Proceedings of Symposia in
PURE MATHEMATICS

Volume 55, Part 2

Motives

Uwe Jannsen
Steven Kleiman
Jean-Pierre Serre
Editors

American Mathematical Society
Providence, Rhode Island

Recent Titles in This Series

55 **Uwe Jannsen, Steven Kleiman, and Jean-Pierre Serre, editors,** Motives (University of Washington, Seattle, July/August 1991)

54 **Robert Greene and S. T. Yau, editors,** Differential geometry (University of California, Los Angeles, July 1990)

53 **James A. Carlson, C. Herbert Clemens, and David R. Morrison, editors,** Complex geometry and Lie theory (Sundance, Utah, May 1989)

52 **Eric Bedford, John P. D'Angelo, Robert E. Greene, and Steven G. Krantz, editors,** Several complex variables and complex geometry (University of California, Santa Cruz, July 1989)

51 **William B. Arveson and Ronald G. Douglas, editors,** Operator theory/operator algebras and applications (University of New Hampshire, July 1988)

50 **James Glimm, John Impagliazzo, and Isadore Singer, editors,** The legacy of John von Neumann (Hofstra University, Hempstead, New York, May/June 1988)

49 **Robert C. Gunning and Leon Ehrenpreis, editors,** Theta functions -- Bowdoin 1987 (Bowdoin College, Brunswick, Maine, July 1987)

48 **R. O. Wells, Jr., editor,** The mathematical heritage of Hermann Weyl (Duke University, Durham, May 1987)

47 **Paul Fong, editor,** The Arcata conference on representations of finite groups (Humboldt State University, Arcata, California, July 1986)

46 **Spencer J. Bloch, editor,** Algebraic geometry – Bowdoin 1985 (Bowdoin College, Brunswick, Maine, July 1985)

45 **Felix E. Browder, editor,** Nonlinear functional analysis and its applications (University of California, Berkeley, July 1983)

44 **William K. Allard and Frederick J. Almgren, Jr., editors,** Geometric measure theory and the calculus of variations (Humboldt State University, Arcata, California, July/August 1984)

43 **François Trèves, editor,** Pseudodifferential operators and applications (University of Notre Dame, Notre Dame, Indiana, April 1984)

42 **Anil Nerode and Richard A. Shore, editors,** Recursion theory (Cornell University, Ithaca, New York, June/July 1982)

41 **Yum-Tong Siu, editor,** Complex analysis of several variables (Madison, Wisconsin, April 1982)

40 **Peter Orlik, editor,** Singularities (Humboldt State University, Arcata, California, July/August 1981)

39 **Felix E. Browder, editor,** The mathematical heritage of Henri Poincaré (Indiana University, Bloomington, April 1980)

38 **Richard V. Kadison, editor,** Operator algebras and applications (Queens University, Kingston, Ontario, July/August 1980)

37 **Bruce Cooperstein and Geoffrey Mason, editors,** The Santa Cruz conference on finite groups (University of California, Santa Cruz, June/July 1979)

36 **Robert Osserman and Alan Weinstein, editors,** Geometry of the Laplace operator (University of Hawaii, Honolulu, March 1979)

35 **Guido Weiss and Stephen Wainger, editors,** Harmonic analysis in Euclidean spaces (Williams College, Williamstown, Massachusetts, July 1978)

34 **D. K. Ray-Chaudhuri, editor,** Relations between combinatorics and other parts of mathematics (Ohio State University, Columbus, March 1978)

33 **A Borel and W. Casselman, editors,** Automorphic forms, representations and L-functions (Oregon State University, Corvallis, July/August 1977)

32 **R. James Milgram, editor,** Algebraic and geometric topology (Stanford University, Stanford, California, August 1976)

(Continued in the back of this publication)

PROCEEDINGS OF THE SUMMER RESEARCH CONFERENCE
ON MOTIVES

HELD AT THE UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON
JULY 20–AUGUST 2, 1991

with the support of the National Science Foundation
Grant DMS–8918200

1991 *Mathematics Subject Classification.*
Primary 14–06; Secondary 11F70, 11G35, 14A20, 19F27.

Library of Congress Cataloging-in-Publication Data

Motives/Uwe Jannsen, Steven L. Kleiman, Jean-Pierre Serre, editors.
p. cm.—(Proceedings of symposia in pure mathematics, ISSN 0082-0717; v. 55)
"Proceedings of the Summer Research Conference on Motives, held at the University of Washington, Seattle, Washington, July 20–August 2, 1991"—T.p. verso.
Includes bibliographical references.
ISBN 0-8218-1635-7 (set: acid-free).—ISBN 0-8218-1636-5 (pt. 1: acid-free).—ISBN 0-8218-1637-3 (pt. 2: acid-free)
1. Motives (Mathematics)—Congresses. I. Jannsen, Uwe. II. Kleiman, Steven L. III. Serre, Jean-Pierre. IV. Summer Research Conference on Motives (1991:University of Washington) V. Series.
QA564.M68 1994 93-38970
516.3′5—dc20 CIP

♾ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
♻ Printed on recycled paper.

Portions of this volume were typeset by the authors using AMS-TeX and AMS-LaTeX,
the American Mathematical Society's TeX macro systems.

10 9 8 7 6 5 4 3 2 1 98 97 96 95 94

Contents

Preface xi

Program xiii

PART 1

The Standard Conjectures
STEVEN L. KLEIMAN 3

Review of ℓ-adic Cohomology
NICHOLAS M. KATZ 21

A Summary of Mixed Hodge Theory
J. H. M. STEENBRINK 31

Crystalline Cohomology
LUC ILLUSIE 43

Conjectures on Algebraic Cycles in ℓ-adic Cohomology
JOHN TATE 71

Some Remarks on the Hodge Type Conjecture
MORIHIKO SAITO 85

Independence of ℓ and Weak Lefschetz
NICHOLAS M. KATZ 101

Décompositions dans la catégorie dérivée
PIERRE DELIGNE 115

Arithmetic Analogs of the Standard Conjectures
H. GILLET AND C. SOULÉ 129

A quoi servent les motifs?
PIERRE DELIGNE 143

Classical Motives
A. J. SCHOLL 163

On the Chow Motive of an Abelian Scheme
KLAUS KÜNNEMANN 189

Weight Filtrations in Algebraic K-Theory
DANIEL R. GRAYSON 207

An Elementary Presentation for K-Groups and Motivic Cohomology
SPENCER BLOCH 239

Motivic Sheaves and Filtrations on Chow Groups
UWE JANNSEN 245

Motivic Complexes
STEPHEN LICHTENBAUM 303

On the Bijectivity of Some Cycle Maps
MORIHIKO SAITO 315

Tannakian Categories
LAWRENCE BREEN 337

Propriétés conjecturales des groupes de Galois motiviques et des représentations ℓ-adiques
JEAN-PIERRE SERRE 377

Motives over Finite Fields
J. S. MILNE 401

Motives for Absolute Hodge Cycles
A. A. PANCHISHKIN 461

CM Motives and the Taniyama Group
NORBERT SCHAPPACHER 485

Structures de Hodge mixtes réelles
PIERRE DELIGNE 509

L-Functions of Mixed Motives
CHRISTOPHER DENINGER 517

L-Functions at the Central Critical Point
BENEDICT H. GROSS 527

Beilinson's Conjectures
JAN NEKOVÁŘ 537

Height Pairings and Special Values of L-Functions
A. J. SCHOLL 571

Autour des conjectures de Bloch et Kato: cohomologie galoisienne et valeurs de fonctions L
JEAN-MARC FONTAINE ET BERNADETTE PERRIN-RIOU 599

Motivic L-Functions and Regularized Determinants
CHRISTOPHER DENINGER 707

On a Result of Deninger Concerning Riemann's Zeta Function
M. SCHRÖTER AND C. SOULÉ 745

PART 2

Classical Polylogarithms
RICHARD M. HAIN 3

Polylogarithms and Motivic Galois Groups
A. B. GONCHAROV 43

Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs
A. BEILINSON ET P. DELIGNE 97

The Elliptic Polylogarithm
A. BEILINSON AND A. LEVIN 123

Iwasawa Theory and p-adic Deformations of Motives
RALPH GREENBERG 193

p-adic Points of Motives
PETER SCHNEIDER 225

Admissible Non-Archimedean Standard Zeta Functions Associated with Siegel Modular Forms
A. A. PANCHISHKIN 251

A p-adic Property of Hodge Classes on Abelian Varieties
DON BLASIUS 293

Drinfeld Modules: Cohomology and Special Functions
DAVID GOSS 309

The Local Langlands Correspondence: The Non-Archimedean Case
STEPHEN S. KUDLA 365

Local Langlands Correspondence: The Archimedean Case
A. W. KNAPP 393

Pure Motives and Automorphic Forms
DINAKAR RAMAKRISHNAN 411

Shimura Varieties and Motives
J. S. MILNE 447

Zeta Functions of Shimura Varieties
DON BLASIUS AND JONATHAN D. ROGAWSKI 525

Hodge-de Rham Structures and Periods of Automorphic Forms
MICHAEL HARRIS 573

Galois Representations Congruent to Those Coming from Shimura Varieties
J. TILOUINE 625

Report on mod ℓ Representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$
KENNETH A. RIBET 639

Preface

The American Mathematical Society, the Institute of Mathematical Statistics, and the Society for Industrial and Applied Mathematics held a joint summer research conference at the University of Washington at Seattle from July 20 to August 2, 1991 on the topic of motives. The conference was organized by Alexander Beilinson (MIT and Moscow), Pierre Deligne (IAS), Uwe Jannsen (Köln), Steven Kleiman (MIT, co-chair), Robert MacPherson (MIT), Jean-Pierre Serre (Collège de France), and Kari Vilonen (Brandeis, co-chair).

The theory of motives was introduced in the middle 1960s by Alexander Grothendieck to explain the analogies among the various cohomology theories for algebraic varieties, to play the role of the missing rational cohomology, and to provide a blueprint for proving Weil's conjectures about the zeta function of a variety over a finite field. Remarkably, over the last ten years or so, researchers in various areas—Hodge theory, algebraic K-theory, polylogarithms, automorphic forms, L-functions, ℓ-adic representations, trigonometric sums, and algebraic cycles—have discovered that an enlarged (and in part conjectural) theory of "mixed" motives indicates and explains phenomena appearing in each area. Thus the theory holds the potential of enriching each area and of unifying them all.

The Seattle conference was the first symposium ever held on motives. It presented a unique opportunity to bring together researchers and students in these diverse areas to exchange ideas and discover common themes. Everyone who applied was invited to attend, and about 140 people from all over the world registered and participated. About a third of the participants were students.

The scientific program ran eleven days. Each day, there were four one-hour lectures; the number was limited to encourage informal discussion. The first lectures introduced and surveyed the entire field; subsequent lectures elaborated on the individual areas. On the last day there was a single one-hour main lecture, followed by six half-hour subsidiary lectures. The lecturers were assigned topics, and were asked to paint panoramic views from their vantage points. A copy of the program is appended.

These volumes contain the proceedings of the conference. They include the revised texts of nearly all the lectures and a number of related works, forty-seven papers in all. There are general introductions, specialized surveys, and research papers. Each paper was refereed and is in final form.

The University of Washington provided a convenient, comfortable, and attractive site, which was conducive to the success of the conference. The AMS did a superb job of administration, freeing the organizing committee to concentrate on the scientific program. In particular, Carole Kohanski, the AMS Conference Coordinator, went far beyond the call of duty. On behalf of the entire organizing committee and all of the participants, the editors wish to express their gratitude to everyone who contributed to the success of the conference and to the production of these proceedings.

Uwe Jannsen
Steven Kleiman
Jean-Pierre Serre

Program

First Week

SUNDAY (Classical motives):
1. Historical introduction (Serre)
2. Standard conjectures (Kleiman)
3. Examples (Scholl)
4. An overview (Deligne)

MONDAY (Cohomology theories):
1. Étale cohomology (Katz)
2. Hodge theory (Steenbrink)
3. Crystalline cohomology (Illusie)
4. The Tate conjectures (Tate)

TUESDAY (Tannakian categories):
1. Tannakian categories and the motivic Galois group (Breen)
2. Motives for absolute Hodge cycles (Panchishkin)
3. CM-motives and the Taniyama group (Schappacher)
4. Motives over finite fields (Milne)

WEDNESDAY (L-functions):
1. Motivic Galois groups (Serre)
2. L-functions (Deninger)
3. The conjectures of Deligne and of Birch/Swinnerton-Dyer (Gross)
4. K-theoretic background (Grayson)

THURSDAY (Beilinson conjectures):
1. Beilinson conjectures I (Soulé)
2. Beilinson conjectures II (Nekovář)
3. Beilinson conjectures III: Reformulation in terms of mixed motives (Scholl)
4. Mixed motives and motivic sheaves (Jannsen)

FRIDAY (Bloch-Kato conjectures, Beilinson-Lichtenbaum complexes):
1. Bloch-Kato conjectures I (Perrin-Riou)
2. Bloch-Kato conjectures II (Fontaine)
3. Beilinson-Lichtenbaum complexes (Lichtenbaum)
4. Higher Chow groups (Bloch)

Second Week

SUNDAY (Mixed Tate motives):

1. Polylogarithms and the line minus three points (Hain)
2. Mixed Tate motives I (MacPherson)
3. Mixed Tate motives II: Zagier's conjecture (Goncharov)
4. Beilinson's work on the Zagier conjecture (Deligne)

MONDAY (Automorphic forms I):

1. The local Langlands conjecture (Kudla)
2. Pure motives and automorphic forms (Ramakrishnan)
3. Shimura varieties and motives (Milne)
4. L-functions of Shimura varieties (Rogawski)

TUESDAY (Automorphic forms II):

1. Hodge-de Rham structures and periods (M. Harris)
2. Mixed motives coming from Shimura varieties (Harder)
3. Galois representations congruent to those arising from Shimura varieties (Tilouine)
4. mod-p Galois representations and Serre's conjectures (Ribet)

WEDNESDAY (p-adic theory, function fields):

1. p-adic L-functions (Coates)
2. Iwasawa theory for motives (Greenberg)
3. p-adic motives (Schneider)
4. Function fields (Goss)

THURSDAY (Miscellaneous topics):

1. Exponential sums (Katz)
2. ℓ-adic representations associated to abelian varieties (Serre)
3. p-adic properties of absolute Hodge cycles (Wintenberger)
4. The motive of an abelian variety (Künnemann)
5. Parshin-Beilinson adèles for schemes (Huber)
6. Hodge modules, questions (M. Saito)
7. F_q-points of a variety and a Hodge-theoretic analogue (Esnault)

Polylogarithms

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Classical Polylogarithms

RICHARD M. HAIN

ABSTRACT. This article is an introduction to classical polylogarithms. After establishing some of their basic properties, we present several examples which originated in the work of Spencer Bloch [8] where the dilogarithm is used to construct the second regulator. We also construct the polylogarithm local systems and show that each underlies a Tate variation of mixed Hodge structure. We conclude by giving an exposition of a motivic description of the polylogarithm local systems.

1. Introduction

Let k be a positive integer. The kth *polylogarithm* $\ln_k x$ is defined by

$$\ln_k x = \sum_{n=1}^{\infty} \frac{x^n}{n^k} . \tag{1}$$

This converges in the unit disk to a holomorphic function. The first polylogarithm, $\ln_1 x$, is just $-\log(1-x)$. The second,

$$\ln_2 x = \sum_{n=1}^{\infty} \frac{x^n}{n^2},$$

is called the *dilogarithm*, and was defined by Leibnitz in a letter to Johann Bernoulli in 1696. The higher polylogarithms were defined by Spence in 1809 (cf. [**31**]).

1991 *Mathematics Subject Classification.* Primary 19-02, 19D55; Secondary 11-02.

Key words and phrases. Dilogarithm, polylogarithm, regulator, Chern class, variation of mixed Hodge structure, K-theory, motive.

Supported in part by grants from the National Science Foundation.

This paper is in final form and no version of it will be submitted for publication elsewhere.

It is generally believed (cf. [5], for example) that the kth regulator

$$c_k : K_m(X) \to H_{\mathscr{D}}^{2k-m}(X, \mathbb{Z}(k))$$

from the algebraic K-theory of a complex algebraic variety X (and therefore all varieties of finite type over $\mathbb{Q}$) to its Deligne cohomology can be expressed in terms of the kth polylogarithm. If true, this would generalize the classical fact that the logarithm occurs as the first Chern class

$$c_1 : K_0(X) \to H^2(X, \mathbb{Z}(1))$$

and its single-valued cousin, $\log|\ | : \mathbb{C}^* \to \mathbb{R}$, occurs as the regulator

$$c_1 : K_1(\mathbb{C}) \approx \mathbb{C}^* \to \mathbb{R} \approx H_{\mathscr{D}}^1(\operatorname{Spec}\mathbb{C}, \mathbb{R}(1)).$$

In the case where X is $\operatorname{Spec}\mathbb{C}$, this should mean that some single-valued cousin $D_k : \mathbb{C} - \{0, 1\} \to \mathbb{R}$ of $\ln_k$ should represent a multiple of the Borel regulator element

$$b_k \in H^{2k-1}(\mathrm{GL}_k(\mathbb{C})^\delta, \mathbb{R}),$$

the cohomology class that gives rise to the regulator [9].[1] The cocycle condition would then be a functional equation satisfied by D_k which generalizes the three-term functional equation satisfied by $D_1 = \log|\ |$.

When X is the spectrum of a number field, Zagier [48] has made a more precise conjecture. He has conjectured that the value of the Dedekind zeta function of a number field F at the positive integer $m > 1$ can be expressed as a certain constant multiple of a determinant of values of D_k at F rational points of $\mathbb{C} - \{0, 1\}$. Such formulas would generalize the formula of Dedekind and Dirichlet for the residue at $s = 1$ of the Dedekind zeta function of F. The case $m = 2$ follows directly from the work of Borel [9] and Suslin [42]. The case $m = 3$ has been proved by Goncharov [21] with the classical trilogarithm, and by Yang [46] with the third higher logarithm of [27].

This, and the results of Suslin, have led Zagier [49] in the case of number fields, and Goncharov [21] for all fields, to conjecture that all of the rational K-theory of a field F should come from $F - \{0, 1\}$, and that the relations among the generators should all correspond to canonical functional equations satisfied by the D_k. For example, for all fields, we may express the familiar fact

$$K_1(F) = F^\times$$

as

$$K_1(F) = \left[\coprod_{x \in F - \{0,1\}} \mathbb{Z}\right] \Big/ \mathscr{R}$$

[1] Here, $\mathrm{GL}_k(\mathbb{C})^\delta$ denotes the general linear group viewed as a discrete group.

where the relations $\mathscr{R}$ are generated by

$$[x] - [xy] + [y] = 0 \quad \text{and} \quad [x] + [x^{-1}] = 0,$$

where $x, y \in F - \{0, 1\}$ and $xy \neq 1$. These are the analogues of the functional equations

$$D_1(x) - D_1(xy) + D_1(y) = 0$$

and

$$D_1(x) + D_1(x^{-1}) = 0.$$

Note also that the functional equation in this case is precisely the condition that D_1 represent an element of $H^1(\mathrm{GL}_1(\mathbb{C}), \mathbb{R})$. The corresponding story for the dilogarithm has been worked out by Bloch [**8**], Dupont and Sah [**20**], and Suslin [**42**]. We give an account of this story in §4.

Most of the results in this paper were discovered by Bloch, Beilinson, Deligne, Dupont, Ramakrishnan, Sah, and Suslin. Their original papers can be consulted for further details and other points of view. Useful references for basic material in this paper include [**34**] and [**22**] for algebraic K-theory, [**31**] for a comprehensive reference on classical aspects of polylogarithms, [**30**] for Tate variations of mixed Hodge structure, and [**24**] for basic facts about iterated integrals and the mixed Hodge theory of the fundamental group. The book [**32**] is a useful reference for other recent developments.

NOTATION. The group of units of a ring R will be denoted by $R^\times$. When Λ is $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$, $\Lambda(k)$ will denote the subgroup $(2\pi i)^k\Lambda$ of $\mathbb{C}$. It will also be used to denote the Hodge structure of type $(-k, -k)$ that has this abelian group as its lattice.

2. Monodromy

An easy power series manipulation yields the formula

$$\ln_k x = \int_0^x \ln_{k-1} z \, \frac{dz}{z},$$

where x lies in the unit disk and $k \geq 2$. It follows, by an induction argument, that each polylogarithm can be analytically continued to a multi-valued holomorphic function on $\mathbb{C} - \{0, 1\}$. In this section we determine the monodromy of the polylogarithms. This was first computed by Ramakrishnan in [**38**].

Set

$$\omega_0 = \frac{dz}{z} \quad \text{and} \quad \omega_1 = \frac{dz}{1 - z}.$$

Let

$$\boldsymbol{\omega} = \begin{pmatrix} 0 & \omega_1 & 0 & \cdots & 0 \\ & \ddots & \omega_0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & \omega_0 \\ 0 & & \cdots & & 0 \end{pmatrix} \in H^0(\mathbb{P}^1, \Omega^1_{\mathbb{P}^1}(\log\{0, 1, \infty\})) \otimes \mathrm{GL}_{n+1}(\mathbb{C}).$$

Consider the first-order linear differential equation

$$d\lambda = \lambda\boldsymbol{\omega} \tag{2}$$

where λ is a possibly multi-valued function $\mathbb{C} - \{0, 1\} \to \mathbb{C}^{n+1}$.

Denote the kth power of the standard logarithm $\log x = \int_1^x \omega_0$ by $\log^k x$. Let

$$\Lambda(x) = \begin{pmatrix} 1 & \ln_1 x & \ln_2 x & \cdots & \cdots & \ln_n x \\ 0 & 2\pi i & 2\pi i \log x & & \cdots & \frac{2\pi i}{(n-1)!}\log^{n-1} x \\ 0 & 0 & (2\pi i)^2 & \ddots & \cdots & \frac{(2\pi i)^2}{(n-2)!}\log^{n-2} x \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & 0 & (2\pi i)^{n-1} & (2\pi i)^{n-1}\log x \\ 0 & \cdots & \cdots & \cdots & 0 & (2\pi i)^n \end{pmatrix} \in \mathrm{GL}_{n+1}(\mathbb{C}).$$

More precisely,

$$\Lambda_{jk}(x) = \begin{cases} \ln_k x & \text{when } j = 0 \text{ and } k > 0, \\ \frac{(2\pi i)^j}{(k-j)!}\log^{k-j} x & \text{when } 0 < j \le k, \\ 0 & \text{when } k < j. \end{cases}$$

We will view this as a multi-valued $\mathrm{GL}_{n+1}(\mathbb{C})$-valued function on $\mathbb{C} - \{0, 1\}$. By the *principal branch* of $\Lambda(x)$ we shall mean the matrix-valued function on the disk $|x - \frac{1}{2}| < \frac{1}{2}$ obtained by taking the standard branches of each of its entries on that disk. (The principal branch of $\ln_k$ on this disk is the one given by the power series expansion (1).)

PROPOSITION 2.1. *The function* $\Lambda(x)$ *is a fundamental solution of* (2). *That is,*

$$d\Lambda = \Lambda\,\boldsymbol{\omega}$$

and $\Lambda(x)$ *is nonsingular for each* $x \in \mathbb{C} - \{0, 1\}$. □

If we analytically continue the principal branch of $\Lambda(x)$ about a loop in $\mathbb{C} - \{0, 1\}$ based at $\frac{1}{2}$, the resulting matrix of functions will still be a fundamental solution of (2). It follows that, for each loop γ based at $\frac{1}{2}$, there is a matrix $M(\gamma) \in \mathrm{GL}_{n+1}(\mathbb{C})$ such that the analytic continuation of (the principal branch of) $\Lambda(x)$ about γ is $M(\gamma)\Lambda(x)$. For a pair of loops α, β based at $1/2$, we have

$$M(\alpha\beta) = M(\alpha)M(\beta).$$

Since the value of $M(\gamma)$ depends only on the homotopy class of γ, we obtain a monodromy representation

$$M : \pi_1(\mathbb{C} - \{0, 1\}, 1/2) \to \mathrm{GL}_{n+1}(\mathbb{C}). \tag{3}$$

Let $\sigma_0, \sigma_1 \in \pi_1(\mathbb{C} - \{0, 1\}, 1/2)$ be the loops defined by

$$\sigma_0(t) = e^{2\pi i t}/2, \quad \sigma_1(t) = 1 - e^{2\pi i t}/2, \quad 0 \le t \le 1.$$

These loops generate $\pi_1(\mathbb{C} - \{0, 1\}, 1/2)$.

PROPOSITION 2.2 **[38]**. *We have*

$$M(\sigma_0) = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & J & \\ 0 & & & \end{array}\right) \quad \textit{and} \quad M(\sigma_1) = \left(\begin{array}{c|ccccc} 1 & -1 & 0 & \cdots & 0 \\ \hline 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{array}\right)$$

where

$$J = \exp \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & & \cdots & & 0 \end{pmatrix}.$$

PROOF. The monodromy around σ_0 is easy to calculate. Indeed, since the principal branch of each polylogarithm is single-valued on the disk $|z - \frac{1}{2}| \le \frac{1}{2}$, each is unchanged when continued along σ_0. The formula for $M(\sigma_0)$ follows since the analytic continuation of $\log x$ about σ_0 is $\log x + 2\pi i$.

Since the principal branch of $\log x$ is defined in a neighbourhood of $\mathbb{R}_+$, it follows that $\log x$ is invariant under analytic continuation along σ_1.

We compute the analytic continuation of $\ln_k$ along σ_1 by induction. When $k = 1$, $\ln_1 x$ changes to $\ln_1 x - 2\pi i$. Assume now that $k \ge 1$ and that the continuation of $\ln_k x$ along σ_1 is

$$\ln_k x - \frac{2\pi i}{(k-1)!} \log^{k-1} x.$$

Denote the integral of $f(z)\,dz$ along the straight line interval in the complex plane between the points a and b by

$$\int_a^b f(z)\,dz.$$

When $|x - \frac{1}{2}| < \frac{1}{2}$, we have

$$\begin{aligned} \ln_{k+1} x &= \int_0^x \ln_k z \frac{dz}{z} \\ &= \int_0^{1/2} \ln_k z \frac{dz}{z} + \int_{1/2}^x \ln_k z \frac{dz}{z}. \end{aligned} \tag{4}$$

The result of analytically continuing this along σ_1 is

$$\int_0^{1/2} \ln_k z \frac{dz}{z} + \int_{\sigma_1} \ln_k z \frac{dz}{z} + \int_{1/2}^x \ln_k z \frac{dz}{z}. \tag{5}$$

It follows from the inductive formula that the difference between (4) and (5) is

$$\int_{\sigma_1} \ln_k z \frac{dz}{z} - \frac{2\pi i}{(k-1)!} \int_{1/2}^x \log^{k-1} z \frac{dz}{z}. \tag{6}$$

For each $\varepsilon \in (0, \frac{1}{2})$, the path that traverses the line segment from $\frac{1}{2}$ to $1 - \varepsilon$, goes around the boundary of the disk $|z - 1| \leq \varepsilon$ in the positive direction, then returns along the interval from $1 - \varepsilon$ to $\frac{1}{2}$ represents the homotopy class σ_1. Again, using the inductive formula for the monodromy of $\ln_k x$ around σ_1, we have

$$\int_{\sigma_1} \ln_k z \frac{dz}{z} = -\frac{2\pi i}{(k-1)!} \int_{1-\varepsilon}^{1/2} \log^{k-1} z \frac{dz}{z} + \int_{-\pi}^{\pi} \ln_k(1 + \varepsilon e^{it}) \frac{d(\varepsilon e^{it})}{1 + \varepsilon e^{it}}.$$

The inductive hypothesis also implies that $\ln_k x$ is bounded in a neighbourhood of 1 when $k > 1$, so that the last integral $\to 0$ as $\varepsilon \to 0$ for all k. (A separate argument is needed when $k = 1$.) Combining this with (6), we see that $\ln_{k+1} x$ changes by

$$-\frac{2\pi i}{(k-1)!} \left[\lim_{\varepsilon \to 0} \int_{1-\varepsilon}^{1/2} \log^{k-1} z \frac{dz}{z} + \int_{1/2}^x \log^{k-1} z \frac{dz}{z} \right] = -\frac{2\pi i}{k!} \log^k x$$

when continued around σ_1. □

The monodromy calculation has several interesting consequences. First, even though it does not make sense, in general, to talk about the value of a multi-valued function at a point, it does make sense to talk about the value of $\ln_k$ at 1.

COROLLARY 2.3. *The value of* $(k-1)! \ln_k 1$ *is well defined modulo* $\mathbb{Z}(k)$ *and is congruent to* $(k-1)!\zeta(k)$. □

The second important consequence of the monodromy calculation is the rationality of the monodromy.

COROLLARY 2.4. *The image of the monodromy representation* (3) *is contained in* $GL_{n+1}(\mathbb{Q})$. □

The significance of this last result is that it implies that the local system over $\mathbb{C}-\{0, 1\}$ which corresponds to the differential equation (2) is defined over $\mathbb{Q}$. This local system is called the *nth polylogarithm local system.* These local systems fit together to form an inverse system of local systems whose limit we call the *polylogarithm local system.* We now describe these local systems in detail.

Define a meromorphic connection ∇ on the trivial bundle

$$\mathbb{P}^1 \times \mathbb{C}^{n+1} \to \mathbb{P}^1 \tag{7}$$

by defining

$$\nabla f = df - f\,\boldsymbol{\omega}$$

where $f : \mathbb{C}-\{0, 1\} \to \mathbb{C}^{n+1}$ is a section. This connection has regular singular points at 0, 1, and ∞, and is flat over $\mathbb{C}-\{0, 1\}$ as $\boldsymbol{\omega}$ satisfies the integrability condition

$$d\,\boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega} = 0$$

(equivalently, because the equation $\nabla f = 0$ is a system of ordinary differential equations). Let $\lambda_0, \lambda_1, \dots, \lambda_n$ be the rows of $\Lambda(x)$. Each of these satisfies (2) and is therefore a flat section of (7). Even though these are multivalued, their $\mathbb{Q}$-linear span is well defined as the monodromy representation is defined over $\mathbb{Q}$.

Suppose that X is a smooth curve and that $\overline{X}$ is a smooth completion of X. Every flat bundle $E \to X$ has a *canonical extension* $\overline{E} \to \overline{X}$. Denote the local monodromy operator about a point $p \in D := \overline{X} - X$ by T_p. When each T_p is unipotent, the canonical extension is characterized by two properties. First, the meromorphic extension $\overline{\nabla}$ of the connection to $\overline{E} \to \overline{X}$ has logarithmic singularities along D. That is,

$$\overline{\nabla} : \overline{\mathscr{E}} \to \overline{\mathscr{E}} \otimes_{\mathscr{O}_{\overline{X}}} \Omega^1_{\overline{X}}(\log D).$$

Second, the residue of $\overline{\nabla}$ at each point of D is nilpotent.

Denote the $\mathbb{Q}$-local system over $\mathbb{C}-\{0, 1\}$ which corresponds to the representation (3) by $\mathbb{V}$.

Since $\boldsymbol{\omega}$ has nilpotent residue at 0, 1, and ∞, we have

PROPOSITION 2.5. *The canonical extension of the flat holomorphic vector bundle* $\mathbb{V} \otimes_{\mathbb{Q}} \mathscr{O}_{\mathbb{C}-\{0,1\}}$ *to* $\mathbb{P}^1$ *is the bundle* (7) *with the connection* ∇ *defined above.* □

3. The Bloch-Wigner function

Define

$$D_2(x) = \Im \ln_2 x + \log|x| \arg(1-x)$$

when $|x - \frac{1}{2}| < \frac{1}{2}$ and where $\ln_2 x$, $\log x$, and $\arg(1-x)$ denote the principal branches of these functions in the disk $|x - \frac{1}{2}| < \frac{1}{2}$. An easy computation using the monodromy calculation (2.2) shows that D_2 is invariant under continuation along the generators σ_0 and σ_1 of $\pi_1(\mathbb{C} - \{0, 1\}, \frac{1}{2})$. Consequently, the function D_2 extends to a single-valued, real analytic function

$$D_2 : \mathbb{C} - \{0, 1\} \to \mathbb{R}.$$

This is called the *Bloch-Wigner function.* If we define

$$D_2(0) = D_2(1) = D_2(\infty) = 0,$$

then D_2 is a continuous function $D_2 : \mathbb{P}^1 \to \mathbb{R}$. The Bloch-Wigner function should be viewed as having the same relation to $\ln_2$ as $D_1 := \log|\ |$ bears to the logarithm.

The boundary of hyperbolic 3-space $\mathbb{H}^3$ is the Riemann sphere $\mathbb{P}^1$. The group of orientation-preserving isometries of $\mathbb{H}^3$ is $\mathrm{PSL}_2(\mathbb{C})$. The induced action on the boundary is just the standard action of $\mathrm{PSL}_2(\mathbb{C})$ on $\mathbb{P}^1$ via fractional linear transformations.

Denote by $\Delta(a_0, a_1, a_2, a_3)$ the ideal tetrahedron in $\mathbb{H}^3$ with vertices at $a_0, a_1, a_2, a_3 \in \mathbb{P}^1$. Since the volume form of hyperbolic space is invariant under the action of the isometry group,

$$\operatorname{vol}\Delta(a_0, a_1, a_2, a_3) = \operatorname{vol}\Delta(\lambda, 1, 0, \infty),$$

where λ is the cross ratio $[a_0 : a_1 : a_2 : a_3]$ of the vertices.

The following result goes back to Lobachevsky (cf. [**35**]). A proof may be found in [**20**, p. 172].

THEOREM 3.1. *For each* $z \in \mathbb{P}^1$, *the volume of* $\Delta(z, 1, 0, \infty)$ *equals* $D_2(z)$. □

COROLLARY 3.2. *If* $a_0, a_1, a_2, a_3, a_4 \in \mathbb{P}^1$, *then*

$$\sum_{j=0}^{4}(-1)^j D_2([a_0 : \cdots : \widehat{a_j} : \cdots : a_4]) = 0.$$

Moreover, for all permutations σ *of* $\{0, 1, 2, 3\}$,

$$D_2([a_{\sigma(0)} : \cdots : a_{\sigma(3)}]) = \operatorname{sgn}(\sigma) D_2([a_0 : \cdots : a_3]).$$

PROOF. The first assertion follows because the ideal polyhedron P with vertices a_0, a_1, a_2, a_3, a_4 decomposes into a union of ideal tetrahedra in two different ways. Namely,

$$P = \Delta(a_1, a_2, a_3, a_4) \cup \Delta(a_0, a_1, a_3, a_4) \cup \Delta(a_0, a_1, a_2, a_3)$$

and

$$P = \Delta(a_0, a_2, a_3, a_4) \cup \Delta(a_0, a_1, a_2, a_4).$$

In each case the pieces intersect along two-dimensional faces. The first assertion follows from Theorem 3.1 by comparing volumes. The second follows because swapping the order of two vertices reverses the orientation of the tetrahedron. □

Taking the five points to be $y, x, 1, 0, \infty$, we obtain the usual form of the functional equation, which is the analogue for D_2 of the Abel-Spence functional equation of $\ln_2$.

COROLLARY 3.3. *If* $y, x, 1, 0, \infty$ *are distinct points of* $\mathbb{P}^1$, *then*

$$D_2(x) - D_2(y) + D_2\left(\frac{y}{x}\right) - D_2\left(\frac{1-y}{1-x}\right) + D_2\left(\frac{1-y^{-1}}{1-x^{-1}}\right) = 0. \quad \square$$

Ramakrishnan [39] showed that all the polylogarithms have such single-valued cousins, the essential point being that one can use the unipotence of the monodromy group and induction to kill off the monodromy of $\ln_k$. Zagier [48], and independently Wojtkowiak [44], gave explicit formulas for such single-valued functions. In fact, there are two canonical ways to produce single-valued cousins of the classical polylogarithms. These are the entries in the top row of the single-valued matrix-valued functions

$$P(x) := \overline{\Lambda(x)^{-1}}\Lambda(x), \quad \log P(x)$$

on $\mathbb{P}^1(\mathbb{C}) - \{0, 1, \infty\}$ (cf. [49, p. 415]). Each of these single-valued cousins is continuous on $\mathbb{P}^1$, and real analytic on $\mathbb{P}^1 - \{0, 1, \infty\}$. The preferred single-valued cousin of the trilogarithm is the function

$$D_3(x) = \Re\left[\ln_3(x) - \log|x| \ln_2 x + (\log^2|x| \ln_1 x)/3\right]. \tag{8}$$

4. The regulator $K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{Q}(2)$

The Deligne cohomology of $\operatorname{Spec}\mathbb{C}$ is

$$H^m_{\mathscr{D}}(\operatorname{Spec}\mathbb{C}, \Lambda(k)) = \begin{cases} 0, & m \neq 1, \\ \mathbb{C}/\Lambda(k), & m = 1, \end{cases}$$

where Λ denotes $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$. So the only nontrivial regulators for $\operatorname{Spec}\mathbb{C}$ with values in Deligne cohomology are

$$K_{2m-1}(\mathbb{C}) \to H^1_{\mathscr{D}}(\operatorname{Spec}\mathbb{C}, \Lambda(m)) \approx \mathbb{C}/\Lambda(m)$$

for each $m \geq 1$. The first regulators

$$K_1(\mathbb{C}) \approx \mathbb{C}^* \to \mathbb{C}/\mathbb{Z}(1) \quad \text{and} \quad K_1(\mathbb{C}) \approx \mathbb{C}^* \to \mathbb{C}/\mathbb{R}(1) \approx \mathbb{R}$$

are given by the functions log and log| |, respectively. In this section we give an account of the construction of the second regulators

$$K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{Q}(2) \quad \text{and} \quad K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{R}(2) \approx \mathbb{R}(1)$$

using $\ln_2$ and D_2 respectively. These results go back to Bloch and Wigner (unpublished, cf. [20]). As is customary, $GL_n(\mathbb{C})^\delta$ signifies that $GL_n(\mathbb{C})$ is viewed as a group with the discrete topology.

LEMMA 4.1. *The Bloch-Wigner function defines a canonical group cohomology class*

$$\mathscr{D}_2 \in H^3(GL_2(\mathbb{C})^\delta, \mathbb{R}).$$

PROOF. Let F be a field and $n \in \mathbb{N}$. For each $k \in \mathbb{N}$, define $C_k(F, n)$ to be the free abelian group with basis the set of ordered $(k+1)$-tuples $(v_0, \ldots, v_k)$ of vectors v_j of F^n in general position. (That is, each $\min(n, k+1)$ of them are independent.) Define a differential $\partial : C_k \to C_{k-1}$ by

$$\partial : (v_0, \ldots, v_k) \mapsto \sum_{j=0}^{k} (-1)^j (v_0, \ldots, \widehat{v_j}, \ldots, v_k).$$

When F is infinite, the complex $C_\bullet(F, n)$ is quasi-isomorphic to $\mathbb{Z}$. Since $GL_n(F)$ acts on this complex, it is a resolution of the trivial module. So, for all $GL_n(F)$ modules M, there is a natural map

$$H^\bullet\left(\mathrm{Hom}_{GL_n(F)}(C_\bullet(F, n), M)\right) \to H^\bullet(GL_n(F), M),$$

provided that F is infinite.

Denote the point in $\mathbb{P}^{n-1}(F)$ determined by the nonzero vector v of F^n by $[v]$. Define a map $f : C_3(\mathbb{C}, 2) \to \mathbb{R}$ by

$$f(v_0, v_1, v_2, v_3) = D_2\left([[v_0] : [v_1] : [v_2] : [v_3]]\right).$$

The cocycle condition is simply the functional equation (3.2). Denote the image of this cohomology class in $H^3(GL_2(\mathbb{C})^\delta, \mathbb{R})$ by $\mathscr{D}_2$. □

A cohomology class $c \in H^m(GL(R), \Lambda)$ defines a map $K_m(R) \to \Lambda$. This map is obtained as the composite

$$K_m(R) = \pi_m(BGL(R)^+) \to H_m(BGL(R)^+) \approx H_m(GL(R)) \xrightarrow{c} \Lambda$$

of the Hurewicz homomorphism with the Λ-valued functional on homology induced by c.

THEOREM 4.2. *The function D_2 can be used to define a map $K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{R}(2)$ which equals the Beilinson regulator, or equivalently, half the Borel regulator.*

PROOF. The proof is an assemblage of results from the literature, and we only give a sketch. First, the cocycle $GL_2(\mathbb{C})^4 \to \mathbb{C}/\mathbb{R}(2)$ induced by iD_2 composed with the cross ratio, represents a continuous cohomology class $i\mathfrak{D}_2$, even though this cocycle is not itself continuous. This apparently contradictory state of affairs arises because the cross ratio $\left(\mathbb{P}^1\right)^4 \to \mathbb{P}^1$ is not everywhere defined. This problem is rectified as follows: the cocycle iD_2

is locally L_2 as a function $\mathrm{GL}_2(\mathbb{C})^4 \to \mathbb{C}/\mathbb{R}(2)$. It therefore determines a class in the locally L_2 cohomology group $H^3_{\mathrm{loc}L_2}(\mathrm{GL}_2(\mathbb{C}, \mathbb{C}/\mathbb{R}(2)))$. Since the natural map from continuous cohomology to locally L_2 cohomology is an isomorphism **[6]**, the dilogarithm iD_2 determines a continuous cohomology class $i\mathfrak{D}_2$. Since locally L_2 cocycles are represented by cocycles that can be changed on sets of measure zero without affecting them, and since elements of $H_3(\mathrm{GL}_2(\mathbb{C}))$ are represented by cycles whose support has measure zero, it is not immediately clear that the image of the class $i\mathfrak{D}_2$ in $H^3(\mathrm{GL}_2(\mathbb{C}), \mathbb{C}/\mathbb{R}(2))$ under the map

$$H^3_{\mathrm{loc}L_2}(\mathrm{GL}_2(\mathbb{C}, \mathbb{C}/\mathbb{R}(2))) \stackrel{\approx}{\leftarrow} H^3_{\mathrm{cts}}(\mathrm{GL}_2(\mathbb{C}, \mathbb{C}/\mathbb{R}(2))) \to H^3(\mathrm{GL}_2(\mathbb{C}, \mathbb{C}/\mathbb{R}(2)))$$

is the class $i\mathscr{D}_2$ of (4.1). This is the case; a proof may be found in **[46]** or **[28]**.

Next, since

$$H^3_{\mathrm{cts}}(\mathrm{GL}_2(\mathbb{C}), \mathbb{C}/\mathbb{R}(2)) \approx H^3_{\mathrm{cts}}(\mathrm{GL}_n(\mathbb{C}), \mathbb{C}/\mathbb{R}(2)) \approx \mathbb{C}/\mathbb{R}(2)$$

for all $n \geq 2$, and since this group is spanned by the continuous cohomology class β_2 used to define the Borel regulator (cf. **[40, 19]**), it follows that there is a real number λ such that

$$i\mathfrak{D}_2 = \lambda\beta_2/2.$$

Denote by c_k the kth Beilinson-Chern class of the universal flat bundle over $\mathrm{BGL}_n(\mathbb{C})^\delta$. This is an element of

$$H^{2k}_{\mathscr{D}}(\mathrm{BGL}_n(\mathbb{C})^\delta, \mathbb{R}(k)) \approx H^{2k-1}(\mathrm{GL}_n(\mathbb{C}), \mathbb{C}/\mathbb{R}(k)).$$

In **[19]** it is shown that, for all n and k, the classes c_k and $\beta_k/2$ in

$$H^k(\mathrm{GL}_n(\mathbb{C})^\delta, \mathbb{C}/\mathbb{R}(2))$$

are equal. It follows that $i\mathfrak{D}_2$ is λ times the Beilinson-Chern class

$$c_2 : K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{R}(2).$$

Denote the rational K-theory $K_\bullet(R) \otimes \mathbb{Q}$ of a ring R by $K_\bullet(R)_\mathbb{Q}$. Since $\mathrm{BGL}(R)^+$ is an H-space, the Hurewicz homomorphism

$$K_m(R)_\mathbb{Q} := \pi_m(\mathrm{BGL}(R)^+) \otimes \mathbb{Q} \to H_m(\mathrm{BGL}(R)^+, \mathbb{Q}) \approx H_m(\mathrm{GL}(R), \mathbb{Q})$$

is injective. Define the *rank filtration*

$$r_1K_m(R) \subseteq r_2K_m(R) \subseteq r_3K_m(R) \subseteq \cdots \subseteq K_m(R)_\mathbb{Q}$$

of $K_m(R)_\mathbb{Q}$ by

$$r_kK_m(R) = K_m(R)_\mathbb{Q} \cap \mathrm{im}\{H_m(\mathrm{GL}_k(R), \mathbb{Q}) \to H_m(\mathrm{GL}(R))\}.$$

Suslin [42] has proved that, for all infinite fields F,

$$K_m(F)_{\mathbb{Q}} = r_m K_m(F)$$

and

$$K_m(F)_{\mathbb{Q}}/r_{m-1}K_m(F) \approx K_m^M(F) \otimes \mathbb{Q},$$

where $K_\bullet^M(F)$ denotes the Milnor K-theory of F.

In particular, for all infinite fields F, there is a canonical isomorphism

$$K_3(F)_{\mathbb{Q}} = K_3^M(F)_{\mathbb{Q}} \oplus r_2 K_3(F).$$

Since all elements of K_3^M of any field are decomposable, the generalization of the Whitney sum formula to K-theory implies that $c_2 : K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{R}(2)$ vanishes on $K_3^M(\mathbb{C})$. It follows that c_2 factors through the projection onto $r_2K_3(\mathbb{C})$.

$$\begin{array}{ccc} K_3(\mathbb{C}) & \xrightarrow{c_2} & \mathbb{C}/\mathbb{R}(2) \\ \text{proj} \downarrow & \nearrow & \\ r_2K_3(\mathbb{C}) & & \end{array}$$

Since $r_2K_3(\mathbb{C})$ comes from $H_3(\mathrm{GL}_2(\mathbb{C}))$, the restriction of c_2 to $r_2K_3(\mathbb{C})$ is given by the class $i\lambda^{-1}\mathfrak{D}_2$. Now Dupont [18] has proved that $i\mathfrak{D}_2$ equals the Cheeger-Simons class

$$\hat{c}_2 \in H^3(\mathrm{GL}_2(\mathbb{C})^\delta, \mathbb{C}/\mathbb{R}(2))$$

of the universal rank 2 flat bundle. But by the main result of [19], this class equals the Beilinson-Chern class of the universal flat bundle. It follows that $\lambda = 1$. □

Much of this story has been extended to the trilogarithm by Goncharov [21] and Yang [46]. In [27] the existence of a canonical single-valued trilogarithm

$$S_3 : \left\{ \begin{array}{c} \text{ordered 6-tuples of points} \\ \text{in } \mathbb{P}^2, \text{ no 3 on a line} \end{array} \right\} \Big/ \text{projective equivalence} \to \mathbb{C}/\mathbb{R}(3) \approx \mathbb{R}$$

was established. It satisfies the seven-term functional equation

$$\sum_{j=0}^{6} (-1)^j S_3(a_0, \dots, \widehat{a_j}, \dots, a_6) = 0$$

where $a_0, \dots, a_6$ are points in $\mathbb{P}^2$, no three of which lie on a line. This equation is an obvious generalization of the five-term equation satisfied by D_2. As in (4.1), this function determines a class in $H^5(\mathrm{GL}_3(\mathbb{C}), \mathbb{R})$. As in the case of the dilogarithm, the cocycle condition is precisely the functional equation. Goncharov has given an explicit construction of this function in terms of the single-valued classical trilogarithm (8). Goncharov [21] and Yang [46] have both shown that this class in $H^5(\mathrm{GL}_3(\mathbb{C}), \mathbb{R})$ is a nonzero rational multiple of the class that corresponds to the Beilinson-Chern class of the universal flat bundle over $BGL_3(\mathbb{C})^\delta$. This multiple has been determined; in [28] it is shown that S_3 represents $\frac{1}{6}$ the Beilinson-Chern class, and $\frac{1}{12}$ *the class of the Borel regulator. The appropriate analogue of Suslin's theorem* is not known for $K_5(\mathbb{C})$. However, Yang [47] has proved the rank conjecture for all K-groups of all number fields except $\mathbb{Q}$.

The regulator $c_2 : K_3(\mathbb{C}) \to \mathbb{C}/\mathbb{Q}(2)$ can be written in terms of the multivalued dilogarithm. This work goes back to unpublished work of Bloch and Wigner (cf. [20]). The construction is similar to, but more complicated than, the construction of the regulator given above. We only sketch the construction. More details can be found in [20].

For x in the disk $|x - \frac{1}{2}| < \frac{1}{2}$, define

$$\rho(x) = \frac{1}{2}\left[\log x \wedge \log(1-x) + 2\pi i \wedge \frac{1}{2\pi i}\left(\ln_2(1-x) - \ln_2(x) - \frac{\pi^2}{6}\right)\right] \in \Lambda^2_{\mathbb{Z}}\mathbb{C}.$$

Here all functions are taken to be the principal branches. This function is invariant under monodromy, and therefore extends to a single-valued function

$$\rho : \mathbb{C} - \{0, 1\} \to \Lambda^2_{\mathbb{Z}}\mathbb{C}.$$

It satisfies a generalization of the five-term equation satisfied by D_2. If $x, y \in \mathbb{C} - \{0, 1\}$ and $x \neq y$, then

$$\rho(x) - \rho(y) + \rho(y/x) - \rho((1-y)/(1-x)) + \rho((1-y^{-1})/(1-x^{-1})) = 0.$$

Define $\mathscr{P}(F)$ to be the free abelian group generated by $F - \{0, 1\}$ subject to the relations

$$[x] - [y] + [y/x] - [(1-y)/(1-x)] + [(1-y^{-1})/(1-x^{-1})] = 0.$$

This is often called the *scissors congruence group*.[2] The function ρ induces a map $\rho : \mathscr{P}(\mathbb{C}) \to \Lambda^2_{\mathbb{Z}}\mathbb{C}$.

There is a natural homomorphism $H_3(SL_2(F)) \to \mathscr{P}(F)$ whose construction is similar to that of the homomorphism

$$H_3(GL_2(F)) \to H_\bullet(C_\bullet(F, n))$$

given in the proof of (4.1). If F is algebraically closed of characteristic 0, then there is an exact sequence

$$0 \to \mathbb{Q}/\mathbb{Z} \to H_3(SL_2(F)) \to \mathscr{P}(F) \to \Lambda^2_{\mathbb{Z}}F^\times \to K_2(F) \to 0.$$

[2]It is also called the *Bloch group* by some, although this term is ambiguous since others use this term to refer to the kernel of the map $\mathscr{P}(F) \to \Lambda^2 F^\times$.

A proof can be found in the appendix of **[20]**. The map $\mathscr{P}(F) \to \Lambda^2_{\mathbb{Z}} F^\times$ takes the generator $[x]$ of $\mathscr{P}(F)$ to $(1-x) \wedge x$. The most right-hand map takes $x \wedge y$ to $\{x, y\}$.

The kernel of the map $\Lambda^2 \exp : \Lambda^2_{\mathbb{Z}}\mathbb{C} \to \Lambda^2_{\mathbb{Z}}\mathbb{C}^*$ is $\mathbb{C}/\mathbb{Q}(2)$, where this is included in $\Lambda^2_{\mathbb{Z}}\mathbb{C}$ by taking the coset of λ to $2\pi i \wedge (\lambda/2\pi i)$. Since the diagram

$$\begin{array}{ccc} \mathscr{P}(\mathbb{C}) & \to & \Lambda^2_{\mathbb{Z}}\mathbb{C}^\times \\ \rho \downarrow & & \downarrow \mathrm{id} \\ \Lambda^2_{\mathbb{Z}}\mathbb{C} & \xrightarrow{\Lambda^2 \exp} & \Lambda^2_{\mathbb{Z}}\mathbb{C}^* \end{array}$$

commutes, ρ induces a homomorphism $H_3(\mathrm{SL}_2(\mathbb{C})) \to \mathbb{C}/\mathbb{Q}(2)$. By **[18]**, this represents the second Cheeger-Simons class $\hat{c}_2$ of the universal flat bundle over $\mathrm{BSL}_2(\mathbb{C})^\delta$. By the main result of **[19]**, this equals the Beilinson-Chern class of the universal flat bundle over $\mathrm{BSL}_2(\mathbb{C})^\delta$. As above, c_2 vanishes on $K_3^M(\mathbb{C})$ and the diagram

$$\begin{array}{ccc} K_3(\mathbb{C}) & \xrightarrow{c_2} & \mathbb{C}/\mathbb{Q}(2) \\ \mathrm{proj} \downarrow & \nearrow_{\rho} & \\ r_2 K_3(\mathbb{C}) & & \end{array}$$

commutes. One can show that the map $H_3(\mathrm{SL}_2(F), \mathbb{Q}) \to r_2K_3(F)$ is surjective. It follows that the map constructed above induces the regulator on all of $K_3(\mathbb{C})$. Alternatively, one can appeal to the theorem of Suslin **[43]** which asserts that there are natural isomorphisms

$$\begin{aligned} H_3(\mathrm{SL}_2(F), \mathbb{Q}) &\approx r_2K_3(F) \\ &\approx \ker\left\{\mathscr{P}(F) \to \Lambda^2_{\mathbb{Z}}F^\times\right\} \otimes \mathbb{Q} \end{aligned}$$

for all fields F. This last isomorphism says that all of the weight 2 part of K_3 of a field comes from $\mathbb{P}^1 - \{0, 1, \infty\}$ and that all the relations come from the functional equation of the dilogarithm. This generalizes the fact, mentioned in the introduction, that the relations in K_1 come from the functional equation of the logarithm. The analogue of this statement for the weight 3 part of K_5 is not known at this time, although Goncharov **[21]** has made significant progress.

5. Iterated integrals

At this stage it is convenient to introduce Chen's iterated integrals **[12]**. Basic references for this section are **[12, 24]**.

Suppose that M is a manifold and that $w_1 \cdots w_r$ are smooth $\mathbb{C}$-valued 1-forms on M. For each piecewise smooth path $\gamma : [0, 1] \to M$, we can define

$$\int_\gamma w_1 \cdots w_r = \int \cdots \int_{0 \le t_1 \le \cdots \le t_r \le 1} f_1(t_1) \cdots f_r(t_r)\, dt_1 \cdots dt_r,$$

where $\gamma^* w_j = f_j(t)\, dt$ for each j. This can be viewed as a $\mathbb{C}$-valued function

$$\int_\gamma w_1 \cdots w_r : PM \to \mathbb{C}$$

on the space of piecewise smooth paths in M. When $r = 1$, $\int_\gamma w$ is just the usual line integral. An *iterated integral* is any function $PM \to \mathbb{C}$ that is a linear combination of a constant function and basic iterated integrals

$$\int_\gamma w_1 \cdots w_r.$$

Now let $M = \mathbb{C} - \{0, 1\}$ and

$$\omega_0 = \frac{dz}{z} \quad \text{and} \quad \omega_1 = \frac{dz}{1-z}.$$

Then

$$\ln_1 x = -\log(1-x) = \int_0^x \omega_1.$$

By induction and the definition, we have, for all $k \ge 2$,

$$\ln_k x = \int_0^x \ln_{k-1} z\, \omega_0 = \int_0^x \omega_1 \overbrace{\omega_0 \cdots \omega_0}^{k-1 \text{ times}}.$$

Here the path of integration must be chosen so that once it has left 0, it never passes through 0 or 1 on its way to x.

The basic properties of iterated integrals are summarized in the following proposition.

PROPOSITION 5.1 **[12, 24]**. *Suppose that $w_1, w_2, \ldots$ are $\mathbb{C}$-valued 1-forms on a manifold M.*

(i) *The value of $\int_\gamma w_1 \cdots w_r$ is independent of the parameterization of γ.*

(ii) *If $\alpha, \beta : [0, 1] \to M$ are composable paths (i.e., $\alpha(1) = \beta(0)$), then*

$$\int_{\alpha\beta} w_1 \cdots w_r = \sum_{j=0}^{r} \int_\alpha w_1 \cdots w_i \int_\beta w_{i+1} \cdots w_r.$$

Here, $\int_\gamma \phi_1 \cdots \phi_m$ is to be interpreted as 1 when $m = 0$.

(iii) *For all paths γ,*

$$\int_{\gamma^{-1}} w_1 \cdots w_r = (-1)^r \int_\gamma w_r \cdots w_1.$$

(iv) *For all paths* α *in* M,

$$\int_\alpha w_1 \cdots w_r \int_\alpha w_{r+1} \cdots w_{r+s} = \sum_\sigma \int_\alpha w_{\sigma(1)} \cdots w_{\sigma(r+s)},$$

where σ *ranges over all shuffles of type* (r, s).

6. The regulator $K_2(X) \to H^1(X, \mathbb{C}^*)$

This section is an exposition of the Beilinson-Deligne construction of the regulator

$$c_2 : K_2(X) \to H^2_{\mathscr{D}}(X, \mathbb{Z}(2))$$

using the dilogarithm (cf. **[2]** and **[15]**). This construction generalizes an earlier construction of Bloch **[8]** in the case when X is an elliptic curve. We have freely incorporated the elegant approaches of Deligne **[16]** and Ramakrishnan **[37, 40]**.

When X is a curve, there is a natural isomorphism

$$H^2(X, \mathbb{Z}(2)) \approx H^1(X, \mathbb{C}/\mathbb{Z}(2)).$$

Identifying $\mathbb{C}/\mathbb{Z}(2)$ with $\mathbb{C}^*$ by the map $\lambda \mapsto \exp[\lambda/(2\pi i)]$, we obtain a canonical identification of $H^2_{\mathscr{D}}(X, \mathbb{Z}(2))$ with $H^1(X, \mathbb{C}^*)$, the group of flat line bundles over X.

We set

$$H_{\mathbb{Z}} = \begin{pmatrix} 1 & \mathbb{Z}(1) & \mathbb{Z}(2) \\ 0 & 1 & \mathbb{Z}(1) \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$H_{\mathbb{C}} = \begin{pmatrix} 1 & \mathbb{C} & \mathbb{C} \\ 0 & 1 & \mathbb{C} \\ 0 & 0 & 1 \end{pmatrix}.$$

There is a natural bundle projection

(9) $$H_{\mathbb{Z}} \backslash H_{\mathbb{C}} \to \mathbb{C}^* \times \mathbb{C}^*;$$

it takes the coset of

$$\begin{pmatrix} 1 & u & w \\ 0 & 1 & v \\ 0 & 0 & 1 \end{pmatrix}$$

to (e^u, e^v). It has fiber $\mathbb{C}/\mathbb{Z}(2)$, which we identify with $\mathbb{C}^*$ as above.

PROPOSITION 6.1. *There is a natural connection on this bundle with curvature* $(dx/x) \wedge (dy/y)/(2\pi i)$, *where* x *and* y *are the coordinates in* $\mathbb{C}^* \times \mathbb{C}^*$.

PROOF. First consider the pull-back of the bundle (9) to $\mathbb{C} \times \mathbb{C}$

$$\mathbb{Z}(2)\backslash H_{\mathbb{C}} \to \mathbb{C} \times \mathbb{C} \tag{10}$$

along the map $(u, v) \mapsto (e^u, e^v)$. The map $H_{\mathbb{C}} \to \mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$ defined by

$$\begin{pmatrix} 1 & u & w \\ 0 & 1 & v \\ 0 & 0 & 1 \end{pmatrix} \mapsto \left(\exp\left(\frac{w}{2\pi i}\right), u, v\right)$$

induces an isomorphism of $\mathbb{Z}(2)\backslash H_{\mathbb{C}}$ with $\mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$ which commutes with the projections to $\mathbb{C} \times \mathbb{C}$. So the bundle (10) is trivial, and sections of it can be identified with maps $\zeta : \mathbb{C} \times \mathbb{C} \to \mathbb{C}^*$. Define a connection on this bundle by

$$\nabla \zeta = d\zeta - \zeta u dv/(2\pi i).$$

This connection is easily seen to be invariant under the left action

$$(n, m) : (\zeta, u, v) \mapsto (e^{nv}\zeta, u + 2\pi i n, v + 2\pi i m)$$

of $\mathbb{Z} \times \mathbb{Z}$ on $\mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$ induced by the left action of $H_{\mathbb{Z}}$ on $H_{\mathbb{C}}$. It therefore descends to a connection on the bundle (9). The connection form of the pull-back bundle is $u\,dv/(2\pi i)$, from which it follows that its curvature is $du \wedge dv/(2\pi i)$ and that the curvature of (9) is $(dx/x) \wedge (dy/y)/(2\pi i)$. □

Now suppose that X is a smooth curve over $\mathbb{C}$. Denote the function field of X by $\mathbb{C}(X)$ and the generic point $\operatorname{Spec} \mathbb{C}(X)$ of X by η_X. We first define the regulator on $K_2(\eta_X) := K_2(\mathbb{C}(X))$. The Deligne cohomology of η_X is defined by

$$H^m_{\mathscr{D}}(\eta_X, \Lambda(k)) = \varinjlim H^m_{\mathscr{D}}(U, \Lambda(k))$$

where the limit is taken over all Zariski open subsets U of X. In particular, we have

$$H^2_{\mathscr{D}}(\eta_X, \mathbb{Z}(2)) = \varinjlim H^1(U, \mathbb{C}^*).$$

This latter group is the group of flat line bundles at the generic point of X—elements of this group are flat line bundles defined on some Zariski open subset of X, and two such are identified if they agree on a smaller open subset. The product is tensor product.

By Matsumoto's Theorem [**34**], $K_2(\eta_X)$ is generated by symbols $\{f, g\}$, where $f, g \in \mathbb{C}(X)^\times$. The only relations that hold between these symbols are bilinearity

$$\{f_1 f_2, g\} = \{f_1, g\}\{f_2, g\}, \qquad \{f, g_1 g_2\} = \{f, g_1\}\{f, g_2\}$$

and the Steinberg relation $\{1 - f, f\} = 1$ whenever f and $1 - f$ are both in $\mathbb{C}(X)^\times$.

Now suppose that $\{f, g\} \in K_2(\eta_X)$. There is a Zariski open subset U of X such that f and g are both defined and invertible on U. They therefore define a regular function $(f, g) : U \to \mathbb{C}^* \times \mathbb{C}^*$. The pull-back of the line bundle $H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ to U is flat since it has curvature a multiple of $(df/f) \wedge (dg/g)$, which is zero because U is a curve. Denote it by $\langle f, g \rangle$. It is an element of $H^1(U, \mathbb{C}^*)$, and therefore of $H^1(\eta_X, \mathbb{C}^*)$.

REMARK 6.2. Observe that this construction makes sense when X is a Riemann surface and $\mathbb{C}(X)$ denotes the field of meromorphic functions of X. It is natural wi^t^h respect to holomorphic maps between Riemann surfaces.

PROPOSITION 6.3. *If f, g are invertible functions on the Zariski open subset U of X, the monodromy of the flat bundle $\langle f, g \rangle$ about a loop γ based at $p \in U$ is*

$$I(f, g, \gamma) := \int_\gamma \frac{df}{f}\frac{dg}{g} - \log g(p) \int_\gamma \frac{df}{f} + \log f(p) \int_\gamma \frac{dg}{g} \in \mathbb{C}/\mathbb{Z}(2).$$

PROOF. We will deduce the result by computing the monodromy about a loop γ in $\mathbb{C}^* \times \mathbb{C}^*$. The assertion will then follow by pulling back the answer to U. Let γ be a path in $\mathbb{C}^* \times \mathbb{C}^\times$ that begins at (x_0, y_0). For $t \in [0, 1]$, denote the path $s \mapsto \gamma(st)$ by γ_t. The horizontal lift of γ to $H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ which begins at the coset of

$$\begin{pmatrix} 1 & \log x_0 & w \\ 0 & 1 & \log y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

is

$$t \mapsto H_{\mathbb{Z}} \begin{pmatrix} 1 & \log x_0 & w \\ 0 & 1 & \log y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \int_{\gamma_t} \frac{dx}{x} & \int_{\gamma_t} \frac{dx}{x}\frac{dy}{y} \\ 0 & 1 & \int_{\gamma_t} \frac{dy}{y} \\ 0 & 0 & 1 \end{pmatrix}.$$

That is,

$$t \mapsto \begin{pmatrix} 1 & \log x_0 + \int_{\gamma_t} \frac{dx}{x} & w + \int_{\gamma_t} \frac{dx}{x}\frac{dy}{y} + \log x_0 \int_{\gamma_t} \frac{dy}{y} \\ 0 & 1 & \log y_0 + \int_{\gamma_t} \frac{dy}{y} \\ 0 & 0 & 1 \end{pmatrix}.$$

Now suppose that γ is a loop. Then the endpoint of the horizontal lift of γ is congruent to the matrix

$$\begin{pmatrix} 1 & -\int_\gamma \frac{dx}{x} & \int_\gamma \frac{dx}{x} \int_\gamma \frac{dy}{y} \\ 0 & 1 & -\int_\gamma \frac{dy}{y} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \log x_0 + \int_\gamma \frac{dx}{x} & w + \int_\gamma \frac{dx}{x}\frac{dy}{y} + \log x_0 \int_\gamma \frac{dy}{y} \\ 0 & 1 & \log y_0 + \int_\gamma \frac{dy}{y} \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \log x_0 & w + \int_\gamma \frac{dx}{x}\frac{dy}{y} + \log x_0 \int_\gamma \frac{dy}{y} - \log y_0 \int_\gamma \frac{dx}{x} \\ 0 & 1 & \log y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

modulo the left action of $H_{\mathbb{Z}}$. It follows that the holonomy about γ is

$$\int_\gamma \frac{dx}{x}\frac{dy}{y} + \log x_0 \int_\gamma \frac{dy}{y} - \log y_0 \int_\gamma \frac{dx}{x} \mod \mathbb{Z}(2).$$

The result follows by pulling the result back to U along the map (f, g). □

PROPOSITION 6.4. *If $f, f_1, f_2, g \in \mathbb{C}(X)^\times$, then*

$$\langle f_1 f_2, g\rangle = \langle f_1, g\rangle\langle f_2, g\rangle \quad \textit{and} \quad \langle f, g\rangle^{-1} = \langle g, f\rangle.$$

Moreover, if $f, 1-f \in \mathbb{C}(X)^*$, then $\langle 1-f, f\rangle = 1$.

PROOF. It is clear from properties of the logarithm and (5.1) that

$$I(f_1 f_2, g, \gamma) = I(f_1, g, \gamma) + I(f_2, g, \gamma).$$

It is not difficult to use (5.1) to prove that

$$I(f, g, \gamma) + I(g, f, \gamma) = 0.$$

These imply the linearity and skew symmetry of the symbol $\langle f, g\rangle$.

It suffices to prove the Steinberg relation in the universal case where $U = \mathbb{C} - \{0, 1\}$ and $f = x$. The line bundle $\langle 1-x, x\rangle$ is trivial as a flat bundle if and only if it has a flat section. The dilogarithm provides such a section. Define $s : \mathbb{C} - \{0, 1\} \to H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ by

$$s(x) = H_{\mathbb{Z}} \begin{pmatrix} 1 & \log(1-x) & -\ln_2 x \\ 0 & 1 & \log x \\ 0 & 0 & 1 \end{pmatrix}.$$

This section is flat, so the Steinberg relation holds. □

THEOREM 6.5 [**2, 3**]. *Taking $\{f, g\}$ to $\langle f, g\rangle$ defines a map*

$$K_2(\eta_X) \to H^1(\eta_X, \mathbb{C}^*),$$

which is the Chern class c_2.

PROOF. The first assertion is an easy consequence of Matsumoto's description of K_2 and (6.4). The second follows from the fact that the symbol $\{f, g\}$ is the cup product of $f, g \in K_1(\eta_X) \approx \mathbb{C}(X)^\times$. Properties of Chern classes then imply that

$$c_2(\{f, g\}) = c_1(f) \cup c_1(g)$$

where the right-hand side is the cup product of

$$c_1(f), c_1(g) \in H^1_{\mathscr{D}}(\eta_X, \mathbb{Z}(1)) \approx \mathbb{C}(X)^\times.$$

Under this isomorphism, c_1 is just the identity.

The formula for the cup product in Deligne cohomology implies that $c_1(f) \cup c_1(g)$ is represented by the element of

$$H^2_{\mathscr{D}}(\eta_X, \mathbb{Z}(2)) \approx H^1(\eta_X, \mathbb{C}/\mathbb{Z}(2))$$

defined by $\gamma \mapsto I(f, g, \gamma)$. □

We now globalize this construction. For each $x \in X$, there is map

$$\delta_x : H^1(\eta_X, \mathbb{C}^*) \to \mathbb{C}^*.$$

To define $\delta_x(l)$, represent l by a flat line bundle $L \to U$ over a Zariski open subset U of X. Choose a small closed disk $\overline{\Delta}$ in X, centered at x, such that $\overline{\Delta} - \{x\}$ is contained in U. Define $\delta(l)$ to be the monodromy of L about the boundary of $\overline{\Delta}$.

Suppose that $\nu : F^\times \to \mathbb{Z}$ is a valuation on a field F. Let $\mathscr{O}$ be the associated valuation ring (i.e., 0 and those elements of $F^\times$ with valuation ≥ 0). Let $\mathfrak{P}$ be the maximal ideal of $\mathscr{O}$. The tame symbol of $f, g \in F^\times$ is defined by

$$(f, g)_\nu = (-1)^{\nu(f)\nu(g)} \frac{f^{\nu(g)}}{g^{\nu(f)}} \mod \mathfrak{P}.$$

(See, e.g., **[35]**.) To each $x \in X$ associate the valuation that takes a function f to its order $\nu_x(f)$ at x. In this way we associate a tame symbol $(\ ,\ )_x$ to each $x \in X$.

PROPOSITION 6.6. *Suppose that $x \in X$. If $f, g \in \mathbb{C}(X)$, then*

$$\delta_x\langle f, g\rangle = (f, g)_x.$$

PROOF. Since both sides of the expression in the statement of the proposition are skew symmetric and bilinear, we can reduce, with the help of (6.2), to the following two cases. First, if z is a local holomorphic parameter about x, then we have to show that $\delta_x\langle z, z\rangle = -1$. Second, if f is a unit in a neighbourhood of x, then

$$\delta_x\langle f, g\rangle = f^{\nu_x(g)}(x).$$

From (6.3) and (5.1)(ii) it follows that

$$I(z, z, p) = \int_{\partial\Delta} \frac{dz}{z}\frac{dz}{z} = \frac{1}{2}\left(\int_{\partial\Delta}\frac{dz}{z}\right)^2 = \frac{(2\pi i)^2}{2} \in \mathbb{C}/\mathbb{Z}(2),$$

where Δ is a sufficiently small imbedded disk in X centered at x. Under the standard isomorphism $\mathbb{C}/\mathbb{Z}(2) \approx \mathbb{C}^*$, this corresponds to $e^{i\pi} = -1$. This proves the first assertion.

To prove the second, write $f = e^{\phi}$, where ϕ is holomorphic in a neighbourhood of x. By (6.3), we have

$$I(f, g, p) = \int_{\partial\Delta} (\phi(z) - \phi(p))\frac{dg}{g} + \phi(p)\int_{\partial\Delta}\frac{dg}{g} = \int_{\partial\Delta}\phi(z)\frac{dg}{g}.$$

By the Residue Theorem this equals $2\pi i \nu_x(g)\phi(x)$. So

$$\delta_x\langle f, g\rangle = \exp(\nu_x(g)\phi(x)) = f(x)^{\nu_x(g)}.$$

This proves the second assertion. □

The following version of the Gysin sequence is easily verified.

PROPOSITION 6.7. *The sequence*

$$0 \to H^1(X, \mathbb{C}^*) \to H^1(\eta_X, \mathbb{C}^*) \xrightarrow{\oplus\delta_x} \bigoplus_{x\in X}\mathbb{C}*$$

is exact. □

The analogue of the Gysin sequence in algebraic K-theory is the localization sequence. In our case, it asserts that the sequence

$$K_2(X) \to K_2(\eta_X) \xrightarrow{\oplus(\ ,\ )_x} \bigoplus_{x\in X}\mathbb{C}^*$$

is exact. By (6.6), the diagram

$$\begin{array}{ccccccc} & & K_2(X) & \to & K_2(\eta_X) & \xrightarrow{\oplus(\ ,\ )_x} & \bigoplus_{x\in X}\mathbb{C}^* \\ & & & & \downarrow c_2 & & \| \\ 0 & \to & H^1(X, \mathbb{C}^*) & \to & H^1(\eta_X, \mathbb{C}^*) & \xrightarrow{\oplus\delta_x} & \bigoplus_{x\in X}\mathbb{C}* \end{array}$$

commutes. Since the rows are exact, the Chern class c_2 induces a map $K_2(X) \to H^1(X, \mathbb{C}^*)$ which must be the Chern class by naturality.

This construction extends easily to give a description of the regulator

$$c_2 : K_2(X) \to H^2_{\mathscr{D}}(X, \mathbb{Z}(2))$$

where X is a smooth variety over $\mathbb{C}$. The construction proceeds in the same way, except that the line bundles are no longer flat. For a mixed Hodge structure H denote by $\mathrm{Hom}_{\mathrm{Hodge}}(\mathbb{Z}, H)$, the set of "Hodge classes" in H of type $(0, 0)$, by ΓH.

PROPOSITION 6.8. *If X is smooth over $\mathbb{C}$, then there is a natural isomorphism between $H^2_{\mathscr{D}}(X, \mathbb{Z}(2))$ and the group that consists of the pairs (L, ∇), where L is a holomorphic line bundle over X and ∇ is a holomorphic connection whose curvature times $2\pi i$ lies in $\Gamma H^2(X, \mathbb{Z}(2))$.* □

7. The polylogarithm variation of mixed Hodge structure

Let X be a smooth complex algebraic curve and $\overline{X}$ a smooth compactification of it. Let $D = \overline{X} - X$. Recall from **[41]** that a *variation of mixed Hodge structure* over X consists of

(i) a $\mathbb{Q}$ local system $\mathbb{V} \to X$ which has a filtration by local systems

$$\cdots \subseteq \mathbb{W}_{l-1} \subseteq \mathbb{W}_l \subseteq \mathbb{W}_{l+1} \subseteq \cdots$$

that exhausts $\mathbb{V}$ and whose intersection is 0. We shall denote the fiber of $\mathbb{V}$ over $x \in X$ by V_x and the fiber of $\mathbb{W}_l$ by $W_l V_x$. We will also assume that each local monodromy operator $T_P : V_P \to V_P$, about each $P \in D$, is unipotent.

(ii) a Hodge filtration

$$\cdots \supseteq \mathcal{F}^{p-1} \supseteq \mathcal{F}^p \supseteq \mathcal{F}^{p+1} \supseteq \cdots$$

of the corresponding holomorphic vector bundle $\mathcal{V} := \mathbb{V} \otimes_{\mathbb{Q}} \mathcal{O}_X$ by holomorphic sub-bundles. These are required to satisfy Griffiths' transversality: If

$$\nabla : \mathcal{V} \to \mathcal{V} \otimes_{\mathcal{O}_X} \Omega^1_X$$

is the natural flat connection, then

$$\nabla(\mathcal{F}^p) \subseteq \mathcal{F}^{p-1} \otimes_{\mathcal{O}_X} \Omega^1_X.$$

Denote the fiber of $\mathcal{F}^p$ over $x \in X$ by $F^p V_x$.

(iii) For each $x \in X$, the filtrations $W_\bullet V_x$ and $F^\bullet V_x$ define a mixed Hodge structure on V_x.

(iv) There is a nondegenerate $\mathbb{Q}$-valued $(-1)^l$- symmetric bilinear form on each of the bundles $\mathrm{Gr}^W_l \mathbb{V}$ which satisfies the Riemann bilinear relations on each fiber.

(v) Denote Deligne's canonical extension of $\mathcal{V}$ to $\overline{X}$ by $\overline{\mathcal{V}} \to \overline{X}$ **[13]**. The Hodge bundles $\mathcal{F}^p$ are required to extend to holomorphic sub-bundles $\overline{\mathcal{F}}^p$ of $\overline{\mathcal{V}}$. (Note that the weight bundles $\mathcal{W}_l := \mathbb{W}_l \otimes_{\mathbb{Q}} \mathcal{O}_X$ automatically extend to sub-bundles $\overline{\mathcal{W}}_l$ of $\overline{\mathcal{V}}$ as they are flat.)

(vi) About each point $P \in D$, there is a relative weight filtration **[41]**. This is an important condition which is rather technical in general. However, in the case where the global monodromy representation

$$\rho_x : \pi_1(X, x) \to \mathrm{GL}(V_x)$$

is unipotent, the condition reduces to the much simpler condition

$$N_P(W_l V_x) \subseteq W_{l-2} V_x$$

for each $P \in D$, where N_P is the local monodromy logarithm

$$N_P = \frac{1}{2\pi i} \log T_P .$$

(See **[29]**.)

THEOREM 7.1. *The nth polylogarithm local system underlies a good variation of mixed Hodge structure whose weight-graded quotients are canonically isomorphic to* $\mathbb{Q}$, $\mathbb{Q}(1), \dots, \mathbb{Q}(n)$.

PROOF. Let $\mathbb{V} \to \mathbb{C} - \{0, 1\}$ be the nth polylogarithm local system, and $\mathscr{V}$ the corresponding holomorphic vector bundle. By (2.5), the canonical extension of this to $\mathbb{P}^1$ is the trivial bundle $\mathbb{P}^1 \times \mathbb{C}^{n+1} \to \mathbb{P}^1$. Denote the standard basis of $\mathbb{C}^{n+1}$ by $e_0, e_1, \dots, e_n$. The fiber V_x is the $\mathbb{Q}$-linear span of $\lambda_0(x), \lambda_1(x), \dots, \lambda_n(x)$, the rows of $\Lambda(x)$. Define

$$W_{-2l+1}\mathbb{C}^{n+1} = W_{-2l}\mathbb{C}^{n+1} = \operatorname{span}\{e_l, \dots, e_n\} \tag{11}$$

and

$$F^{-p}\mathbb{C}^{n+1} = \operatorname{span}\left\{e_0, \dots, e_p\right\}. \tag{12}$$

Define

$$\overline{\mathscr{F}}^p = \mathbb{P}^1 \times F^p\mathbb{C}^{n+1} \subseteq \overline{\mathscr{V}} \quad \text{and} \quad \overline{\mathscr{W}}_l = \mathbb{P}^1 \times W_l\mathbb{C}^{n+1} \subseteq \overline{\mathscr{V}}.$$

Observe that the weight filtration comes from a filtration defined on $\mathbb{V}$:

$$W_{-2l+1}V_x = W_{-2l}V_x = \operatorname{span}\{\lambda_l(x), \dots, \lambda_n(x)\}$$

The polarizations on the weight-graded quotients are simply the ones that give each vector $(2\pi i)e_k$ length 1. □

Suppose that $\mathbb{V} \to X$ is a good variation of mixed Hodge structure with unipotent monodromy about each point of $D = \overline{X} - X$. Let $P \in D$. For each nonzero tangent vector $\vec{v}$ of X at p, there is a canonical mixed Hodge structure on $V_{\mathbb{C}}$, the fiber of $\overline{\mathscr{V}}$ over P. This is called the *limit mixed Hodge structure associated to* $\vec{v}$. The Hodge and weight filtrations on $\mathbb{V}_{\mathbb{C}}$ are defined by letting $F^pV_{\mathbb{C}}$ and $W_lV_{\mathbb{C}}$ be the fibers of $\overline{\mathscr{F}}^p$ and $\overline{\mathscr{W}}_l$ over P, respectively. To construct the limit mixed Hodge structure, we have to construct a rational form $V_{\mathbb{Q}}$ of $V_{\mathbb{C}}$ and show that the weight filtration defined above is the complexification of a filtration of $V_{\mathbb{Q}}$.

To construct $V_{\mathbb{Q}}$, choose an imbedded closed disk $\overline{\Delta}$ in $\overline{X}$ centered at P. Let t be a holomorphic parameter in $\overline{\Delta}$ such that $t(P) = 0$ and $|t| = 1$ is $\partial\overline{\Delta}$. By choosing the disk to be small enough, we may suppose that $\overline{\Delta} - \{0\} \subseteq X$.

We first consider the case where $\vec{v} = \partial/\partial t$. Let $x \in \overline{\Delta}$ be the point corresponding to $t = 1$. Choose a $\mathbb{Q}$ basis $v_1, \dots, v_m$ of V_x, the fiber of $\mathbb{V}$ over x. Let $v_1(t), \dots, v_m(t)$ be flat (possibly multi-valued) sections of

$\mathbb{V}$ over $\overline{\Delta}^*$ that satisfy $v_j(1) = v_j$ for each j. Let $T : V_x \to V_x$ be the local monodromy operator, and $N = \log T/(2\pi i)$ be the local monodromy logarithm. For each j, define

$$s_j(t) = v_j(t)t^{-N}.$$

Then each $s_j(t)$ is a single-valued section of $\overline{\mathcal{V}}$ over $\overline{\Delta}^*$. In fact, by the construction of the canonical extension $\overline{\mathcal{V}} \to \overline{X}$, the s_j comprise a local framing of $\overline{\mathcal{V}}$ over $\overline{\Delta}$. In particular, $s_1(0), \ldots, s_m(0)$ is a $\mathbb{C}$ basis of $V_{\mathbb{C}}$. Define the rational form $V_{\mathbb{Q}}$ of $V_{\mathbb{C}}$ that corresponds to $\partial/\partial t$ to be the $\mathbb{Q}$-linear span of $s_1(0), \ldots, s_m(0)$. By choosing the basis $v_1, \ldots, v_m$ of V_x to be adapted to the weight filtration, one can easily show that the weight filtration of $V_{\mathbb{C}}$ is the complexification of a filtration of $V_{\mathbb{Q}}$. The $\mathbb{Q}$ structure on $V_{\mathbb{C}}$ that corresponds to the tangent vector $\vec{v} = \lambda\,\partial/\partial t$ is defined to be

$$V_{\mathbb{Q}}(\vec{v}) = V_{\mathbb{Q}}\lambda^N.$$

It is not difficult to show that this rational structure depends only on the tangent vector $\vec{v}$, and not on the choice of the parameter t.

If the weight-graded quotients of $\mathbb{V} \to X$ are constant as variations of Hodge structure (e.g., the polylogarithm variations), it is not difficult to show that

$$((V_{\mathbb{Q}}(\vec{v}), W_\bullet), (V_{\mathbb{C}}, F^\bullet))$$

is a mixed Hodge structure, and that $N : V \to V(-1)$ is a morphism of mixed Hodge structures.

The following result is due to Deligne **[16]** in the case $n = 2$. It is a straightforward computation using the monodromy computation (2.2) and the procedure described above.

THEOREM 7.2. *Let z be the natural coordinate function on $\mathbb{C} - \{0, 1\}$. The limit mixed Hodge structure on the nth polylogarithm variation at the tangent vector $\partial/\partial z$ at 0 has rational structure spanned by the vectors*

$$\begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ \vdots \\ s_n \end{pmatrix} = \left(\begin{array}{c|ccccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 2\pi i & 0 & \cdots & 0 \\ \vdots & 0 & (2\pi i)^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & (2\pi i)^n \end{array}\right) \begin{pmatrix} e_0 \\ e_1 \\ \vdots \\ \vdots \\ e_n \end{pmatrix}.$$

The limit mixed Hodge structure associated with the tangent vector $-\partial/\partial z$ at 1 has rational structure spanned by the vectors

$$\begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ \vdots \\ s_n \end{pmatrix} = \left(\begin{array}{c|ccccc} 1 & 0 & \zeta(2) & \cdots & \zeta(n) \\ \hline 0 & 2\pi i & 0 & \cdots & 0 \\ \vdots & 0 & (2\pi i)^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & (2\pi i)^n \end{array}\right) \begin{pmatrix} e_0 \\ e_1 \\ \vdots \\ \vdots \\ e_n \end{pmatrix}$$

where $\zeta(s)$ denotes the Riemann zeta function. In both cases, the Hodge and weight filtrations are defined as in (11) *and* (12).

8. Mixed Hodge structure on π_1

In this section we give the construction of the mixed Hodge structure on the fundamental group $\pi_1(X, x)$ in the special case where X is a smooth variety over $\mathbb{C}$ whose $H^1(X)$ is pure of weight 2. This purity condition may be restated in several equivalent ways. For example, $H^1(X)$ is pure of weight 2 if and only if one and (and hence all) smooth completions $\overline{X}$ of X have first Betti number 0. In particular, all Zariski open subsets of Grassmannians have this property.

To put a mixed Hodge structure on $\pi_1(X, x)$, it is necessary to linearize it. The first step is to replace the fundamental group by its group algebra $\mathbb{Q}\pi_1(X, x)$. This object is not well enough behaved, and we need to linearize it further, which we do by completion. Let J be the augmentation ideal, that is, J is the kernel of the augmentation $\varepsilon : \mathbb{Q}\pi_1(X, x) \to \mathbb{Q}$ which takes each $g \in \pi_1(X, x)$ to 1. The powers of J define a topology on $\mathbb{Q}\pi_1(X, x)$, and we consider

$$\mathbb{Q}\pi_1(X, x)\hat{} = \varinjlim \mathbb{Q}\pi_1(X, x)/J^m,$$

the J-adic completion.

The completed group ring is a complete Hopf algebra, that is, it is a topological algebra and has a continuous algebra homomorphism

$$\Delta : \mathbb{Q}\pi_1(X, x)\hat{} \to \mathbb{Q}\pi_1(X, x)\hat{} \mathbin{\hat{\otimes}} \mathbb{Q}\pi_1(X, x)\hat{}.$$

Here $\hat{\otimes}$ denotes completed tensor product. This algebra homomorphism is induced by the usual diagonal

$$\mathbb{Q}\pi_1(X, x) \to \mathbb{Q}\pi_1(X, x) \otimes \mathbb{Q}\pi_1(X, x)$$

which takes each $g \in \mathbb{Q}\pi_1(X, x)$ to $g \otimes g$.

The Malcev Lie algebra $\mathfrak{g}(X, x)$ associated to $\pi_1(X, x)$ is, by definition, the set of primitive elements

$$\{X \in \mathbb{Q}\pi_1(X, x)^\wedge : \Delta X = 1 \otimes X + X \otimes 1\}$$

of $\mathbb{Q}\pi_1(X, x)^\wedge$. This is a complete topological Lie algebra. The bracket of the elements A, B of $\mathfrak{g}$ is their commutator $AB - BA$, which is again primitive. The topology is induced from that of $\mathbb{Q}\pi_1(X, x)^\wedge$. The completed group ring may be recovered from its primitive elements as its completed enveloping algebra $U^\wedge\mathfrak{g}(X, x)$ (see [**36**, Appendix A]).

A mixed Hodge structure on $\mathbb{Q}\pi_1(X, x)^\wedge$ is, by definition, a compatible sequence of mixed Hodge structures

$$\cdots \to \mathbb{Q}\pi_1(X, x)/J^3 \to \mathbb{Q}\pi_1(X, x)/J^2 \to \mathbb{Q}\pi_1(X, x)/J \to \mathbb{Q} \to 0$$

on the truncations of the group ring. Note that each of these is a finite-dimensional vector space.

If the product and diagonal of $\mathbb{Q}\pi_1(X, x)^\wedge$ are morphisms of mixed Hodge structure, then, since $\mathfrak{g}(X, x)$ is the kernel of the reduced diagonal

$$\overline{\Delta} : \mathbb{Q}\pi_1(X, x)^\wedge \to [\mathbb{Q}\pi_1(X, x)^\wedge/\mathbb{Q}] \widehat{\otimes} [\mathbb{Q}\pi_1(X, x)^\wedge/\mathbb{Q}] \approx J\widehat{\otimes}J,$$

the Malcev Lie algebra inherits a mixed Hodge structure compatible with its Lie algebra structure. This mixed Hodge structure determines the one on $\mathbb{Q}\pi_1(X, x)^\wedge$ since $\mathfrak{g}(X, x)$ generates $\mathbb{Q}\pi_1(X, x)^\wedge$ topologically.

THEOREM 8.1 [**33, 25**]. *If (X, x) is a complex algebraic variety, then $\mathbb{Q}\pi_1(X, x)^\wedge$ and $\mathfrak{g}(X, x)$ have canonical mixed Hodge structures which are compatible with their algebraic structures.*

It is important to note that if $\mathfrak{g}(X, x)$ is nonabelian, then these mixed Hodge structures depend nontrivially on the basepoint $x \in X$ (cf. [**24**, §§6,7]).

We shall sketch the construction of these mixed Hodge structures in the case when $H^1(X)$ is pure of weight 2. Write $X = \overline{X} - D$, where $\overline{X}$ is smooth and complete, and D is a divisor with normal crossings in $\overline{X}$. For convenience, we set

$$\Omega^\bullet(X) = H^0(\overline{X}, \Omega^\bullet_{\overline{X}}(\log D)).$$

By results of Deligne [**14**, (3.2.14)], every element of $\Omega^\bullet(X)$ is closed and the obvious map $\Omega^\bullet(X) \to H^\bullet(X, \mathbb{C})$ is injective. The tensor algebra

$$T := \bigoplus_{n\in\mathbb{N}} \Omega^1(X)^{\otimes(-n)}$$

on the dual of $\Omega^1(X)$ is a Hopf algebra; the diagonal is defined as the algebra homomorphism that takes each $X \in \Omega^1(X)^*$ to $1 \otimes X + X \otimes 1$. The ideal

$(\operatorname{im}\delta)$ in T generated by the image of the dual of the cup product

$$\delta : \Omega^2(X)^* \to \Lambda^2\Omega^1(X)^* \subseteq \Omega^1(X)^{\otimes(-2)}$$

is a Hopf ideal. It follows that $A := T/(\operatorname{im}\delta)$ is a Hopf algebra whose diagonal is induced from that of the tensor algebra. The set of primitive elements of A is

$$PA = \mathbb{L}(\Omega^1(X)^*)/(\operatorname{im}\delta),$$

where $\mathbb{L}(V)$ denotes the free Lie algebra generated by the vector space V.

Denote the ideal generated by $\Omega^1(X)^*$ by I. The powers of I define a topology on A. Denote the I-adic completion of A by $A^\frown$. The set of primitive elements of $A^\frown$ is the I-adic completion of PA.

The following result is a special case of a theorem of K.-T. Chen [**12**, (3.5)].

PROPOSITION 8.2. *For each $x \in X$ there is a canonical isomorphism*

$$\Theta_x : \mathbb{C}\pi_1(X, x)^\frown \to A^\frown$$

of complete Hopf algebras.

PROOF. Let $\omega \in \Omega^1(X) \otimes \Omega^1(X)^*$ be the element that corresponds to the identity $\Omega^1(X) \to \Omega^1(X)$. This can be viewed as an element of $\Omega^1(X) \otimes PA^\frown$. This form is integrable, that is,

$$d\omega + \omega \wedge \omega = 0.$$

It follows that the value of the A-valued iterated integral

$$1 + \int \omega + \int \omega\omega + \int \omega\omega\omega + \cdots$$

on each path in X depends only on its homotopy class relative to its end points (cf. [**24**, §3], for example). It follows from this and (5.1)(ii) that this map induces a well-defined homomorphism from $\pi_1(X, x)$ into the group of units of $A^\frown$. This extends to an algebra homomorphism

$$\mathbb{C}\pi_1(X, x) \to A^\frown.$$

Since the augmentation ideal of $\mathbb{C}\pi_1(X, x)$ is mapped into the ideal I of $A^\frown$, it follows that this homomorphism extends to a continuous algebra homomorphism

$$\Theta_x : \mathbb{C}\pi_1(X, x)^\frown \to A^\frown.$$

The property (5.1)(iv) implies that Θ_x commutes with the diagonals; that is, Θ_x is a Hopf algebra homomorphism.

The graded module associated to the filtration of $\mathbb{C}\pi_1(X, x)^\frown$ by powers of J is generated by J/J^2, which is isomorphic to $H_1(X)$. Similarly, the graded module associated to the filtration of $A^\frown$ by the powers of I is generated by I/I^2, which is also isomorphic to $H_1(X)$. The map Θ_x induces the isomorphism $J/J^2 \approx I/I^2$ that corresponds to the identifications of each with $H_1(X)$. Since both algebras are complete, this implies that Θ_x is surjective. One can show, without too much difficulty, that the map $J^2/J^3 \to I^2/I^3$ induced by Θ_x is also an isomorphism (cf. [**24**, (6.1)]). It is then relatively straightforward to show that Θ_x must be injective. The idea is that $A^\frown$ contains no other relations other than those that are consequences of the quadratic ones, while $\mathbb{C}\pi_1(X, x)^\frown$ has at least these relations. Since Θ_x is well defined, it must be an isomorphism. □

The next step in constructing the mixed Hodge structure on $\mathbb{Q}\pi_1(X, x)^\frown$ is to define the Hodge and weight filtrations. We do this by defining them on $A^\frown$ and transferring them to $\mathbb{Q}\pi_1(X, x)^\frown$ via the isomorphism Θ_x.

The ring A is graded because the ideal $(\operatorname{im}\delta)$ is graded. Write $A = \bigoplus_{n\in\mathbb{N}} A_n$ where A_n is the image of $\Omega^1(X)^{\otimes(-n)}$ in A. The assumption that $H^1(X)$ be pure of weight 2 implies that it has Hodge type $(1, 1)$. Consequently, $\Omega^1(X)^* \approx H_1(X)$ has Hodge type $(-1, -1)$. It is therefore natural to define Hodge and weight filtrations on A by

$$F^{-p}A = \bigoplus_{n\leq p} A_n \quad \text{and} \quad W_{-m}A = \bigoplus_{n\geq m/2} A_n .$$

Now define Hodge and weight filtrations on $\mathbb{Q}\pi_1(X, x)/J^l$ by transferring the Hodge and weight filtrations from A/I^l via the isomorphism

$$\Theta_x : \mathbb{Q}\pi_1(X, x)/J^l \to A/I^l.$$

This data defines a mixed Hodge structure on $\mathbb{Q}\pi_1(X, x)/J^l$. To see this, first observe that

$$\operatorname{Gr}^W_m \mathbb{Q}\pi_1(X, x)/J^l = \begin{cases} J^r/J^{r+1} & \text{when } m = -2r \text{ and } 0 \leq r < l, \\ 0 & \text{otherwise.} \end{cases}$$

The Hodge filtration induced on $\operatorname{Gr}^W_{-2p}$ satisfies

$$F^{-p}\left[J^p/J^{p+1}\right] = J^p/J^{p+1}, \qquad F^{-p+1}\left[J^p/J^{p+1}\right] = 0$$

when $0 \leq p < l$. Because the weight filtration is defined over $\mathbb{Q}$, it follows that these filtrations define a mixed Hodge structure on $\mathbb{Q}\pi_1(X, x)/J^l$ whose weight-graded quotients are all of even weight, and where the $2p$th graded quotient is of type (p, p). Since the multiplication and comultiplication of A preserve the filtrations and are defined over $\mathbb{Q}$, they are morphisms of mixed Hodge structure. It follows that $\mathfrak{g}(X, x)$, endowed with the induced filtrations, is a mixed Hodge structure. □

We conclude this section by relating this mixed Hodge structure to unipotent variations of mixed Hodge structure. A variation of mixed Hodge structure over a smooth variety X is *good* if its restriction to every curve satisfies the conditions in §7. A good variation of mixed Hodge structure $\mathbb{V} \to X$

over a smooth variety X is *unipotent* if one (and hence all) monodromy representations

$$\rho_x : \pi_1(X, x) \to \operatorname{Aut} V_x \tag{13}$$

are unipotent. This condition is equivalent to the condition that each of the variations of Hodge structure $\operatorname{Gr}_m^W \mathbb{V}$ be constant.

The monodromy representation (13) induces a map

$$\theta_x : \mathbb{Q}\pi_1(X, x) \to \operatorname{End} V_x .$$

Since the representation is unipotent, there exists l such that J^l is contained in $\ker \theta_x$. It follows that there is an algebra homomorphism

$$\theta_x : \mathbb{Q}\pi_1(X, x)/J^l \to \operatorname{End} V_x . \tag{14}$$

Both sides of this last equation have natural mixed Hodge structures.

THEOREM 8.3 [**29**]. *For each $x \in X$, the representation* (14) *is a morphism of mixed Hodge structures.*

Define the category of Hodge theoretic representations of $\pi_1(X, x)$ to be the set of pairs (V, ρ), where V is a mixed Hodge structure and ρ is a unipotent representation $\pi_1(X, x) \to \operatorname{Aut} V$ which induces a morphism of mixed Hodge structure

$$\theta_x : \mathbb{Q}\pi_1(X, x)/J^l \to \operatorname{End} V$$

for l sufficiently large. Theorem 8.3 implies that taking the fiber at x defines a functor from the category of unipotent variations of mixed Hodge structure over X to the category of Hodge theoretic representations of $\pi_1(X, x)$.

THEOREM 8.4 [**29**]. *This functor is an equivalence of categories.*

The proofs of Theorems 8.3 and 8.4 in the case when $H^1(X)$ is pure of weight 2 are considerably simpler than in the general case, and may be found in [**30**].

A good variation of mixed Hodge structure $\mathbb{V}$ over a smooth variety X is called a *Tate variation of mixed Hodge structure* if all of its weight-graded quotients are constant and of even weight, and if each of the variations $\operatorname{Gr}_{2p}^W \mathbb{V}$ is of type (p, p). The polylogarithm variations are examples of Tate variations of mixed Hodge structure. Now suppose that $\overline{X}$ is any smooth compactification of X where $D = \overline{X} - X$ is a divisor with normal crossings in $\overline{X}$. Let $\overline{\mathscr{V}} \to \overline{X}$ be the canonical extension of $\mathbb{V}$ to $\overline{X}$. The following result is a straightforward consequence of Theorem 8.3 and [**23**, (6.4)].

THEOREM 8.5. *If $\mathbb{V} \to X$ is a Tate variation of mixed Hodge structure, then its canonical extension $\overline{\mathscr{V}} \to \overline{X}$ is trivial as a holomorphic vector bundle, so* *that there is a complex vector space* V *and an isomorphism*

$$\begin{array}{ccc} \overline{\mathscr{V}} & \rightarrow & V \times \overline{X} \\ \downarrow & & \downarrow \\ \overline{X} & = & \overline{X} \end{array}$$

of holomorphic vector bundles. Moreover, there are filtrations $F^\bullet$ *and* $W_\bullet$ *of* V *such that the extended Hodge and weight bundles* $\overline{\mathscr{F}}^p$ *and* $\overline{\mathscr{W}}_l$ *correspond to* $F^p \times \overline{X}$ *and* $W_l \times \overline{X}$, *respectively, under the bundle isomorphism.* □

One important example of a unipotent variation of mixed Hodge structure over a smooth variety X is the one whose fiber over $x \in X$ is the truncated group ring $\mathbb{Q}\pi_1(X, x)/J^l$. This is a good variation because the monodromy representation

$$\mathbb{Q}\pi_1(X, x)/J^l \rightarrow \operatorname{Aut} \mathbb{Q}\pi_1(X, x)/J^l$$

can be written in terms of left and right multiplication, and is thus a morphism of mixed Hodge structure. Such variations form an inverse system of variations, and we call the inverse limit the *tautological variation* over X. In the case when $H^1(X)$ has weight 2, this variation is a Tate variation of mixed Hodge structure, and can be described explicitly. View $A^\frown$ as a subalgebra of End $A^\frown$ via the right regular representation.

PROPOSITION 8.6. *If* $H^1(X)$ *is of weight 2, then the tautological variation over* X *has canonical extension* $A^\frown \times \overline{X} \rightarrow \overline{X}$. *The connection form of the canonical flat connection on this bundle is given by the* $PA^\frown$ *valued* 1*-form*

$$\omega \in \Omega^1(X) \otimes H_1(X) \subseteq \Omega^1(X) \otimes PA^\frown \subseteq \Omega^1(X) \otimes \operatorname{End} A^\frown$$

which corresponds to the canonical isomorphism $\Omega^1(X) \approx H^1(X)$. *The extended Hodge and weight bundles are* $F^p A^\frown \times \overline{X}$ *and* $W_l A^\frown \times \overline{X}$. □

9. Hodge theoretic interpretation of regulators

In this section we give a Hodge theoretic interpretation of the regulators constructed in §§4 and 6. These interpretations are due to Deligne **[16]**.

Throughout this section, Λ will denote $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$. Suppose that

$$V = (V_\Lambda, (V_{\Lambda\otimes\mathbb{Q}}, W_\bullet), (V_\mathbb{C}, F^\bullet))$$

is a mixed Hodge structure where the underlying lattice V_Λ is torsion free. The ring of endomorphisms End V has a mixed Hodge structure whose

Hodge and weight filtrations are defined by

$$F^p \operatorname{End}_{\mathbb{C}} V = \left\{ \phi \in \operatorname{End}_{\mathbb{C}} V : \phi(F^q V) \subseteq F^{p+q} V \right\}$$

and

$$W_l \operatorname{End}_{\Lambda\otimes\mathbb{Q}} V = \left\{ \phi \in \operatorname{End}_{\Lambda\otimes\mathbb{Q}} V : \phi(W_m V) \subseteq W_{m+l} V \right\}.$$

Set $\mathfrak{g} = W_{-1} \operatorname{End} V$. This is a nilpotent Lie algebra with a mixed Hodge structure—the bracket being the commutator $[\phi, \psi] = \phi\psi - \psi\phi$. The subspace $F^0\mathfrak{g}$ is a Lie subalgebra. Denote the simply connected Lie groups which correspond to $\mathfrak{g}_{\mathbb{C}}$ and $F^0\mathfrak{g}$ by G and F^0G, respectively. These are unipotent subgroups of $\operatorname{Aut}_{\mathbb{C}} V$. Set $G_\Lambda = G \cap \operatorname{End}_\Lambda V$. We view G_Λ as acting on the right of V_Λ.

For each $g \in G$, the triple $(V_\Lambda g, (V_{\Lambda\otimes\mathbb{Q}} g, W_\bullet g), (V_{\mathbb{C}}, F^\bullet))$ is a mixed Hodge structure whose weight-graded quotients are canonically isomorphic to those of V. It is not difficult to show that every mixed Hodge structure with torsion free lattice and weight-graded quotients canonically isomorphic to those of V can be constructed this way. More generally, we have the following result which is easily proved (cf. **[10]**).

PROPOSITION 9.1. *The set of Λ-mixed Hodge structures whose weight-graded quotients are canonically isomorphic to those of V is naturally isomorphic to $G_\Lambda \backslash G / F^0 G$. The double coset of $g \in G$ corresponds to the mixed Hodge structure*

$$V = (V_\Lambda g, (V_{\Lambda\otimes\mathbb{Q}} g, W_\bullet g), (V_{\mathbb{C}}, F^\bullet)). \quad \square$$

This identification can be used to compute extension groups of mixed Hodge structures. Suppose that A and B are Λ-Hodge structures whose underlying Λ module is torsion free. Suppose that the weight of A is greater than that of B. If we take $V = A \oplus B$ then, for $R = \Lambda, \mathbb{C}$,

$$G_R = \operatorname{Hom}_R(A, B).$$

In this case the moduli space of mixed Hodge structures with weight-graded quotients canonically isomorphic to A and B is the group $\operatorname{Ext}^1_{\mathscr{H}}(A, B)$ of extensions of A by B in the category $\mathscr{H}$ of mixed Hodge structures. Applying Proposition (9.1) to the split mixed Hodge structure $A \oplus B$, we obtain the well-known formula for $\operatorname{Ext}^1_{\mathscr{H}}$ (cf. **[10]**).

PROPOSITION 9.2. *With A and B as above, there is a canonical isomorphism*

$$\operatorname{Ext}^1_{\mathscr{H}}(A, B) \approx \frac{\operatorname{Hom}_{\mathbb{C}}(A, B)}{\operatorname{Hom}_\Lambda(A, B) + F^0 \operatorname{Hom}_{\mathbb{C}}(A, B)}.$$

An important special case is where $A = \mathbb{Z}$, $B = \mathbb{Z}(n)$, and $n \geq 1$. (Recall that $\mathbb{Z}(n)$ is the Hodge structure of type $(-n, -n)$ whose lattice is the subgroup $(2\pi i)^n\mathbb{Z}$ of $\mathbb{C}$.) In this case we have

$$\operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}, \mathbb{Z}(n)) \approx \mathbb{C}/\mathbb{Z}(n).$$

Following through the construction, we see that the mixed Hodge structure that corresponds to $\lambda \in \mathbb{C}/\mathbb{Z}(n)$ can be described as follows. Denote the standard basis of $\mathbb{C}^2$ by e_0, e_n. These have type $(0, 0)$, $(-n, -n)$, respectively. The Hodge and weight filtrations on $\mathbb{C}^2$ are defined by

$$W_l\mathbb{C}^2 = \operatorname{span}\{e_j : -2j \leq l\} \quad \text{and} \quad F^p\mathbb{C}^2 = \operatorname{span}\{e_j : -j \geq p\}.$$

The mixed Hodge structure that corresponds to λ has integral basis spanned by the two vectors

$$\begin{pmatrix} 1 & \lambda \\ 0 & (2\pi i)^n \end{pmatrix} \begin{pmatrix} e_0 \\ e_n \end{pmatrix}.$$

In particular, the extension of $\mathbb{Z}$ by $\mathbb{Z}(1)$ that corresponds to $x \in \mathbb{C}^* \approx \mathbb{C}/\mathbb{Z}(1)$ has integral basis spanned by the vectors

$$\begin{pmatrix} 1 & \log x \\ 0 & 2\pi i \end{pmatrix} \begin{pmatrix} e_0 \\ e_1 \end{pmatrix}.$$

A unipotent variation of mixed Hodge structure over a smooth variety X whose weight-graded quotients are canonically isomorphic to $\mathbb{Z}$ and $\mathbb{Z}(m)$, $(m \geq 1)$ will determine a *classifying map* $X \to \operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}, \mathbb{Z}(m))$.

PROPOSITION 9.3. *When $m > 1$ the classifying map is constant. When $m = 1$, a map*

$$X \to \operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}, \mathbb{Z}(1)) \approx \mathbb{C}^*$$

is the classifying map of a good variation of mixed Hodge structure if and only if it is an algebraic function on X.

PROOF. In both cases, the canonical extension of the variation to a good compactification $\overline{X}$ of X is a trivial bundle (8.5), as are the extended Hodge and weight bundles. In the first case, Griffiths' transversality forces the integral lattice to be constant. In the second, the regularity of the connection of the canonical extension corresponds to the classifying map $X \to \mathbb{C}^*$ of the variation having poles at infinity. □

Since there is a canonical isomorphism

$$\operatorname{Ext}^1_{\mathscr{H}}(\Lambda, \Lambda(m)) \approx \mathbb{C}/\Lambda(m),$$

the regulator $K_m(\mathbb{C}) \to \mathbb{C}/\Lambda(m)$ can then be interpreted as a map

$$K_m(\mathbb{C}) \to \operatorname{Ext}^1_{\mathscr{H}}(\Lambda, \Lambda(m)).$$

A motivic description of this regulator in the case when $m = 3$ is given in [**4**].

The regulator $c_2 : K_2(X) \to H^2_{\mathscr{D}}(X, \mathbb{Z}(2))$ also admits a Hodge theoretic interpretation. This time we take our reference Hodge structure V to be the direct sum of $\mathbb{Z}(0)$, $\mathbb{Z}(1)$, and $\mathbb{Z}(2)$. The moduli space of mixed Hodge structures whose weight-graded quotients are canonically isomorphic to these Hodge structures is $H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$, where H denotes the Heisenberg group defined in §6. The bundle projection $H_{\mathbb{Z}}\backslash H_{\mathbb{C}} \to \mathbb{C}^* \times \mathbb{C}^*$ may be interpreted as the map that takes a mixed Hodge structure $V \in H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ to

$$(V/\mathbb{Z}(2), W_2 V) \in \operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}, \mathbb{Z}(1)) \times \operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}(1), \mathbb{Z}(2)) \approx \mathbb{C}^* \times \mathbb{C}^*.$$

We next consider the problem of determining which maps $X \to H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ classify variations of mixed Hodge structure.

PROPOSITION 9.4. *A function $f : X \to H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ is the classifying map of a variation of mixed Hodge structure over X with weight-graded quotients canonically isomorphic to $\mathbb{Z}$, $\mathbb{Z}(1)$, and $\mathbb{Z}(2)$ if and only if*

(i) *f is holomorphic;*
(ii) *the composite $X \xrightarrow{f} H_{\mathbb{Z}}\backslash H_{\mathbb{C}} \to \mathbb{C}^* \times \mathbb{C}^*$ of f with the canonical projection is algebraic;*
(iii) *the map $f : X \to H_{\mathbb{Z}}\backslash H_{\mathbb{C}}$ is a flat section of the bundle $H_{\mathbb{Z}}\backslash H^*_{\mathbb{C}} \to \mathbb{C}^* \times \mathbb{C}^*$.*

PROOF. The first statement corresponds to the fact that the connection on the bundle $\mathscr{V} = \mathbb{V} \otimes_{\mathbb{Z}} \mathscr{O}_X$ is holomorphic. The second follows from (9.3) and the fact that if $\mathbb{V}$ is a variation, then so are $\mathbb{V}/\mathbb{Z}(2)$ and $W_2\mathbb{V}$. The last condition corresponds to Griffiths' transversality. One needs to use the fact that the canonical extension of $\mathbb{V}$ to a good compactification $\overline{X}$ of X is trivial, and that the extended Hodge and weight bundles are also trivial (8.5). □

This result allows us to give an interpretation of the regulator

$$c_2 : K_2(X) \to H^2_{\mathscr{D}}(X, \mathbb{Z}(2))$$

constructed in §6: If f, g are invertible functions on X, then $c_2(\{f, g\})$ is the obstruction to finding a good variation of mixed Hodge structure $\mathbb{V}$ over X with weight-graded quotients $\mathbb{Z}$, $\mathbb{Z}(1)$, $\mathbb{Z}(2)$ and whose subquotients $\mathbb{V}/\mathbb{Z}(2)$ and $W_{-2}\mathbb{V}$ are classified by

$$f : X \to \mathbb{C}^* \approx \operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}, \mathbb{Z}(1)) \quad \text{and} \quad g : X \to \mathbb{C}^* \approx \operatorname{Ext}^1_{\mathscr{H}}(\mathbb{Z}(1), \mathbb{Z}(2)).$$

More on extensions of variations of mixed Hodge structure can be found in **[11]** and **[26]**.

Denote by V_x the variation of mixed Hodge structure over $\mathbb{C} - \{0, 1\}$ which is the extension of $\mathbb{Q}$ by $\mathbb{Q}(1)$ which corresponds to the invertible function $x \in \mathscr{O}^{\times}(\mathbb{C} - \{0, 1\})$. The following proposition is easily verified.

Proposition 9.5. *As a variation of mixed Hodge structure, the nth polylogarithm local system is an extension of* $\mathbb{Q}$ *by the shift* $\left(\mathrm{Sym}^{n-1} V_x\right) \otimes \mathbb{Q}(1)$ *of the* $(n-1)$*st symmetric power of* V_x. □

10. The polylogarithm quotient of $\pi_1(\mathbb{P}^1 - \{0, 1, \infty\})$

The polylogarithm quotient of the fundamental group of $\mathbb{P}^1 - \{0, 1, \infty\}$ is the image of the monodromy representation

$$\pi_1(\mathbb{P}^1 - \{0, 1, \infty\}, x) \to \mathrm{Aut}\, P_x$$

where $P \to \mathbb{P}^1 - \{0, 1, \infty\}$ is the polylogarithm variation of mixed Hodge structure.

Since the monodromy representation of the nth polylogarithm local system is unipotent, it induces a representation

$$\mathbb{Q}\pi_1(\mathbb{C} - \{0, 1\}, x)/J^{n+1} \to \mathrm{GL}_{n+1}(\mathbb{C}).$$

Denote the image of the composite

$$\mathfrak{g}(\mathbb{C} - \{0, 1\}, x) \to \mathbb{Q}\pi_1(\mathbb{C} - \{0, 1\}, x)/J^{n+1} \to \mathrm{GL}_{n+1}(\mathbb{C})$$

by $\mathfrak{p}_n(x)$.

Proposition 10.1. *For each* $x \in \mathbb{C} - \{0, 1\}$, $\mathfrak{p}_n(x)$ *has a natural mixed Hodge structure compatible with its Lie algebra structure. The local system of the* $\mathfrak{p}_n(x)$ *forms a good unipotent variation of mixed Hodge structure over* $\mathbb{C} - \{0, 1\}$. *Finally, these local systems form an inverse system of variations of mixed Hodge structure.*

Proof. The first assertion is an immediate consequence of (7.1) and (8.3). The second is a consequence of (8.4), or alternatively, that this local system is a Hodge quotient of the tautological variation. The last assertion is clear. □

By the construction given in §8, the Malcev Lie algebra of $\pi_1(\mathbb{C} - \{0, 1\}, x)$ is the completion of the free Lie algebra generated by $H_1(\mathbb{C} - \{0, 1\}, \mathbb{C}) \approx \Omega^1(\mathbb{C} - \{0, 1\})^*$. Let X_0, X_1 be the basis of $H_1(\mathbb{C} - \{0, 1\}, \mathbb{Z})$ consisting of the homology classes of the loops σ_0, σ_1 defined in §2. This is dual to the basis $\omega_0/(2\pi i)$, $-\omega_1/(2\pi i)$ of $\Omega^1(\mathbb{C} - \{0, 1\})$, where ω_0, ω_1 are the forms defined in §2. The completed group ring $\mathbb{C}\pi_1(\mathbb{C} - \{0, 1\}, x)^\wedge$ is isomorphic to the completion $\mathbb{C}\langle\langle X_0, X_1\rangle\rangle$ of the free associative algebra generated by X_0, X_1. The set of primitive elements of $\mathbb{C}\langle\langle X_0, X_1\rangle\rangle$ is $\mathfrak{f} = \mathbb{L}(X_0, X_1)^\wedge$, the completion of the free Lie algebra generated by X_0 and

X_1. The isomorphism

$$\mathbb{C}\pi_1(\mathbb{C} - \{0, 1\}, x)^\wedge \to \mathbb{C}\langle\langle X_0, X_1\rangle\rangle$$

is induced by the map $\pi_1(\mathbb{C} - \{0, 1\}, x)$ which takes γ to

$$1 + \int_\gamma \omega + \int_\gamma \omega\omega + \int_\gamma \omega\omega\omega + \cdots$$

where ω is the $\mathfrak{f}$-valued 1-form

$$\omega = \omega_0 X_0 - \omega_1 X_1. \tag{15}$$

PROPOSITION 10.2. *The monodromy representation*

$$\pi_1(\mathbb{C} - \{0, 1\}, x) \to \mathrm{GL}_{n+1}(\mathbb{C})$$

of the polylogarithm local system is induced by the homomorphism $\mathfrak{f} \to gl_{n+1}$ *defined by*

$$X_0 \mapsto \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & & \cdots & & 0 \end{pmatrix}, \qquad X_1 \mapsto \left(\begin{array}{c|cccc} 0 & -1 & 0 & \cdots & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & 0 \end{array}\right).$$

PROOF. This follows since the connection matrix of the polylogarithm local system is

$$\begin{pmatrix} 0 & \omega_1 & 0 & \cdots & 0 \\ & \ddots & \omega_0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & \omega_0 \\ 0 & & \cdots & & 0 \end{pmatrix}$$

which is simply the $\mathfrak{f}$-valued form ω (15) composed with the homomorphism $\mathfrak{f} \to gl_{n+1}(\mathbb{C})$ defined in the statement of the proposition. □

COROLLARY 10.3. *The complex form of the polylogarithm quotient has presentation*

$$\mathfrak{p} = \mathbb{L}(X_0, X_1)^\wedge/(\mathrm{ad}(X_1)[\mathbb{L}, \mathbb{L}])$$

and a topological basis $\{X_0, \mathrm{ad}(X_0)^n(X_1) : n \in \mathbb{N}\}$. □

The generator X_0 of $\mathfrak{p}_\mathbb{C}$ is dual to $\log x$, and $\mathrm{ad}(X_0)^n(X_1)$ is dual to $(-1)^n \ln_{n+1} x$.

11. Motivic description of the polylogarithm variation

In this section we give a motivic description of the polylogarithm variation. This description goes back to Deligne.

First suppose that X is a topological space. Denote the space of paths $\gamma : [0, 1] \to X$ by PX. There is a canonical projection $PX \to X \times X$ that takes a path γ to its endpoints $(\gamma(0), \gamma(1))$. Denote the fiber of this map over (x, y) by $P_{x,y}X$.

The group $H_0(P_{x,y}X)$ is the free abelian group on the set of homotopy classes of paths in X from x to y. These form a local system

$$\left\{H_0(P_{x,y}X)\right\}_{(x,y)} \to X \times X. \tag{16}$$

When $x = y$ there is a canonical isomorphism $H_0(P_{x,y}X; \mathbb{Q}) \approx \mathbb{Q}\pi_1(X, x)$. This has a filtration given by the powers of the augmentation ideal J. This filtration extends to a flat filtration of the local system (16). Denote the completion of $H_0(P_{x,y}X; \mathbb{Q})$ in the corresponding topology by $H_0(P_{x,y}X; \mathbb{Q})^\wedge$.

THEOREM 11.1 **[29]**. *If X is a smooth algebraic variety, the local system*

$$\left\{H_0(P_{x,y}X; \mathbb{Q})^\wedge\right\}_{(x,y)} \to X \times X$$

underlies a good variation of mixed Hodge structure whose fiber over (x, x) is the canonical mixed Hodge structure on $\mathbb{Q}\pi_1(X, x)^\wedge$. □

We call this the *canonical variation of mixed Hodge structure* associated to X. Although this construction appears to be outside the domain of algebraic geometry, it can be made motivic. There are several equivalent ways of doing this. One can be found in **[17]**; the other is a construction in topology of a cosimplicial model of PX and the fibration $PX \to X \times X$. It is called the *cobar construction* and dates back to the paper **[1]** of F. Adams. It makes sense for varieties over any base, as was noted by Wojtkowiak **[45]**. Briefly, it is the cosimplicial space

$$P^\bullet = X^{\Delta[1]_\bullet}$$

where $\Delta[1]_\bullet$ is the standard simplicial model of the unit interval. The projection $PX \to X \times X$ corresponds to the map $X^{\Delta[1]_\bullet} \to X^{\partial\Delta[1]_\bullet} = X^2$ induced by the inclusion of the boundary of the interval.

Since the set of m-simplices of $\Delta[1]_\bullet$ is the set of order preserving maps $\{0, 1, \dots, m\} \to \{0, 1\}$, and since there are precisely $m+2$ of these,

$$P^m = X \times X^m \times X.$$

The coface maps $P^m \to P^{m+1}$ are the various diagonals. Applying the de Rham functor to this cosimplicial space yields the bar construction on the de Rham complex of X, which is essentially Chen's complex of iterated integrals on PX (**[12]**, see also **[25**, §§1,2**]**).

Suppose that $X = \overline{X} - D$ is a smooth curve and that $\vec{v}$ is a nonzero tangent vector at a point $P \in D$. Suppose that $x \in X$. Denote by $P_{\vec{v},x}X$ the set of paths $\gamma : [0, 1] \to \overline{X}$ that have the property that $\gamma'(0) = \vec{v}$, $\gamma(1) = x$, and $\gamma(]0, 1]) \subseteq X$. This space is easily seen to be homotopy equivalent to $P_{z,x}X$ where z is a point in X that is sufficiently close to P in the direction of $\vec{v}$. More generally, one can define $P_{\vec{v}_1,\vec{v}_2}X$ where $\vec{v}_1$ and $\vec{v}_2$ are nonzero tangent vectors to points of D. Deligne defines $\pi_1(X, \vec{v})$ to be the set of path components of $P_{\vec{v},\vec{v}}X$. It is canonically isomorphic to $\pi_1(X, z)$ when $z \in X$ is sufficiently close to P in the direction of $\vec{v}$ (cf. [**17**]).

It is useful to think of the limit mixed Hodge structure of the local system

$$\left\{H_0(P_{z,x}X; \mathbb{Q})^\wedge\right\}_{z\in X} \to X$$

associated to $\vec{v}$ as a mixed Hodge structure on $H_0(P_{\vec{v},x}X; \mathbb{Q})^\wedge$.

If $\mathbb{V} \to X$ is a unipotent variation of mixed Hodge structure, there is a canonical map

$$\Theta_{x,y} : V_x \otimes H_0(P_{x,y}X)^\wedge \to V_y \tag{17}$$

induced by parallel transport. When $x = y$, this is just the monodromy representation $V_x \otimes \mathbb{Q}\pi_1(X, x)^\wedge \to V_x$.

Proposition 11.2. *If $\mathbb{V} \to X$ is a unipotent variation of mixed Hodge structure over a smooth curve X and if $x, y \in X$, then the map* (17) *is a morphism of mixed Hodge structures.*

This is easily proved using (8.3) and (8.4): First show that the map from the trivial variation of type $(0, 0)$ over $X \times X$ to the unipotent variation over $X \times X$ whose fiber over (x, y) is

$$\mathrm{Hom}_{\mathbb{Q}}\left(H_0(P_{x,y}X)^\wedge, \mathrm{Hom}_{\mathbb{Q}}(V_x, V_y)\right)$$

given by the map (17) is a morphism of variations of mixed Hodge structure.

Denote the tangent vector $\partial/\partial z$ of $\mathbb{P}^1$ at 0 by $\vec{v}$. Denote the nth polylogarithm variation over $\mathbb{C} - \{0, 1\}$ by V. Then the parallel transport map

$$V_{\vec{v}} \otimes H_0(P_{\vec{v},x}\mathbb{C} - \{0, 1\})^\wedge \to V_x$$

is a morphism of mixed Hodge structure. By (7.2),

$$V_{\vec{v}} = \mathbb{Q}(0) \oplus \mathbb{Q}(1) \oplus \cdots \oplus \mathbb{Q}(n),$$

so there is a canonical inclusion of mixed Hodge structures $\mathbb{Q}(0) \to V_{\vec{v}}$. Applying the parallel transport map to the generator e_0 of $\mathbb{Q}(0)$ yields a canonical morphism of mixed Hodge structure

$$H_0(P_{\vec{v},x}\mathbb{C}-\{0,1\})^\wedge \to V_x. \tag{18}$$

Denote by $\hat{\sigma}_0$ and $\hat{\sigma}_1$ the elements of $\pi_1(\mathbb{C}-\{0,1\},\vec{v})$ obtained from the generators σ_0, σ_1 of $\pi_1(\mathbb{C}-\{0,1\},\frac{1}{2})$ defined in §2 by moving the base point back along the interval $[0,\frac{1}{2}]$.

THEOREM 11.3. *The polylogarithm variation is the quotient of the variation of mixed Hodge structure*

$$\left\{H_0(P_{\vec{v},x}\mathbb{C}-\{0,1\};\mathbb{Q})^\wedge\right\}_{x\in\mathbb{C}-\{0,1\}} \to \mathbb{C}-\{0,1\}$$

whose fiber over $\vec{v}$ is the quotient of $\mathbb{Q}\pi_1(\mathbb{C}-\{0,1\},\vec{v})^\wedge$ by the right ideal generated by

$$\hat{\sigma}_0 - 1, \quad J(\hat{\sigma}_1 - 1);$$

here J denotes the augmentation ideal of $\mathbb{Q}\pi_1(\mathbb{C}-\{0,1\},\vec{v})^\wedge$.

SKETCH OF PROOF. The copy of $\mathbb{Q}(0)$ in V, the nth polylogarithm variation, is spanned by the vector $e_0 \in V$. Since (18) is a right $\pi_1(\mathbb{C}-\{0,1\},\vec{v})$-module map, it follows that its kernel is the right ideal of $\mathbb{Q}\pi_1(\mathbb{C}-\{0,1\},\vec{v})^\wedge$ of elements that annihilate e_0. Since the monodromy representation

$$V_x \otimes \mathbb{Q}\pi_1(\mathbb{C}-\{0,1\},x)^\wedge \to V_x$$

is flat as x varies over $\mathbb{C}-\{0,1\}$, it follows from (7.2) that $\pi_1(\mathbb{C}-\{0,1\},\vec{v})$ acts on $V_{\vec{v}}$ via the matrices in (2.2) with respect to the $\mathbb{Q}$ basis $e_0, 2\pi i e_1, \ldots (2\pi i)^n e_n$ of $V_{\vec{v}}$. This implies that (18) is surjective and that its kernel is generated by $\hat{\sigma}_0 - 1$, $J(\hat{\sigma}_1 - 1)$, J^n. The result now follows by taking inverse limits. □

REFERENCES

1. F. Adams, *On the cobar construction*, Colloque de Topologie Algébrique (Louvain, 1956), George Thone, Liège; Masson, Paris, 1957, pp. 81–87.
2. A. Beilinson, *Higher regulators and values of L-functions of curves*, Funct. Anal. Appl. **14** (1980), 116–118.
3. ———, *Higher regulators and values of L-functions*, Sovrem. Probl. Mat. Fund. Naprav. **24** (1984), 181–238; English transl. in J. Soviet Math. **30** (1985), 2036–2070.
4. A. Beilinson, A. Goncharov, V. Schechtman, and A. Varchenko, *Aomoto dilogarithms, mixed Hodge structures and motivic cohomology of pairs of triangles in the plane*, in The Grothendieck Festschrift : A Collection of Articles Written in Honor of the 60th Birthday of Alexander Grothendieck (P. Cartier et al., eds.), Birkhäuser, Boston, 1990.
5. A. Beilinson, R. MacPherson, and V. Schechtman, *Notes on motivic cohomology*, Duke Math. J. **54** (1987), 679–710.
6. P. Blanc, *Sur la cohomologie continue des groupes localement compacts,* Ann. Sci. École Norm. Sup. (4) **12** (1979), 137–168.
7. S. Bloch, *Applications of the dilogarithm function in algebraic K-theory and algebraic geometry*, Internat. Sympos. Algebraic Geometry (Kyoto), 1977, 103–114.
8. ———, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, unpublished manuscript, 1978.

9. A. Borel, *Cohomologie de* SL_n *et valeurs de fonctions de zeta*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **7** (1974), 613–636.
10. J. Carlson, *The geometry of the extension class of a mixed Hodge structure*, Proc. Sympos. Pure Math., vol. 46, Amer. Math. Soc., Providence, RI, 1987, pp. 199–222.
11. J. Carlson and R. Hain, *Extensions of variations of mixed Hodge structure*, Astérisque **179–180** (1989), 39–65.
12. K.-T. Chen, *Iterated integrals*, Bull. Amer. Math. Soc. (N. S.) **83** (1977), 831–879.
13. P. Deligne, *Équations différentielles à points singuliers réguliers*, Lecture Notes in Math., vol. 163, Springer-Verlag, Berlin and New York, 1970.
14. ______, *Theorie de Hodge*. II, Inst. Hautes Études Sci. Publ. Math. **40** (1971), 5–58.
15. ______, *Le symbole modéré*, unpublished notes, 1979.
16. ______, *Letter to Spencer Bloch*, April 3, 1984.
17. ______, *Le groupe fondamental de la droite projective moins trois points*, Galois groups over $\mathbb{Q}$ (Proc. Workshop, 1987, K. Ribet, ed.) Springer-Verlag, New York, 1989.
18. J. Dupont, *The dilogarithm as a characteristic class for flat bundles*, J. Pure Appl. Algebra **44** (1987), 137–164.
19. J. Dupont, R. Hain, and S. Zucker, *Regulators and characteristic classes of flat bundles*, preprint, 1992.
20. J. Dupont and C.-H. Sah, *Scissors congruences*. II, J. Pure Appl. Algebra **24** (1982), 159–195.
21. A. Goncharov, *Geometry of configurations, polylogarithms and motivic cohomology*, MPI preprint, 1991.
22. D. Grayson, *Weight filtrations in algebraic K-theory*, these Proceedings, vol. 1, pp. 207–237.
23. R. Hain, *On a generalization of Hilbert's 21st problem*, Ann. Sci. École Norm. Sup. (4) **19** (1986), 609–627.
24. R. Hain, *The geometry of the mixed Hodge structure in the fundamental group*, Proc. Sympos. Pure Math., vol. 46, Amer. Math. Soc., Providence, RI, 1987, pp. 247–282.
25. ______, *The de Rham homotopy theory of complex algebraic varieties*. I, K-Theory **1** (1987), 271–324.
26. ______, *Algebraic cycles and extensions of variations of mixed Hodge structure*, Proc. Sympos. Pure Math., vol. 53, Amer. Math. Soc., Providence, RI, 1991, pp. 175–221.
27. R. Hain and R. MacPherson, *Higher logarithms*, Illinois J. Math. **34** (1990), 392–475.
28. R. Hain and J. Yang, *Higher logarithms and Chern classes*, in preparation.
29. R. Hain and S. Zucker, *Unipotent variations of mixed Hodge structure*, Invent. Math. **88** (1987), 83–124.
30. ______, *A Guide to unipotent variations of mixed Hodge structure*, Hodge Theory (Proc. U.S. Spain Workshop, Sant Cugat, Spain, 1985), Lecture Notes in Math., vol. 1246, Springer-Verlag, Berlin and New York, 1987.
31. L. Lewin: *Polylogarithms and associated functions*, North-Holland, Amsterdam, 1981.
32. ______, *Structural Properties of Polylogarithms*, Math. Surveys Monographs, vol. 37, Amer. Math. Soc., Providence, RI, 1991.
33. J. Morgan, *The algebraic topology of smooth algebraic varieties*, Inst. Hautes Études Sci. Publ. Math. **48** (1978), 137–204; correction, **64** (1986), 185.
34. J. Milnor, *Introduction to algebraic K-theory*, Ann. of Math. Stud., no. 72, Princeton Univ. Press, Princeton, NJ, 1971.
35. ______, *Hyperbolic geometry: The first 150 years*, Bull. Amer. Math. Soc. (N. S.) **6** (1982), 9–24.
36. D. Quillen, *Rational homotopy theory*, Ann. of Math. (2) **90** (1969), 205–295.
37. D. Ramakrishnan, *A regulator for curves via the Heisenberg group*, Bull. Amer. Math. Soc. (N. S.) **5** (1981), 191–195.
38. ______, *On the monodromy of higher logarithms*, Proc. Amer. Math. Soc. **85** (1982), 596–599.
39. ______, *Analogs of the Bloch-Wigner function for higher polylogarithms*, Applications of Algebraic K-theory to Algebraic Geometry and Algebraic Number Theory, Part I, Contemp. Math., vol. 55, Amer. Math. Soc., Providence, RI, 1986, pp. 371–376.

40. ______, *Regulators, algebraic cycles, and values of L-functions*, Contemp. Math., vol. 83, Amer. Math. Soc., Providnece, RI, 1989, pp. 183–310.
41. J. Steenbrink and S. Zucker, *Variations of mixed Hodge structure*. I. Invent. Math. **80** (1985), 489–542.
42. A. Suslin, *Homology of* GL_n, *characteristic classes and Milnor K-theory*, Algebraic K-Theory, Number Theory, Geometry and Analysis (Proc. Bielefeld 1982), Lecture Notes in Math., vol. 1046, Springer-Verlag, Berlin and New York, pp. 357–375.
43. ______, *Algebraic K-theory of fields*, Proc. Internat. Congr. Math., 1986, Amer. Math. Soc., 1987, pp. 222–244.
44. Z. Wojtkowiak, *A construction of analogs of the Bloch-Wigner function*, Math. Scand. **65** (1989), 140–142.
45. ______, *Cosimplicial objects in algebraic geometry*, preprint.
46. J. Yang, *Algebraic K-groups of number fields and the Hain-MacPherson trilogarithm*, Ph.D. thesis, Univ. Washington, 1991.
47. ______, *On the real cohomology of arithmetic groups and the rank conjecture for number fields*, Ann. Sci. École Norm. Sup. (4) 25 (1992).
48. D. Zagier, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. **286** (1990), 613–624.
49. D. Zagier, *Polylogarithms, Dedekind zeta functions, and the algebraic K-theory of fields*, Arithmetic and Algebraic Geometry (G. van der Geer et al., eds.), Progr. Math. vol. 89, Birkhäuser, Boston, 1991, pp. 391–430.

Duke University, Durham, North Carolina 27706
E-mail address: `hain@math.duke.edu`

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Polylogarithms and Motivic Galois Groups

A. B. GONCHAROV

Sections 1 and 2 of this paper are an enlarged version of the lecture given at the AMS conference "Motives" in Seattle, July 1991. More details can be found in **[G2]**.

My aim is to formulate a precise conjecture about the structure of the Galois group $\mathrm{Gal}(\mathscr{M}_T(F))$ of the category $\mathscr{M}_T(F)$ of mixed Tate motivic sheaves over $\mathrm{Spec}\, F$, where F is an arbitrary field. This conjecture implies (and in fact is equivalent to) a construction of complexes $\Gamma(F, n)_{\mathbb{Q}}$ that should satisfy all the Beilinson-Lichtenbaum axioms modulo torsion.

In particular, we get a hypothetical description of $K_n(F)\otimes\mathbb{Q}$ by generators and relations that generalize the definition of Milnor's K-groups. This can be considered as a (hypothetical) generalization of the computation of $K_2(F)$ by Matsumoto-Moore **[Ma, Mo]** and $K_3(F)$ by Suslin (**[S3]**, see also **[S2]**) to all Quillen K-groups of arbitrary fields. In the case when F is a number field, this together with the Borel theorem **[Bo2]** would imply

ZAGIER'S CONJECTURE **[Z1]**. *The value of the Dedekind zeta-function $\zeta_F(s)$ of an arbitrary number field F at the point n is expressed by a determinant whose entries are rational linear combinations of values of the* classical *n-logarithms at* (*the complex embedding of*) *some elements of this field.*

In §3 I give a proof of Zagier's conjecture in the case $n = 3$.

The classical polylogarithms invented by Leibniz and Johannes Bernoulli almost 300 years ago (see **[Lei]**) are defined on the unit disc $|z| \leq 1$ by the absolutely convergent series

$$Li_n(z) = \sum_{k=1}^{\infty} \frac{z^k}{n^k}.$$

These can be continued analytically to multi-valued functions on $\mathbb{C}P^1\setminus\{0, 1, \infty\}$. Their properties, including the differential and functional

1991 *Mathematics Subject Classification.* Primary 11R70, 14F99, 19D99.

This paper is in final form and no version of it will be submitted for publication elsewhere.

0082-0717/94 $1.00 + $.25 per page

equations, play the key role in all our considerations. However, the special role of the *projective line* and classical polylogarithms in the theory of mixed Tate motives remains absolutely mysterious. Formulas that led me to the conjectures about $\Gamma(F, n)_{\mathbb{Q}}$ and $\mathrm{Gal}(\mathscr{M}_T(F))$ are discussed in §4.

In §5 I shall construct explicitly a regulator map r_3 from the motivic complex $\Gamma(X; 3)_{\mathbb{Q}}$ attached to any algebraic variety over $\mathbb{C}$ to the third Deligne complex of $X(\mathbb{C})$. (For a generalization of this construction to motivic complexes $\Gamma(X; n)_{\mathbb{Q}}$ see **[G3]**). Then an explicit formula for the universal motivic Chern class $c_3 \in H^6_{\mathscr{M}}(\mathrm{BGL}_3(F)_\bullet, \mathbb{Q}(3))$ will be given. Applying the regulator we get a realization of c_3 in the real Deligne cohomology. I need the last result in order to complete the proof of Zagier's conjecture.

I would like to express my deep gratitude to Sasha Beilinson and Don Zagier for many valuable discussions, suggestions, and interest. I am also grateful to Christophe Soulé for useful remarks on the previous version of the work. This paper was written during my stay at Harvard University and completed at MIT. I am grateful to both institutions for their hospitality and to Sarah Warren for excellent printing of the manuscript and pictures. Finally, I am grateful to Jan Nekovář for an extremely careful reading of the manuscript.

1. Conjectures

First of all we need to explain how to think about $\mathrm{Gal}(\mathscr{M}_T(F))$. So for convenience of the reader I reproduce basic definitions from **[BD]**.

1. Mixed Tate categories (**[BD]**, see also **[BMS, B2, D2]**). A mixed Tate category is a Tannakian $\mathbb{Q}$-category $\mathscr{M}$ together with a fixed invertible object $\mathbb{Q}(1)_{\mathscr{M}}$ such that

(a) any simple object in $\mathscr{M}$ is isomorphic to

$$\mathbb{Q}(m)_{\mathscr{M}} := \mathbb{Q}(1)_{\mathscr{M}}^{\otimes m}, \qquad m \in \mathbb{Z};$$

(b) $\dim \mathrm{Hom}_{\mathscr{M}}(\mathbb{Q}(o)_{\mathscr{M}}, \mathbb{Q}(m)_{\mathscr{M}}) = \delta_{o,m}$,

$$\mathrm{Ext}^1_{\mathscr{M}}(\mathbb{Q}(o)_{\mathscr{M}}, \mathbb{Q}(m)_{\mathscr{M}}) = 0 \quad \text{for } m \leq 0.$$

(I recall that "Tannakian" means, in particular, that there is a $\otimes$-product in $\mathscr{M}$; the functor $\mathscr{F} \mapsto \mathscr{F} \otimes \mathbb{Q}(1)_{\mathscr{M}}$ is an equivalence of categories.)

A Tate functor $F: \mathscr{M}_1 \to \mathscr{M}_2$ between mixed Tate categories is an exact $\otimes$ functor such that $F(\mathbb{Q}(1)_{\mathscr{M}_1}) = \mathbb{Q}(1)_{\mathscr{M}_2}$. Sometimes I shall write $\mathbb{Q}(m)$ instead of $\mathbb{Q}(m)_{\mathscr{M}}$.

An object $\mathscr{F}$ of $\mathscr{M}$ has a canonical finite increasing filtration $\cdots \subset \mathscr{F}_{\leq i} \subset \mathscr{F}_{\leq i+1} \subset \cdots$ such that $\mathscr{F}_i := \mathscr{F}_{\leq i}/\mathscr{F}_{\leq i-1}$ is isomorphic to a direct sum of $\mathbb{Q}(-i)$'s. There is a canonical fiber functor to the category of finite-dimensional graded $\mathbb{Q}$-vector spaces $\omega_{\mathscr{M}}: \mathscr{M} \to \mathrm{Vect}^\bullet_{\mathbb{Q}}$:

$$\omega_{\mathscr{M}}(\mathscr{F}_i) := \mathrm{Hom}_{\mathscr{M}}(\mathbb{Q}(-i), \mathscr{F}_i), \qquad \omega_{\mathscr{M}}(\mathscr{F}) := \bigoplus_i \omega_{\mathscr{M}}(\mathscr{F})_i.$$

Let $L(\mathscr{M})$ be the space of all derivations of $\omega_{\mathscr{M}}$; an element $\varphi \in L(\mathscr{M})$ is a natural endomorphism of the functor $\omega_{\mathscr{M}}$ such that $\varphi_{\mathscr{F}\otimes G} = \varphi_{\mathscr{F}} \otimes \mathrm{id}_{\omega(G)} + \mathrm{id}_{\omega(\mathscr{F})} \otimes \varphi_G$. Then $L(\mathscr{M})$ is canonically equipped with the structure of a graded pro-Lie algebra: $L(\mathscr{M}) = \bigoplus L(\mathscr{M})_i$, where

$$L(\mathscr{M})_i := \{\varphi \in L(\mathscr{M}) \mid \varphi(\mathscr{F})\colon \omega_{\mathscr{M}}(\mathscr{F})_\bullet \to \omega_{\mathscr{M}}(\mathscr{F})_{\bullet+i}\}.$$

(Recall that "graded pro-Lie algebra" is a projective limit of finite-dimensional graded Lie algebras.) It is easy to prove that $L(\mathscr{M})_i = 0$ for $i \geq 0$. Such Lie algebras are called mixed Tate pro-Lie algebras. For any mixed Tate Lie algebra L the category L-mod of finite-dimensional graded continuous L-modules is a mixed Tate category. The object $\mathbb{Q}(1)$ in this category is $\mathbb{Q}$ with trivial action of L placed in degree -1; the fiber functor $\omega\colon L\text{-mod} \to \mathrm{Vect}^\bullet_{\mathbb{Q}}$ is just forgetting of L-action functor. For any mixed Tate category $\mathscr{M}$ the fiber functor $\omega_{\mathscr{M}}$ lifts canonically to the Tate functor $\omega_{\mathscr{M}}\colon \mathscr{M} \to L(\mathscr{M})$-mod. It is easy to prove that $\omega_{\mathscr{M}}$ is an equivalence of categories. Note that any Tate functor $F\colon \mathscr{M}_1 \to \mathscr{M}_2$ commutes with ω's and so defines the morphism $F_\bullet\colon L(\mathscr{M}_2)_\bullet \to L(\mathscr{M}_1)_\bullet$ of corresponding mixed Tate algebras. For an object $\mathscr{F} \in \mathscr{M}$

$$H^\bullet_{\mathscr{M}}(\mathscr{F}) := \mathrm{Ext}^\bullet_{\mathscr{M}}(\mathbb{Q}(o), \mathscr{F}) = H^\bullet(L(\mathscr{M})_\bullet, \omega_{\mathscr{M}}(\mathscr{F})). \tag{1.1}$$

REMARK. Let $G(\mathscr{M})$ be a pro-unipotent group with the Lie algebra $L(\mathscr{M})$. Note that $G(\mathscr{M})$ acts on any continuous $L(\mathscr{M})$-module. There is a semidirect product $G_m \times G(\mathscr{M})$ where G_m is the multiplicative group and the action of G_m on $G(\mathscr{M})$ provides the grading on $L(\mathscr{M})$-modules. So the category of finite-dimensional graded continuous $L(\mathscr{M})$-modules is canonically isomorphic to the category of $G_m \times G(\mathscr{M})$ finite-dimensional continuous modules.

2. The motivic Lie algebra $L(F)_\bullet$. Beilinson conjectured **[B1]** that for arbitrary field F there exists a mixed Tate category $\mathscr{M}_T(F)$ of mixed motivic Tate sheaves over $\operatorname{Spec} F$ such that

$$\mathrm{Ext}^i_{\mathscr{M}_T(F)}(\mathbb{Q}(o), \mathbb{Q}(n)) \cong \mathrm{gr}^n_\gamma K_{2n-i}(F)_{\mathbb{Q}} \tag{1.2}$$

where γ is the γ-filtration on K-groups (see **[So]**) and for an abelian group A we put $A_{\mathbb{Q}} := A \otimes \mathbb{Q}$. Let $L(F)_\bullet = \bigoplus_{n=1}^\infty L(F)_{-n}$ be the corresponding mixed Tate Lie algebra. Its cohomology $H^i(L(F)_\bullet, \mathbb{Q})$ (we shall often use a notation $H^i(L(F)_\bullet)$) has a natural grading by positive integers because $L(F)_\bullet$ itself is a negatively graded Lie algebra (see Example 1.2). Let us denote by $H^i_{(n)}(L(F)_\bullet)$ the part of degree n with respect to this grading. Then axiom (1.2) means that

$$H^i_{(n)}(L(F)_\bullet) = \mathrm{gr}^n_\gamma K_{2n-i}(F)_{\mathbb{Q}}. \tag{1.3}$$

Moreover, this isomorphism should be compatible with natural products on $H^*(L(F)_\bullet)$ and $K_*(F)$. It also should be functorial with respect to embeddings of fields $i: F \hookrightarrow E$. (More precisely the corresponding morphism of schemes $\tilde{i}$: $\operatorname{Spec} E \to \operatorname{Spec} F$ should lift to a morphism of mixed Tate categories $\tilde{i}^*: \mathscr{M}_T(F) \to \mathscr{M}_T(E)$ commuting with the fiber functors and so provide us with a homomorphism of the Lie algebras $\tilde{i}_\bullet: L(E)_\bullet \to L(F)_\bullet$.) The Galois group $\operatorname{Gal}(\mathscr{M}_T(F))$ is by definition the semidirect product $G_m \times G(\mathscr{M}_T(F))$ (see above).

This conjecture gives a new point of view on algebraic K-theory. Let me give some examples demonstrating how powerful it is.

EXAMPLE 1.1. $H^i(L(F)_\bullet) = 0$ for $i < 0$ and $H^0_{(n)}(L(F)_\bullet) = 0$ for $n > 0$. So $\operatorname{gr}^n_\gamma K_m(F)_\mathbb{Q} = 0$ for $m \geq 2n > 0$. But this is just the Beilinson-Soulé conjecture.

EXAMPLE 1.2. The degree n part of the standard cochain complex $(\Lambda^\bullet(L(F)^\vee_\bullet), \partial)$ of the Lie algebra $L(F)_\bullet$ forms a subcomplex $(\Lambda^\bullet_{(n)}(L(F)^\vee_\bullet, \partial)$:

$$L^\vee_{-n} \xrightarrow{\partial} \cdots \xrightarrow{\partial} L^\vee_{-2} \otimes \Lambda^{n-2} L^\vee_{-1} \xrightarrow{\partial} \Lambda^n L^\vee_{-1} \tag{1.4}$$

(we write L_{-n} instead of $L(F)_{-n}$). In particular, it is concentrated in degrees $[1, n]$. $(\Lambda^m_{(n)}(L(F)^\vee_\bullet) = 0$ for $m > n$ because $L(F)_\bullet$ is graded by strictly negative integers.) So according to (1.3) $\operatorname{gr}^n_\gamma K_m(F)_\mathbb{Q} = 0$ for $m < n$. This is a well-known theorem in K-theory that follows from results of Suslin **[S1]** (see **[So]**).

EXAMPLE 1.3 (Relation with Milnor K-theory). Applying (1.3) in the simplest case $i = n = 1$ we get

$$H^1_{(1)}(L(F)_\bullet) \overset{\text{def}}{=} L(F)^\vee_{-1} \overset{(1.3)}{=} K_1(F)_\mathbb{Q} = F^*_\mathbb{Q}. \tag{1.5}$$

Here $W \to W^\vee$ is the duality between $\varprojlim$ and $\varinjlim$ of finite-dimensional $\mathbb{Q}$-vector spaces $(W^\vee)^\vee = W$. The structure of a $\varinjlim$ of finite-dimensional $\mathbb{Q}$-vector spaces on $F^*_\mathbb{Q}$ is defined as follows. Let $\mathbb{Z}[P^1_F]$ be the free abelian group generated by symbols $\{x\}$ where x runs through all F-points of the projective line P^1. Let us denote by $\mathscr{R}_1(F)$ the subgroup generated by symbols $\{\infty\}$, $\{0\}$, $\{xy\} - \{x\} - \{y\}$ $(x, y \in F^*)$. Then there is a canonical isomorphism

$$\mathbb{Z}[P^1_F]/\mathscr{R}_1(F) \to F^*; \qquad \{x\} \mapsto x; \quad \{\infty\}, \{0\} \mapsto 1.$$

Both $\mathbb{Q}[P^1_F] := \mathbb{Z}[P^1_F] \otimes \mathbb{Q}$ and $\mathscr{R}_1(F)_\mathbb{Q}$ are $\varinjlim$ of finite-dimensional $\mathbb{Q}$-vector spaces; so we get the same structure on $F^*_\mathbb{Q}$.

Now look at the degree 2 part of the cochain complex of $L(F)_\bullet$. (We use (1.5)):

$$L^\vee_{-2} \xrightarrow{\partial} \Lambda^2 F^*_\mathbb{Q}.$$

According to (1.3) $\operatorname{Coker} \partial = K_2(F)_\mathbb{Q}$. So by the Matsumoto-Moore theorem, $\operatorname{Im} \partial$ is generated by symbols $(1 - x) \wedge x$. Hence, we get a homomorphism of complexes, where $\delta\colon \{x\} \mapsto (1-x)\wedge x$

$$\begin{array}{ccc} \mathbb{Q}[P^1_F] & \xrightarrow{\delta} & \bigwedge^2 F^*_{\mathbb{Q}} \\ \downarrow & & \| \\ L(F)^\vee_{-2} & \xrightarrow{\partial} & \bigwedge^2 F^*_{\mathbb{Q}}. \end{array}$$

Further,

$$\partial\left(L^\vee_{-2}\otimes \bigwedge^{n-2} L^\vee_{-1}\right) = \partial(L^\vee_{-2})\wedge \bigwedge^{n-2} L^\vee_{-1};$$

so

$$H^n_{(n)}(L(F)_\bullet) = K^M_n(F)_{\mathbb{Q}} := \frac{\bigwedge^n F^*}{(1-x)\wedge x\wedge \bigwedge^{n-2} F^*}\otimes \mathbb{Q}.$$

(Here $K^M_*(F)$ is the Milnor ring of the field F; see **[M]**). Comparing with (1.3) we obtain $\mathrm{gr}^n_\gamma K_n(F)_{\mathbb{Q}} = K^M_n(F)_{\mathbb{Q}}$. More precisely we get the following: multiplication in $K_*(F)$ induces a map $m\colon K_1(F)\times\cdots\times K_1(F)\to K_n(F)$ that factorizes through a map $s\colon K^M_n(F)\to K_n(F)$

$$\begin{array}{ccc} F^*\times\cdots\times F^* & \xrightarrow{m} & K_n(F) \\ & \searrow \quad K^M_n(F) \quad \nearrow_s & \end{array}$$

Then the composition $K^M_n(F)\to K_n(F)\to \mathrm{gr}^n_\gamma K_n(F)$ is an isomorphism modulo torsion. But this is the well-known theorem of Suslin **[S1]**. (In fact, Suslin proved that it is an isomorphism modulo $(n-1)!$.)

Example 1.4. Complexes $(\bigwedge^\bullet_{(n)}(L(F)^\vee_\bullet), \partial)$ should satisfy the Beilinson-Lichtenbaum axioms modulo torsion (see **[B1]** and **[L1]**).

More precisely, the (hypothetical) properties of the Lie algebra $L(F)_\bullet$ provide most of all axioms: these complexes are concentrated in degrees $[1, n]$ by definition; relation with algebraic K-theory is given by (1.3); the DGA structure of $\bigwedge^\bullet(L(F)^\vee_\bullet)$ gives a morphism of complexes

$$\left(\bigwedge^\bullet_{(n)}(L(F)^\vee_\bullet), \partial\right)\otimes\left(\bigwedge^\bullet_{(m)}(L(F)^\vee_\bullet), \partial\right)\to\left(\bigwedge^\bullet_{(n+m)}(L(F)^\vee_\bullet), \partial\right),$$

and Example 1.3 shows that

$$H^n\left(\bigwedge^\bullet_{(n)}(L(F)^\vee_\bullet)\right) = K^M_n(F)_{\mathbb{Q}}.$$

The only axiom that remains unclear from this point of view is the existence of the transfer

$$\left(\bigwedge^\bullet_{(n)}(L(E)^\vee_\bullet)\right)\to\left(\bigwedge^\bullet_{(n)}(L(F)^\vee_\bullet)\right)$$

for a finite extension of fields $F\subset E$. On the other hand, if we know something about $K_*(F)_{\mathbb{Q}}$, then conjecture (1.3) provides us with some information about the structure of the Lie algebra $L(F)_\bullet$.

EXAMPLE 1.5. Let F be a number field. Then it is well known that $\mathrm{gr}_\gamma^n K_m(F)_\mathbb{Q} \neq 0$ only if $m = 2n-1$. So $H^i(L(F)_\bullet) = 0$ for $i \geq 2$; hence, $L(F)_\bullet$ is a free graded Lie algebra. Further, Borel proved ([**Bo1, Bo2**], see also subsection 2 of §2) that for $m > 1$

$$\dim K_{2m-1}(F)_\mathbb{Q} = d_m := \begin{cases} r_1 + r_2 & \text{if } m \text{ is odd,} \\ r_2 & \text{if } m \text{ is even.} \end{cases} \tag{1.6}$$

So $L(F)_\bullet$ is generated by $(F_\mathbb{Q}^*)^\vee$ in degree -1 and vector spaces of dimension d_m in degrees $-m = -2, -3, \ldots$.

EXAMPLE 1.6. F is a finite field. Then $K_*(F)_\mathbb{Q} = 0$ (see [**Q2**]); so $L(F)_\bullet = 0$. This agrees with the fact that the category $\mathscr{M}_T(F)$ should be semisimple because Frobenius acts on simple objects $\mathbb{Q}(j)$ with different eigenvalues q^{-j}.

Let us denote by F_0 the subfield of constants in a field F (i.e., F_0 is the algebraic closure in F of the prime field).

RIGIDITY CONJECTURE 1.7 (Beilinson). *The canonical map* $K_*(F_0) \to K_*(F)$ *induces an isomorphism* $\mathrm{gr}_\gamma^n K_{2n-1}1(F_0) \xrightarrow{\sim} \mathrm{gr}_\gamma^n K_{2n-1}(F)$ *for* $n \geq 2$.

EXAMPLE 1.8. Now let $\mathrm{char}\, F = p > 0$. Then Example 1.6 together with the rigidity conjecture imply that $\mathrm{gr}_\gamma^n K_{2n-1}(F_0)_\mathbb{Q}$ should be zero for $n \geq 2$. This means that $L(F)_\bullet$ is generated by $(F_\mathbb{Q}^*)^\vee$ sitting in degree -1.

3. The structure of $L(F)_\bullet$. Set

$$I(F)_\bullet := \bigoplus_{n=2}^{\infty} L(F)_{-n}.$$

CONJECTURE 1.9. $I(F)_\bullet$ *is a free graded Lie algebra.*

Our next aim is to construct explicitly the quotient $L_\bullet/[I_\bullet, I_\bullet]$. There is the extension

$$0 \to I_\bullet/[I_\bullet, I_\bullet] \to L_\bullet/[I_\bullet, I_\bullet] \to L_\bullet/I_\bullet \to 0. \tag{1.7}$$

Let $\mathfrak{n}$ be a nilpotent Lie algebra. Then $H_1(\mathfrak{n}) = \mathfrak{n}/[\mathfrak{n}, \mathfrak{n}]$ can be interpreted as a space of generators of $\mathfrak{n}$ (as a Lie algebra) and $H_2(\mathfrak{n})$ as a space of relations between generators; $\mathfrak{n}$ is free if and only if $H_2(\mathfrak{n}) = 0$. If $\mathfrak{n}$ is free then $H_i(\mathfrak{n}) = 0$ for $i \geq 2$.

Returning to (1.7) we see that the left space in (1.6) is just the space of generators of $I_\bullet$. So Conjecture 1.9 together with explicit construction of extension (1.7) will give us, in particular, a complete description of the ideal $I_\bullet$. The quotient $L_\bullet/I_\bullet$ is abelian and as a $\mathbb{Q}$-vector space is isomorphic to $L_{-1}^\vee \cong (F_\mathbb{Q}^*)^\vee$ (see (1.5)). The inclusion $L_{-1} \hookrightarrow L_\bullet$ provides canonical splitting $s\colon L_\bullet/I_\bullet \to L_\bullet/[I_\bullet, I_\bullet]$ of (1.7) as an extension of $\mathbb{Q}$-vector spaces; the action of $L_\bullet$ on $I_\bullet$ gives the action of $L_\bullet/I_\bullet$ on $H_1(I_\bullet)$. Let $H_1^{(-n)}(I_\bullet)$

be the component of grading $-n$ of $H_1(I_\bullet)$. Then to construct $L_\bullet/[I_\bullet, I_\bullet]$ we need the following data:

(1.8a) (i) A graded $\mathbb{Q}$-vector space $H_1(I_\bullet) = \bigoplus_{n=+2}^{\infty} H_1^{(-n)}(I_\bullet)$,

(1.8b) (ii) A map $(F_{\mathbb{Q}}^*)^\vee \wedge (F_{\mathbb{Q}}^*)^\vee \to H_1^{(-2)}(I_\bullet)$
(this will be the commutator $[s(L_\bullet/I_\bullet), s(L_\bullet/I_\bullet)]$),

(1.8c) (iii) Maps $(F_{\mathbb{Q}}^*)^\vee \otimes H_1^{(-(n-1))}(I_\bullet) \to H_1^{(-n)}(I_\bullet)$.

Dualizing (1.8) we get

$$f_2 \colon H^1_{(2)}(I_\bullet) \to \bigwedge^2 F_{\mathbb{Q}}^*, \tag{1.9a}$$

$$f_n \colon H^1_{(n)}(I_\bullet) \to H^1_{(n-1)}(I_*) \otimes F_{\mathbb{Q}}^*. \tag{1.9b}$$

These data will be defined in the next subsections.

4. The groups $\mathcal{R}_n(F)$. Let us define by induction subgroups $\mathcal{R}_n(F) \subset \mathbb{Z}[P_F^1]$, $n \geq 1$. Set

$$\mathcal{B}_n(F) := \mathbb{Z}[P_F^1]/\mathcal{R}_n(F).$$

The subgroup $\mathcal{R}_1(F)$ was already defined in such a way that $\mathcal{B}_1(F) = F^*$:

$$\mathcal{R}_1(F) := (\{x\} + \{y\} - \{xy\}, (x, y \in F^*); \{0\}; \{\infty\}).$$

Consider homomorphisms

$$\mathbb{Z}[P_F^1] \xrightarrow{\delta_n} \begin{cases} \mathcal{B}_{n-1}(F) \otimes F^*, & n \geq 3, \\ \bigwedge^2 F^*, & n = 2; \end{cases}$$

$$\delta_n \colon \{x\} \mapsto \begin{cases} \{x\}_{n-1} \otimes x, & n \geq 3, \\ (1-x) \wedge x, & n = 2; \end{cases} \tag{1.10}$$

$$\delta_n \colon \{\infty\}, \{0\}, \{1\} \mapsto 0.$$

Here $\{x\}_n$ is the projection of $\{x\}$ in $\mathcal{B}_n(F)$. Set

$$\mathcal{A}_n(F) := \operatorname{Ker} \delta_n.$$

Any element $\alpha(t) = \sum n_i \{f_i(t)\} \in \mathbb{Z}[P_{F(t)}^1]$ has a specialization $\alpha(t_0) := \sum n_i \{f_i(t_0)\} \in \mathbb{Z}[P_F^1]$, $t_0 \in P_F^1$. (It is correctly defined even if t_0 is a pole of $f_i(t)$; in this case $f_i(t_0) = \infty \in P_F^1$.)

Definition 1.10. $\mathcal{R}_n(F)$ is generated by elements $\alpha(0) - \alpha(1)$ where $\alpha(t)$ runs through all elements of $\mathcal{A}_n(F(t))$, and also $\{\infty\}$, $\{0\}$.

Lemma 1.11. $\delta_n(\mathcal{R}_n(F)) = 0$.

Proof. See proof of Lemma 1.16 in [G2].

So we get

$$\delta \colon \mathcal{B}_n(F) \to \begin{cases} \mathcal{B}_{n-1}(F) \otimes F^*, & n \geq 3, \\ \bigwedge^2 F^*, & n = 2. \end{cases}$$

Let me give some examples of elements of $\mathscr{R}_n(F)$.

EXAMPLE 1.12. $\{x\} + \{x^{-1}\}$ and $\{x\} + \{1-x\} \in \mathscr{R}_2(F)$. Indeed, $\delta_2(\{x\} + \{x^{-1}\}) = (1-x) \wedge x + (1-x^{-1}) \wedge x^{-1} = 0$ in $\bigwedge^2 F(t)^*$ modulo 2-torsion. On the other hand, $\{x\} + \{x^{-1}\}|_{x=\infty} \in \mathscr{R}_2(F)$ by definition. The same arguments work for $\{x\} + \{1-x\}$.

EXAMPLE 1.13. $\{x\} + (-1)^n\{x^{-1}\} \in \mathscr{R}_n(F)$. Indeed, by induction, $\delta_n(\{x\} + (-1)^n\{x^{-1}\}) = (\{x\} + (-1)^{n-1}\{x\}) \otimes x \in \mathscr{R}_{n-1}(F(t)) \otimes F(t)^*$ and $\{x\} + (-1)^n\{x^{-1}\}|_{x=\infty} \in \mathscr{R}_n(F)$ by definition. In particular, $2 \cdot \{1\} \in \mathscr{R}_{2m}(F)$. (Put $x = 1$, $n = 2m$.) We shall prove in the next subsection that $\{1\} \notin \mathscr{R}_{2m+1}(\mathbb{C})$ (see Example 1.18).

5. Motivation: polylogarithms. The classical n-logarithm can be defined on the unit disk $|z| \leq 1$ by absolutely convergent series

$$Li_n(z) := \sum_{k=1}^{\infty} \frac{z^k}{k^n}.$$

We have

$$\begin{aligned} Li_1(z) &= -\log(1-z), \\ dLi_n(z) &= Li_{n-1}(z)\, d\log z. \end{aligned} \tag{1.11}$$

So using the formula

$$Li_n(z) = \int_0^z Li_{n-1}(w) \frac{dw}{w}$$

we can continue analytically $Li_n(z)$ to a multi-valued analytic function on $\mathbb{C}P^1 \backslash \{0, 1, \infty\}$. However, n-logarithm has a remarkable single-valued version:

$$\mathscr{L}_n(z) := \begin{matrix} \mathrm{Re}(n\colon \text{odd}) \\ \mathrm{Im}(n\colon \text{even}) \end{matrix} \left(\sum_{k=0}^{n} \frac{B_k \cdot 2^k}{k!} \log^k |z| \cdot Li_{n-k}(z) \right), \qquad n \geq 2,$$

$$\mathscr{L}_1(z) := \log|z|.$$

Here $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}, \ldots$ are Bernoulli numbers. Let me note that

$$\mathscr{L}_2(z) = \mathrm{Im}(Li_2(z)) + \arg(1-z)\log|z| \tag{1.12}$$

is the well-known Bloch-Wigner function, and

$$\mathscr{L}_3(z) = \mathrm{Re}(Li_3(z) - \log|z| \cdot Li_2(z) - \tfrac{1}{3}\log^2|z| \log(1-z))$$

was used in **[G1]**. The functions $\mathscr{L}_n(z)$ for arbitrary n were written by Zagier **[Z1]**, who proved the following:

THEOREM 1.14. *$\mathscr{L}_n(z)$ is single valued and continuous on $\mathbb{C}P^1$ for $n \geq 2$.*

It is clear that then $\mathscr{L}_n(z)$ is real-analytic on $\mathbb{C}P^1 \backslash \{0, 1, \infty\}$.

The Hodge-theoretical interpretation of these functions was given by Beilinson and Deligne (see, e.g., **[D2]**).

Any real-valued function, and in particular $\mathscr{L}_n(z)$, defines a homomorphism

$$\widetilde{\mathscr{L}}_n : \mathbb{Z}[P^1_{\mathbb{C}}] \to \mathbb{R}, \quad \{z\} \mapsto \mathscr{L}_n(z).$$

THEOREM-MOTIVATION 1.15. $\widetilde{\mathscr{L}}_n(\mathscr{R}_n(\mathbb{C})) = 0$.

PROOF. Let us first prove the theorem for $n = 2$.

LEMMA 1.16. *Let* $\alpha(t) = \sum n_i\{f_i(t)\} \in \mathbb{Z}[P^1_{\mathbb{C}(t)}]$. *If*

$$\delta_2\alpha(t) := \sum n_i(1 - f_i(t)) \wedge f_i(t) = 0$$

in $\bigwedge^2 \mathbb{C}(t)^*$ *then* $d(\sum n_i \mathscr{L}_2(f_i(z))) = 0$.

It follows immediately from the lemma that $\widetilde{\mathscr{L}}_2(\alpha(0) - \alpha(1)) = 0$ and so $\widetilde{\mathscr{L}}_2(\mathscr{R}_2(\mathbb{C})) = 0$. □

PROOF OF LEMMA 1.16. Let us consider the diagram

$$\text{(1.13)} \qquad \begin{array}{ccc} \mathbb{Z}[P^1_{\mathbb{C}(t)}] & \xrightarrow{\delta_2} & \bigwedge^2 \mathbb{C}(t)^* \\ \widetilde{\mathscr{L}}_2 \downarrow & & \downarrow r_2 \\ S^0(\mathbb{C}P^1) & \xrightarrow{d} & S^1(\mathbb{C}P^1) \end{array}$$

$$r_2(f \wedge g) := -\log|f| d\arg g + \log|g| d\arg f.$$

Here $S^i(\mathbb{C}P^1)$ is the space of smooth i-forms each defined on an appropriate Zariski open domain of $\mathbb{C}P^1$ $(= C^\infty$ i-forms at the generic point of $\mathbb{C}P^1)$.

The formula

$$d\mathscr{L}_2(z) = -\log|1 - z| d\arg z + \log|z| d\arg(1 - z)$$

provides the commutativity of the diagram (1.13). So if $\alpha(t) \in \mathscr{A}_2(\mathbb{C}(t))$, then

$$0 = r_2 \circ \delta_2(\alpha(t)) = d \circ \widetilde{\mathscr{L}}_2(\alpha(t)) \overset{\text{def}}{=} d\left(\sum n_i \mathscr{L}_2(f_i(z))\right).$$

Set

$$\widehat{\mathscr{L}}_n(z) = \begin{cases} \mathscr{L}_n(z), & n \text{ odd}, \\ i\mathscr{L}_n(z), & n \text{ even}. \end{cases}$$

Then we have for $n \geq 3$

(1.14)

$$d\widehat{\mathscr{L}}_n(z) = \widehat{\mathscr{L}_{n-1}}(z) d(i \arg z) - \sum_{k=2}^{n-2} \frac{B_k \cdot 2^k}{k!} \log^{k-1}|z| \cdot \widehat{\mathscr{L}_{n-k}}(z) \cdot d\log|z|$$
$$- \frac{B_{n-1} \cdot 2^{n-1}}{(n-1)!} \log^{n-2}|z| (\log|z| d\log|1 - z| - \log|1 - z| d\log|z|).$$

It is interesting that in this formula the same coefficients appear as in (1.12).

The proof of the theorem in the case $n \geq 3$ is based on this formula and the following commutative diagram it provides:

$$\begin{array}{ccc} \mathbb{Z}[P^1_{\mathbb{C}(t)}] & \longrightarrow & B_{n-1}(\mathbb{C}(t)) \otimes \mathbb{C}(t)^* \\ {\scriptstyle \widetilde{\widehat{\mathscr{L}}}_n}\downarrow & & \downarrow{\scriptstyle r_n} \\ S^0(\mathbb{C}P^1) & \xrightarrow{\;d\;} & S^1(\mathbb{C}P^1) \end{array}$$

where

$$\begin{aligned} r_n(\{f(t)\}_{n-1} \otimes g(t)) := {} & \widehat{\mathscr{L}_{n-1}}(f(t))\, d\, i \arg g(t) \\ & - \sum_{k=2}^{n-2} \frac{B_k \cdot 2^k}{k!} \log^{k-1} |f(t)| \cdot \widehat{\mathscr{L}}_{n-k}(f(t))\, d \log |g(t)| \\ & - \frac{B_{n-1} \cdot 2^{n-1}}{(n-1)!} \log |g(t)| \cdot \log^{n-3} |f(t)| \\ & \cdot (\log |f(t)| d \log |1 - f(t)| - \log |1 - f(t))| d \log |f(t)|). \end{aligned}$$

There are three terms in this formula. Each of them is a homomorphism from $B_{n-1}(\mathbb{C}(t)) \otimes \mathbb{C}(t)^*$ to $S^1(\mathbb{C}P^1)$—the first by induction; the second because it is a composition of the homomorphism

$$\begin{array}{ccc} B_{n-1}(\mathbb{C}(t)) \otimes \mathbb{C}(t)^* & \longrightarrow & B_{n-k}(\mathbb{C}(t)) \otimes S^{k-1}\mathbb{C}(t)^* \otimes \mathbb{C}(t)^* \\ {\scriptstyle \delta(k-1)\otimes \mathrm{id}} \searrow & & \nearrow {\scriptstyle \mathrm{id} \otimes \text{projection} \otimes \mathrm{id}} \\ & B_{n-k}(\mathbb{C}(t)) \otimes \underbrace{\mathbb{C}(t)^* \otimes \cdots \otimes \mathbb{C}(t)^*}_{k \text{ times}} & \end{array}$$

($\delta(1) := \delta$ and $\delta(k) := (\delta \otimes \mathrm{id}) \circ \delta(k-1)$) with the obvious homomorphism from $B_{n-k}(\mathbb{C}(t)) \otimes S^{k-1}\mathbb{C}(t)^* \otimes \mathbb{C}(t)$ to $S^1(\mathbb{C}P^1)$; and finally the third is the composition of the homomorphism

$$\begin{array}{ccc} B_{n-1}(\mathbb{C}(t)) \otimes \mathbb{C}(t)^* & \longrightarrow & \bigwedge^2 \mathbb{C}(t)^* \otimes S^{n-2}\mathbb{C}(t)^* \\ {\scriptstyle \delta(n-3)\otimes \mathrm{id}} \searrow & & \nearrow {\scriptstyle \delta \otimes \text{projection}} \\ & B_2(\mathbb{C}(t)) \otimes \mathbb{C}(t)^* \otimes \cdots \otimes \mathbb{C}(t)^* & \end{array}$$

with $r_2 \otimes \sqcap \log |\cdot|$.

For another formula for $d\mathscr{L}_n(z)$ (without Bernoulli numbers on the right-hand side) see **[Z1]**, where Zagier suggests a slightly different definition of the "subgroup of functional equations" for $\mathscr{L}_n(z)$.

THEOREM 1.17. *Suppose that for some* $f_i(t) \in \mathbb{C}(t)$ *we have* $\sum_i a_i \mathscr{L}_n(f_i(t)) = 0$. *Then for any* $z \in \mathbb{C}$

$$\sum_i a_i(\{f_i(z)\} - \{f_i(0)\}) \in \mathscr{R}_n(\mathbb{C}).$$

See Proposition 4.9 for the case $n = 2$. The proof in the general case follows the same idea—to study singularities of $d(\sum a_i \mathscr{L}_n(f_i(t)))$ using formula (1.14). $\square$

Theorem 1.5 permits us to prove that the quotient $\mathscr{A}_n(F)/\mathscr{R}_n(F)$ can be nontrivial. The simplest example is:

EXAMPLE 1.18. $\{1\} \notin \mathscr{R}_{2n+1}(\mathbb{C})$ because $\mathscr{L}_{2n+1}(1) = \zeta_{\mathbb{Q}}(2n+1) \neq 0$. (Compare with Example 1.13 where we proved that $2 \cdot \{1\} \in \mathscr{R}_{2n}(F)$.)

REMARK. Let us denote by $F(X)$ the field of rational functions on a curve X/F. The proof of Theorem 1.15 suggests

DEFINITION 1.19. *$\mathscr{R}'_n(F)$ is generated by elements $\alpha(t_0) - \alpha(t_1)$ where t_0, t_1 run through all F-points of X, X runs through all curves over F, and $\alpha(t)$ runs through all elements of $\mathscr{A}_n(F(X))$.*

The previous definition uses only P^1 instead of all curves over F. However, I believe that the natural map $\mathscr{R}_n(F) \to \mathscr{R}'_n(F)$ is an isomorphism. In fact, this is equivalent to the Rigidity Conjecture 1.7 (see subsection 9 of §1 in [G2]).

6. The main conjecture. Now we are ready to formulate the conjecture about the structure of the Lie algebra $L(F)_\bullet$. As explained in subsection 3 to describe the ideal $I_\bullet$ and extension

$$0 \to H_1(I_\bullet) \to L_\bullet/[I_\bullet, I_\bullet] \to L_\bullet/I_\bullet \to 0, \qquad L_\bullet/I_\bullet = (F^*_{\mathbb{Q}})^\vee$$

it is sufficient to define the following data (see (1.19)):

$$\text{(1.15)} \quad \begin{aligned} &\text{(i)} \quad H^1(I_\bullet) = \bigoplus_{n=2}^{\infty} H^1_{(n)}(I_\bullet), \\ &\text{(ii)} \quad f_2 \colon H^1_{(2)}(I_\bullet) \to \bigwedge^2 F^*_{\mathbb{Q}}, \\ &\text{(iii)} \quad f_n \colon H^1_{(n)}(I_\bullet) \to H^1_{(n-1)}(I_\bullet) \otimes F^*_{\mathbb{Q}}. \end{aligned}$$

CONJECTURE 1.20. *For an arbitrary field F*

(a) *$I(F)_\bullet$ is a free graded pro-Lie algebra;*

(b) *$H^1_{(n)}(I(F)_\bullet) \cong \mathscr{B}_n(F)_{\mathbb{Q}}$, $n \geq 2$, i.e., $I(F)_\bullet$ is generated as a graded Lie algebra by the spaces $\mathscr{B}_n(F)^\vee$ sitting in degree $-n$; and*

(c) *$L_\bullet/I_\bullet \cong (F^*_{\mathbb{Q}})^\vee$ and f_n coincides with*

$$\delta \colon \mathscr{B}_n(F)_{\mathbb{Q}} \to \begin{cases} \mathscr{B}_{n-1}(F)_{\mathbb{Q}} \otimes F^*_{\mathbb{Q}}, & n \geq 3, \\ \bigwedge^2 F^*_{\mathbb{Q}}, & n = 2. \end{cases}$$

2. Corollaries

1. A candidate for the Beilinson-Lichtenbaum complexes. Let us compute $H^*_{(n)}(L(F)_\bullet)$ using the Hochshild-Serre spectral sequence for the ideal $I_\bullet$ and

Conjecture 1.20. We get

$$E_1^{p,q} = C^p(L_\bullet/I_\bullet, H^q_{(n-p)}(I_\bullet)) = \begin{cases} \bigwedge^p F^*_{\mathbb{Q}} \otimes \mathscr{B}_{n-p}(F)_{\mathbb{Q}}; & q = 1, \\ \bigwedge^n F^*_{\mathbb{Q}}, & q = 0, \ n = p, \\ 0, & \text{otherwise.} \end{cases}$$

The action of $L_\bullet/I_\bullet$ on $\bigoplus_{m=2}^{\infty} H_1^{(-m)}(I_\bullet)$ is given by maps f_m^* dual to f_m $(m \geq 3)$. So the differential

$$d_1^{p,q} : \mathscr{B}_{n-p}(F)_{\mathbb{Q}} \otimes \bigwedge^p F^*_{\mathbb{Q}} \to \mathscr{B}_{n-p-1}(F)_{\mathbb{Q}} \otimes \bigwedge^{p+1} F^*_{\mathbb{Q}}$$

is given by the formula $(n - p \geq 3)$

$$\delta : \{x\}_{n-p} \otimes y_1 \wedge \cdots \wedge y_p \mapsto \{x\}_{n-p-1} \otimes x \wedge y_1 \wedge \cdots \wedge y_p$$

$$E_1^{p,q}:$$

$$q = 1 \quad \mathscr{B}_n \xrightarrow{\delta} \mathscr{B}_{n-1} \otimes F^* \xrightarrow{\delta} \mathscr{B}_{n-2} \otimes \bigwedge^2 F^* \xrightarrow{\delta} \cdots \xrightarrow{\delta} \mathscr{B}_2 \otimes \bigwedge^{n-2} F^* \overset{\delta}{\searrow} \bigwedge^n F^*$$

(axes: q vertical, p horizontal)

The only nontrivial higher differential is

$$d_1^{n-2,1} : \mathscr{B}_2(F)_{\mathbb{Q}} \otimes \bigwedge^{n-2} F^*_{\mathbb{Q}} \to \bigwedge^n F^*_{\mathbb{Q}}$$
$$\{x\}_2 \otimes y_1 \wedge \cdots \wedge y_{n-2} \mapsto (1-x) \wedge x \wedge y_1 \wedge \cdots \wedge y_{n-2}.$$

So we get the following complex $\Gamma(F, n)$:

$$\mathscr{B}_n \xrightarrow{\delta} \mathscr{B}_{n-1} \otimes F^* \xrightarrow{\delta} \mathscr{B}_{n-2} \otimes \bigwedge^2 F^* \xrightarrow{\delta} \cdots \xrightarrow{\delta} \mathscr{B}_2 \otimes \bigwedge^{n-2} F^* \xrightarrow{\delta} \bigwedge^n F^*$$

where $\mathscr{B}_n \equiv \mathscr{B}_n(F)$ placed in degree 1 and

$$\delta : \{x\}_p \otimes \bigwedge_{i=1}^{n-p} y_i \to \delta(\{x\}_p) \wedge \bigwedge_{i=1}^{n-p} y_i$$

has degree $+1$. Conjecture 1.20 together with (1.3) imply

CONJECTURE 2.1. $H^i(\Gamma(F, n)_{\mathbb{Q}}) \cong \mathrm{gr}_\gamma^n K_{2n-i}(F)_{\mathbb{Q}}$.

This conjecture gives a description of K-groups in terms of symbols.

EXAMPLE. Let $F = \mathbb{Q}$. We showed in Example 1.18 that $\{1\} \notin \mathscr{R}_{2n+1}(\mathbb{Q})$ for $n \geq 1$. So $\{1\}$ should represent a nontrivial element in $\mathrm{gr}_\gamma^{2n+1} K_{4n+1}(\mathbb{Q})$. Note that $\dim K_m(\mathbb{Q}) = 1$ for $m = 4n+1$, 0 otherwise.

Complexes $\Gamma(F, n)_{\mathbb{Q}}$ should satisfy Beilinson-Lichtenbaum axioms.

In fact, Conjecture 2.1 is equivalent to Conjecture 1.20 if we assume (1.3). More precisely, let us suppose that there exist homomorphisms $\varphi_n: \mathscr{B}_n(F)_{\mathbb{Q}} \to L(F)^{\vee}_{-n}$ such that the following diagrams are commutative:

$$\begin{array}{ccc} \mathscr{B}_2(F)_{\mathbb{Q}} & \xrightarrow{\delta} & \bigwedge^2 F^*_{\mathbb{Q}} \\ \varphi_2 \downarrow & & \downarrow \bigwedge^2 \varphi_1 \\ L(F)^{\vee}_{-2} & \xrightarrow{\partial} & \bigwedge^2 L(F)^{\vee}_{-1} \end{array} \tag{2.1a}$$

$$\begin{array}{ccc} \mathscr{B}_n(F)_{\mathbb{Q}} & \xrightarrow{\delta} & \mathscr{B}_{n-1}(F)_{\mathbb{Q}} \otimes F^*_{\mathbb{Q}} \\ \varphi_n \downarrow & & \downarrow \varphi_{n-1} \otimes \varphi_1 \\ L(F)^{\vee}_{-n} & \xrightarrow{\partial_{(1)}} & L(F)^{\vee}_{-(n-1)} \otimes L(F)^{\vee}_{-1}. \end{array} \tag{2.1b}$$

Here $\partial_{(1)}$ is the $L^{\vee}_{-(n-1)} \otimes L^{\vee}_{-1}$-component of δ. Then we get a homomorphism of complexes

$$\Phi_n: \Gamma(F, n)_{\mathbb{Q}} \to \bigwedge^{\bullet}_{(n)}(L(F)^{\vee}_{\bullet}). \tag{2.2}$$

THEOREM 2.2. *Suppose that there exists a graded Lie algebra* $L_{\bullet} = \bigoplus_{n=1}^{\infty} L_{-n}$ *and homomorphisms* $\varphi_n: \mathscr{B}_n \to L^{\vee}_{-n}$ *such that diagrams* (2.1a), (2.1b) *are commutative and* Φ_n *is a quasi-isomorphism for* $n \geq 1$. *Then*

(a) $I_{\bullet} := \bigoplus_{n=2}^{\infty} L_{-n}$ *is a free graded Lie algebra,*

(b) $\varphi_n: \mathscr{B}_n \to H^1_{(n)}(I_{\bullet})$ *is an isomorphism for any* $n \geq 2$,

(c) *maps* f_n *describing the quotient* $L_{\bullet}/[I_{\bullet}, I_{\bullet}]$ (*see* (1.15)) *coincide with*

$$\delta: \mathscr{B}_n \to \begin{cases} \mathscr{B}_{n-1} \otimes F^*_{\mathbb{Q}}, & n \geq 3, \\ \bigwedge^2 F^*_{\mathbb{Q}}, & n = 2. \end{cases}$$

For the proof of the theorem, see proof of Proposition 1.26 in **[G2]**. □

Our next purpose will be to show that Conjecture 2.1 in the case when F is a number field implies Zagier's conjecture about the values of Dedekind zeta functions $\zeta_F(n)$. But first of all we need to recall the Borel theorems.

2. The Borel theorems. Set $\mathbb{R}(n) = (2\pi i)^n \mathbb{R} \subset \mathbb{C}$ and $X_F := \mathbb{Z}^{\mathrm{Hom}(F, \mathbb{C})}$. Let us define the Borel regulator $r_m: K_{2m-1}(F) \to X_F \otimes \mathbb{R}(m-1)$. The Hurewicz map gives a canonical homomorphism

$$\begin{aligned} K_{2m-1}(F) := \pi_{2m-1}(\mathrm{BGL}(F)^+) &\to H_{2m-1}(\mathrm{BGL}(F)^+, \mathbb{Z}) \\ &= H_{2m-1}(\mathrm{GL}(F), \mathbb{Z}). \end{aligned} \tag{2.3}$$

For every embedding $\sigma: F \hookrightarrow \mathbb{C}$ we have a homomorphism

$$H_{2m-1}(\mathrm{GL}(F), \mathbb{Z}) \to H_{2m-1}(\mathrm{GL}(\mathbb{C}), \mathbb{Z}). \tag{2.4}$$

There is a canonical pairing

$$H^{2m-1}(\mathrm{GL}(\mathbb{C}), \mathbb{R}(m-1)) \times H_{2m-1}(\mathrm{GL}(\mathbb{C}), \mathbb{Z}) \xrightarrow{\langle , \rangle} \mathbb{R}(m-1). \tag{2.5}$$

Let us define a canonical element

$$b_{2m-1} \in H^{2m-1}_{\mathrm{cts}}(\mathrm{GL}(\mathbb{C}), \mathbb{R}(m-1)) \subset H^{2m-1}(\mathrm{GL}(\mathbb{C}), \mathbb{R}(m-1)).$$

Recall that (cf. **[Bo1]**) $H^*_{\mathrm{cts}}(\mathrm{GL}(\mathbb{C}), \mathbb{R}) \cong H^*_{\mathrm{top}}(U, \mathbb{R})$ where $H^*_{\mathrm{top}}(U, \mathbb{R})$ is the cohomology of the infinite unitary group, considered as a topological space. Further,

$$H^*_{\mathrm{top}}(U, \mathbb{Z}) = H^*(S^1 \times S^3 \times S^5 \times \cdots, \mathbb{Z}) = \bigwedge^*(u_1, u_3, \dots)$$

where u_i denotes the class of the sphere S^i.

Combining the above isomorphisms we get an isomorphism

$$\varphi\colon H^*_{\mathrm{cts}}(\mathrm{GL}(\mathbb{C}), \mathbb{R}) \xrightarrow{\sim} \bigwedge^*(u_1, u_3, \dots) \otimes \mathbb{R}. \tag{2.6}$$

Set $b'_{m-1} := 2\pi \cdot \varphi^{-1}(u_{2m-1})$ and

$$b_{2m-1} := (2\pi i)^{m-1} \cdot b'_{2m-1} \in H^*_{\mathrm{cts}}(\mathrm{GL}(\mathbb{C}), \mathbb{R}(m-1)).$$

So combining this with (2.3)–(2.5) we get

$$K_{2m-1}(F) \to \bigoplus_{\mathrm{Hom}(F, \mathbb{C})} K_{2m-1}(\mathbb{C}) \to X_F \otimes \mathbb{R}(m-1).$$

It is known that if $\lambda \in H^d_{\mathrm{cont}}(\mathrm{GL}(\mathbb{C}), \mathbb{R})$ and c^* denotes the involution defined by complex conjugation c, then in (2.6) $c^*\varphi(\lambda) = (-1)^d\varphi(c^*\lambda)$, where c acts also on $S^{2m-1} \subset \mathbb{C}^m$. Note that $c^*u_{2m-1} = (-1)^m u_{2m-1}$. So we see that

$$r_m\colon K_{2m-1}(F) \to [X_F \otimes \mathbb{R}(m-1)]^+ = \mathbb{R}^{d_m}$$

where on the right-hand side stands the subspace of invariants of the action of c and

$$d_m = \begin{cases} r_1 + r_2 & \text{if } m \text{ is odd,} \\ r_2 & \text{if } m \text{ is even} \end{cases}$$

is its dimension. Here r_1 (resp. r_2) denotes the number of real (resp. complex) places; so $[F : Q] = r_1 + 2r_2$.

In fact, we have constructed a homomorphism

$$r^{(n)}_m\colon \operatorname{Prim} H_{2m-1}(\mathrm{GL}_n(F), \mathbb{Z}) \to [X_F \otimes \mathbb{R}(m-1)]^+.$$

For any lattice Λ of $(X_F \otimes \mathbb{R}(m-1))^+$ define its (co)volume $\operatorname{vol}\Lambda$ by

$$\det(\Lambda) = \operatorname{vol}(\Lambda) \cdot \det[X_F(m-1)^+].$$

THEOREM 2.3 (Borel **[Bo1]**, **[Bo2]**). *For every $m \geq 2$ and for sufficiently large n*

(a) $\operatorname{Im} r^{(n)}_m$ *is a lattice in* $(X_F \otimes \mathbb{R}(m-1))^+$ *and*

(b) $R_m := \mathrm{vol}(\mathrm{Im}\, r_m^{(n)}) \sim \mathbb{Q}^* \cdot \lim_{s\to 1-m}(s-1+m)^{-d_m}\zeta_F(s)$.

Here $a \sim \mathbb{Q}^* b$ means that $a = \kappa b$ for some $\kappa \in \mathbb{Q}^*$.

REMARK 2.4. The functional equation for $\zeta_F(s)$ shows that

$$\zeta_F(m) \sim \mathbb{Q}^* \cdot \pi^{(r_1+2r_2-d_m)\cdot m} \cdot |d_F|^{-1/2} \cdot R_m$$

where d_F is the discriminant of F.

3. Zagier's conjecture. According to Conjecture 2.1 we have an isomorphism

$$H^1(\Gamma(\mathbb{C}, n)_{\mathbb{Q}}) \equiv \mathrm{Ker}(\mathscr{B}_n(\mathbb{C})_{\mathbb{Q}} \xrightarrow{\delta} \mathscr{B}_{n-1}(\mathbb{C})_{\mathbb{Q}} \otimes \mathbb{C}^*) \cong \mathrm{gr}_\gamma^n K_{2n-1}(\mathbb{C})_{\mathbb{Q}}.$$

Recall that there is a homomorphism $\mathscr{L}_n \colon \mathscr{B}_n(\mathbb{C}) \to R$. We expect that the restriction of this homomorphism to the subgroup $H^1(\Gamma(\mathbb{C}, (n)_{\mathbb{Q}}) \subset \mathscr{B}_n(\mathbb{C})_{\mathbb{Q}}$ coincides with the Borel regulator (the reasons can be found in §1 of **[G2]**). So applying the Borel theorem we come to the following conjecture.

CONJECTURE 2.5. *Let F be a number field and σ_j the set of all possible embeddings $F \hookrightarrow \mathbb{C}$ $(1 \le j \le r_1 + 2r_2)$ numbered so that $\sigma_{r_1+k} = \overline{\sigma_{r_1+r_2+k}}$. Then there exist elements*

$$y_1, \dots, y_{d_n} \in \mathrm{Ker}(\mathscr{B}_n(F)_{\mathbb{Q}} \xrightarrow{\delta} \mathscr{B}_{n-1}(F)_{\mathbb{Q}} \otimes F_{\mathbb{Q}}^*)$$

such that

$$\zeta_F(n) = \pi^{(r_1+2r_2-d_n)\cdot n}|d_F|^{-1/2} \det|\mathscr{L}_n(\sigma_j(y_i))| \qquad (1 \le i, j \le d_n).$$

This conjecture was stated by Don Zagier, who proved it for $n = 2$ **[Z2]** and using a computer gave an impressive list of numerical examples (see **[Z1]**). The case $n = 2$ follows also from the Borel theorem and the results of Bloch **[Bl1]** and Suslin **[S1]**. A complete proof for the case $n = 3$ will be given in §3 (see also **[G1]** and **[G2]**).

4. A topological consequence of Conjecture 1.9. We shall show that in Beilinson's World (*a world where his conjectures are theorems*) Conjecture 1.9 implies that the commutant of the maximal Tate quotient of the pronilpotent completion of the classical fundamental group of the generic point of an arbitrary complex variety over $\mathbb{C}$ should be a free graded pro-Lie algebra.

Recall that Beilinson conjectured **[B1]** that for an arbitrary scheme X there exists a mixed Tate category $\mathscr{M}_T(X)$ of mixed motivic Tate sheaves over X. In the special case $X = \mathrm{Spec}\, F$, F is a field, $\mathscr{M}_T(\mathrm{Spec}\, F)$ is just the category $\mathscr{M}_T(F)$ discussed in subsections 1–2 of §1. Let us denote by $L(X)_\bullet$ the corresponding mixed Tate Lie algebra. Any morphism of schemes $f\colon X \to Y$ defines a Tate functor $f^*\colon \mathscr{M}_T(Y) \to \mathscr{M}_T(X)$ ("inverse image" of mixed Tate sheaves) such that $\omega_{\mathscr{M}_T(X)} f^* = \omega_{\mathscr{M}_T(Y)}$ ($\omega_{\mathscr{M}}$ is the canonical fiber functor for a mixed Tate category $\mathscr{M}$). So we have a morphism $f_*\colon L(X)_\bullet \to L(Y)_\bullet$ of the corresponding mixed Tate Lie algebras. In particular, if X is a scheme over a field F, we have the map

$p_*: L(X)_\bullet \to L(\operatorname{Spec} F)_\bullet$ that should be surjective because p^* is fully faithful. Put $L(X)^g_\bullet := \operatorname{Ker} p_*$ (the "geometric part of $L(X)_\bullet$"). We get the following exact sequence

$$0 \to L(X)^g_\bullet \to L(X)_\bullet \xrightarrow{p_*} L(\operatorname{Spec} F)_\bullet \to 0.$$

Note that

$$[L(X)^g_\bullet, L(X)^g_\bullet] \subset L(X)^g_{\leq -2}.$$

(It was proved in **[B2]** (see Lemma 1.2.1) that $L(X)^g_\bullet$ is generated by $L(X)^g_{-1}$; so $[L(X)^g_\bullet, L(X)^g_\bullet] = L(X)^g_{\leq -2}$, but we shall not use this fact.)

Let $\eta = \operatorname{Spec} F(X)$ be the generic point of X. Then according to Conjecture 1.9 the (graded) Lie algebra $L(\eta)_{\leq -2}$ is free. Therefore, its subalgebra $[L(\eta)^g_\bullet, L(\eta)^g_\bullet]$ is also free.

Now let X be a smooth algebraic variety over $\mathbb{C}$. I need to explain what is the *maximal Tate quotient of the pronilpotent completion of* $\pi_1(\operatorname{Spec} \mathbb{C}(X))$. In **[HZ]** Hain and Zucker defined the category $\mathscr{H}^{\mathrm{un}}_X$ of good unipotent variations of mixed $\mathbb{R}$-Hodge structures over X ("good" means some growth conditions at infinity).

Fix any $x \in X$. Let $V \in Ob\mathscr{H}^{\mathrm{un}}_X$ and V_x be the fiber of the local system underlying V at point x. Then the monodromy representation $\rho: \pi_1(X, x) \to \operatorname{Aut}(V_x)$ is unipotent and hence defines an algebra homomorphism $\bar{\rho}: \mathbb{C}\pi_1(X, x)^\wedge \to \operatorname{Aut}(V_x)$, where $\mathbb{C}\pi_1(X, x)^\wedge := \varprojlim \mathbb{C}[\pi_1(X, x)]/J^r$ (J is the kernel of the usual augmentation homomorphism). It is well known that $\mathbb{C}\pi_1(X, x)^\wedge$ is a Hopf algebra in the category $\mathscr{H}$ of mixed $\mathbb{R}$-Hodge structures and $\bar{\rho}$ is a mixed Hodge-theoretic representation (i.e., representation in the category $\mathscr{H}$). Hain and Zucker proved the following theorem.

THEOREM 2.6. *The monodromy representation functor* $V \in \mathscr{H}^{\mathrm{un}}_X \mapsto V_x$ *defines an equivalence of categories*

$$\mathscr{H}^{\mathrm{un}}_X \to \left\{ \begin{array}{l} \textit{category of mixed Hodge-theoretic} \\ \textit{representations of } \mathbb{C}\pi_1(X, x)^\wedge. \end{array} \right.$$

The vector space underlying a Hodge structure $H \in \mathscr{H}$ is a fiber functor on the category $\mathscr{H}$. Composition of the functor $s_x: \mathscr{H}^{\mathrm{un}}_X \to \mathscr{H}$, $s_x: V \mapsto V_x$ with this fiber functor gives a fiber functor on $\mathscr{H}^{\mathrm{un}}_X$. Let us denote by $L(\mathscr{H})$ and $L(\mathscr{H}^{\mathrm{un}}_X, x)$ the corresponding fundamental Lie algebras. We get an embedding $s_x: L(\mathscr{H}) \to L(\mathscr{H}^{\mathrm{un}}_X)$. There is a canonical functor $c: \mathscr{H} \to \mathscr{H}^{\mathrm{un}}_X$, where $c(H)$ is a constant variation of the mixed Hodge structure H over X. So we get an epimorphism $c: L(\mathscr{H}^{\mathrm{un}}_X, x) \to L(\mathscr{H})$. It is clear that $c \circ s_x = \mathrm{id}$. Set $L(\mathscr{H}^{\mathrm{un}}_X, x)^g := \operatorname{Ker} c$. We get the following split exact sequence

$$0 \to L(\mathscr{H}^{\mathrm{un}}_X, x)^g \to L(\mathscr{H}^{\mathrm{un}}_X, x) \underset{c}{\overset{s_x}{\leftrightarrows}} L(\mathscr{H}) \to 0.$$

Note that $s_x(L(\mathscr{H}))$ acts on the ideal $L(\mathscr{H}_X^{\mathrm{un}}, x)^g$, and hence $L(\mathscr{H}_X^{\mathrm{un}}, x)^g$ is equipped with canonical mixed Hodge structure. Further, an $L(\mathscr{H}_X^{\mathrm{un}}, x)$-module is just a mixed Hodge-theoretic representation of $L(\mathscr{H}_X^{\mathrm{un}}, x)^g$.

We have $\mathbb{C}\pi_1(X, x)^\wedge = \mathbb{C} \oplus \widehat{J}$. The set of primitive elements

$$\mathfrak{G}_x := \{v \in \widehat{J} : \Delta(v) = v \hat{\otimes} 1 + 1 \hat{\otimes} v\}$$

is a Lie algebra (Δ is the coproduct). The forgetting functor $\mathscr{H}_X^{\mathrm{un}} \to$ {local systems on X} provides a homomorphism of Lie algebras $f_x : \mathfrak{G}_x \to L(\mathscr{H}_X^{\mathrm{un}}, x)$ such that $c \circ f_x = 0$. So $f_x : \mathfrak{G}_x \to L(\mathscr{H}_X^{\mathrm{un}}, x)^g$. Mixed Hodge structures $\mathfrak{G}_x$ form a good variation of mixed Hodge structures over X. So f_x is a morphism of mixed Hodge structures. Now it follows from Theorem 2.6 that $f_x : \mathfrak{G}_x \xrightarrow{\sim} L(\mathscr{H}_X^{\mathrm{un}}, x)^g$ is an isomorphism.

Let $\mathscr{H}_X^T \subset \mathscr{H}_X^{\mathrm{un}}$ be a subcategory of variations of mixed Hodge-Tate structures (i.e., $\mathrm{gr}_{2n-1}^W V_x = 0$, $\mathrm{gr}_{2n}^W V_x$ is a Hodge structure of type (n, n)). Then $L(\mathscr{H}_X^T, x)^g$ is a maximal Tate quotient of $L(\mathscr{H}_X^{\mathrm{un}}, x)^g$. If $\mathfrak{G}_x^T(X)$ is a maximal Tate quotient of $\mathfrak{G}_x$, $\mathfrak{G}_x^T(X) \xrightarrow{\sim} L(\mathscr{H}_X^T, x)^g$ is an isomorphism. There is another fiber functor on the category $\mathscr{H}_X^T$ that does not involve the choice of $x \in X$: $H \in \mathscr{H}_X^T \mapsto \bigoplus_n \mathrm{gr}_{2n}^W H$. Let us denote the corresponding geometric Lie algebra by $L(\mathscr{H}_X^T)^g$. Of course, $L(\mathscr{H}_X^T)^g \cong \mathfrak{G}_x^T(X)$. Set

$$L(\mathscr{H}_\eta^T)^g := \varprojlim_{U \subset X} L(\mathscr{H}_U^T)^g.$$

This is the *definition* of the maximal Tate quotient of the pronilpotent completion of the fundamental group of a generic point of a complex algebraic variety.

CONJECTURE 2.7. *The commutant of the Lie algebra* $L(\mathscr{H}_\eta^T)^g$ *is free.*

The Hodge-realization functor $\mathscr{M}_T(X) \to \mathscr{H}_X^T$ induces a morphism $L(\mathscr{H}_X^T) \to L(\mathscr{M}_T(X))$ that should be an isomorphism. (This follows from Beilinson's definition of mixed Hodge structure on $\mathbb{C}\pi_1(X, x)^\wedge$ and standard conjectures including the Hodge one; see **[B2]** and **[BD]**). Therefore, Conjecture 2.7 is a corollary of Conjecture 1.9 in Beilinson's World.

It is interesting to compare Conjecture 2.7 with the following one stated by F. A. Bogomolov.

CONJECTURE 2.8. *Let* $\mathrm{Gal}\, K$ *be the maximal pro-p-quotient of the Galois group of the field* K *containing a nontrivial algebraically closed subfield. Then the commutant* $[\mathrm{Gal}\, K, \mathrm{Gal}\, K]$ *is free as a pro-p-group.*

It is also reminiscent of the following Shafarevich conjecture.

CONJECTURE 2.9. $[\mathrm{Gal}\, \overline{\mathbb{Q}}/\mathbb{Q}, \mathrm{Gal}\, \overline{\mathbb{Q}}/\mathbb{Q}]$ *is free as a profinite group.*

3. A proof of Zagier's conjecture about $\zeta_F(3)$

1. The Grassmanian complex ([S1], see also [BMS]). We shall say that an m-tuple of vectors in an n-dimensional vector space V^n is in a generic position if any $k \leq n$ vectors are linearly independent. *Configurations* of m vectors in V^n are n-tuples of vectors considered modulo $\mathrm{GL}(V^n)$-equivalence. Let us denote by $\widetilde{C}_m(n)$ the free abelian group generated by m-tuples of vectors in V^n in generic position. Let $C_m(n) := \widetilde{C}_m(n)_{\mathrm{GL}(V^n)}$ be the group of co-invariants of the natural action of $\mathrm{GL}(V^n)$ on $\widetilde{C}_m(n)$. Then $C_m(n)$ is a free abelian group generated by configurations of m vectors in generic position in V^n. There is a differential

$$d\colon \widetilde{C}_m(n) \to \widetilde{C}_{m-1}(n); \quad d\colon (l_1, \dots, l_m) \mapsto \sum_{i=1}^{m} (-1)^{i-1}(l_1, \dots, \hat{l}_i, \dots, l_m).$$

We get a complex $(\widetilde{C}_*(n), d)$ where $\widetilde{C}_m(n)$ placed in degree $m-1$.

LEMMA 3.1. *$H_i(\widetilde{C}_*(n)) = 0$ for $i \geq 1$ and $\mathbb{Z}$ for $i = 0$ if F is an infinite field.*

PROOF. If $d(\sum n_j(l_1^{(j)}, \dots, l_m^{(j)})) = 0$, choose a vector v in a generic position with respect to all $l_k^{(j)}$. Then $d(\sum n_j(v, l_1^{(j)}, \dots, l_m^{(j)})) = \sum n_j(l_1^{(j)}, \dots, l_m^{(j)})$. $\square$

So $\widetilde{C}_*(n)$ is a resolution of $\mathbb{Z}$, and therefore we have a map

$$H_i(\mathrm{GL}_n(F)) \to H_i(C_*(n)). \tag{3.1}$$

2. Our strategy. We shall work modulo 6-torsion. In the next section we shall construct a homomorphism of complexes

$$\begin{array}{ccccccc}
C_7(3) & \xrightarrow{d} & C_6(3) & \xrightarrow{d} & C_5(3) & \xrightarrow{d} & C_4(3) \\
 & & \downarrow f_6(3) & & \downarrow f_5(3) & & \downarrow f_4(3) \\
0 & \longrightarrow & \mathscr{B}_3(F) & \xrightarrow{\delta} & \mathscr{B}_2(F) \otimes F^* & \xrightarrow{\delta} & \bigwedge^3 F^*
\end{array} \tag{3.2}$$

and hence get a map

$$c_i(3)\colon H_i(\mathrm{GL}_3(F)) \to H^{6-i}(\Gamma(F, 3)), \qquad i = 3, 4, 5.$$

Then we shall construct a map $c_i(N)\colon H_i(\mathrm{GL}_N(F)) \to H^{6-i}(\Gamma(F, 3))$ such that the following diagram is commutative:

$$\begin{array}{ccc}
H_i(\mathrm{GL}_3(F)) & \xrightarrow{c_i(3)} & H^{6-i}(\Gamma(F, 3)) \\
 & \searrow \qquad \nearrow c_i(N) & \\
 & H_i(\mathrm{GL}_N(F)) &
\end{array}$$

and $\operatorname{Im} c_i(N) = \operatorname{Im} c_i(3)$.

Recall that $H_n(\mathrm{GL}_n(F)) = H_n(\mathrm{GL}(F))$ (see **[S1]**), so

$$K_n(F)_{\mathbb{Q}} = \mathrm{Prim}\, H_n(\mathrm{GL}(F), \mathbb{Q}) = \mathrm{Prim}\, H_n(\mathrm{GL}_n(F), \mathbb{Q}).$$

Put

$$\begin{aligned} K_n^{(j)}(F)_{\mathbb{Q}} &:= \mathrm{Im}(H_n(\mathrm{GL}_{n-j}(F), \mathbb{Q}) \to H_n(\mathrm{GL}_n(F), \mathbb{Q})) \\ &\quad \cap \mathrm{Prim}\, H_n(\mathrm{GL}_n(F), \mathbb{Q}), \\ K_n^{[j]}(F)_{\mathbb{Q}} &:= K_n^{(j)}(F)_{\mathbb{Q}} / K_n^{(j+1)}(F)_{\mathbb{Q}}. \end{aligned}$$

CONJECTURE 3.2 (A. A. Suslin, unpublished). $K_n^{[j]}(F)_{\mathbb{Q}} \cong \mathrm{gr}_\gamma^{n-j} K_n(F)_{\mathbb{Q}}$.

(Yang **[Y1]** proved this conjecture for number fields with only one exception: $F = \mathbb{Q}$. Another proof for all number fields in the case $n = 5$ follows from results of this paper.) We show that $c_i(3)$ vanishes on the image of $H_i(\mathrm{GL}_2(F), \mathbb{Q})$; hence, we get canonical homomorphisms

$$C_i^{[i-3]}: K_i^{[i-3]}(F)_{\mathbb{Q}} \to H^{6-i}(\Gamma(F, 3) \otimes \mathbb{Q}) \qquad (i = 3, 4, 5).$$

Suslin proved that $K_n^{[0]}(F)_{\mathbb{Q}} \cong K_n^M(F)_{\mathbb{Q}}$. So $C_3^{[0]}$ is an isomorphism. $C_4^{[1]}$ and $C_5^{[2]}$ also should be isomorphisms. In any case $C_5^{[2]}: K_5^{[2]}(F) \to H^1(\Gamma(F, 3) \otimes \mathbb{Q})$. We shall construct a homomorphism $c_5: K_5(F) \to H^1(\Gamma(F, 3) \otimes \mathbb{Q})$ and show that the composition

$$K_5(\mathbb{C}) \xrightarrow{c_5} H^1(\Gamma(\mathbb{C}, 3) \otimes \mathbb{Q}) \xrightarrow{\widetilde{\mathscr{L}}_3} R$$

coincides with the Borel regulator **[Bo2]**. This implies immediately Zagier's conjecture about $\zeta_F(3)$.

3. Construction of homomorphism of complexes 3.2. Choose a volume form $\omega \in \bigwedge^3 (V^3)^*$. Set $\Delta(l_i, l_j, l_k) := \langle \omega, l_i \wedge l_j \wedge l_k \rangle \in F^*$. Put

$$(3.3) \qquad f_4(3): (l_1, \dots, l_4) \mapsto \mathrm{Alt}\, \Delta(l_1, l_2, l_3) \wedge \Delta(l_1, l_2, l_4) \wedge \Delta(l_1, l_3, l_4).$$

Here

$$\mathrm{Alt}\, f(l_1, \dots, l_n) := \sum_{\sigma \in S_n} (-1)^{|\sigma|} f(l_{\sigma(1)}, \dots, l_{\sigma(n)}).$$

LEMMA 3.3. *$f_4(3)$ does not depend on the choice of ω.*

PROOF. Let $\omega' = \lambda\omega$, $\lambda \in F^*$. Then the difference between the right-hand sides of (3.3) computed using ω' and ω is $\mathrm{Alt}(\lambda \wedge \Delta(l_1, l_2, l_4) \wedge \Delta(l_1, l_3, l_4))$. But this is 0 because we alternate an expression that is symmetric with respect to permutation of 1 and 4. □

For a nonzero vector $l \in V^3$ let us denote by $\bar{l}$ the corresponding point in $P(V^3) = P^2$. Let us denote by $(\bar{l}_1 \,|\, \bar{l}_2, \dots, \bar{l}_4)$ the configuration of 4 points on P^1 obtained by projection of points $\bar{l}_2, \dots, \bar{l}_5$ with the center at point

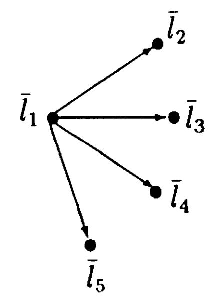


FIGURE 3.1

$\bar{l}_1$; see Figure 3.1. (All lines passing through $\bar{l}_1$ form a projective line; any point $m \neq \bar{l}_1$ defines a point on this line.)

Now let $(m_1, \dots, m_4) \in C_4(2)$. Let us define the cross-ratio $r(\overline{m}_1, \dots, \overline{m}_4)$ as

$$r(\overline{m}_1, \dots, \overline{m}_4) := \frac{\Delta(m_1, m_3)\Delta(m_2, m_4)}{\Delta(m_1, m_4)\Delta(m_2, m_3)}. \tag{3.4}$$

It is clear that the right-hand side of (3.4) does not depend on the length of m_i. We have

(3.5)

$$r(\overline{m}_1, \overline{m}_2, \overline{m}_3, \overline{m}_4) = r(\overline{m}_2, \overline{m}_1, \overline{m}_3, \overline{m}_4)^{-1} = r(\overline{m}_1, \overline{m}_2, \overline{m}_4, \overline{m}_3)^{-1} = 1 - r(\overline{m}_1, \overline{m}_3, \overline{m}_2, \overline{m}_4).$$

The last equality is proved using the identity

$$\Delta(m_1, m_4)\Delta(m_2, m_3) - \Delta(m_1, m_2)\Delta(m_3, m_4) = \Delta(m_1, m_3)\Delta(m_2, m_4).$$

Set

$$f_5(3)(l_1, \dots, l_5) := \tfrac{1}{2}\operatorname{Alt}(\{r(\bar{l}_1 \mid \bar{l}_2, \dots, \bar{l}_5)\}_2 \otimes \Delta(l_1, l_2, l_3)). \tag{3.6}$$

Here $\{x\}_2$ means the image of $\{x\}$ in $\mathscr{B}_2(F)$.

PROPOSITION 3.4. *$f_5(3)$ does not depend on ω.*

PROOF. The difference between the right-hand sides of (3.6) computed using $\lambda \cdot \omega$ and ω is proportional to

$$\sum_{i=1}^{5}(-1)^i\{r(\bar{l}_i \mid \bar{l}_1, \dots, \hat{\bar{l}}_i, \dots, \bar{l}_5)\}_2 \otimes \lambda$$

because $\{r(m_1, \dots, m_4)\}_2 \in \mathscr{B}_2(F)_{\mathbb{Q}}$ is skew-symmetric with respect to permutation of points m_i; see (3.5) and Example 1 in subsection 4 of §1. So we need to prove the following.

LEMMA 3.5. *Let $x_1, \dots, x_5$ be five points on P^2 in generic position. Then*

$$\sum_{i=1}^{5}(-1)^i\{r(x_i \mid x_1, \dots, \hat{x}_i, \dots, x_5)\} \in \mathscr{R}_2(F).$$

This lemma follows from

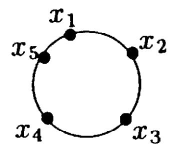


FIGURE 3.2

LEMMA 3.6. *Let* $m_1, \ldots, m_5$ *be five different points on* P^1. *Then*

$$(3.7)\qquad R_2(m_1, \ldots, m_5) := \sum_{i=1}^{5} (-1)^i \{r(m_1, \ldots, \hat{m}_i, \ldots, m_5)\} \in \mathcal{R}_2(F).$$

Indeed, let us consider the unique conic (a curve of order two) passing through the points $x_1, \ldots, x_5$ as a projective line. It remains to apply Lemma 3.6 to these points on the projective line in Figure 3.2.

PROOF OF LEMMA 3.6. Consider the following homomorphism of complexes

$$(3.8)\qquad \begin{array}{ccccc} C_5(2) & \xrightarrow{d} & C_4(2) & \xrightarrow{d} & C_3(2) \\ & & \downarrow{\scriptstyle f_4(2)} & & \downarrow{\scriptstyle f_3(2)} \\ & & \mathbb{Z}[P^1_F] & \xrightarrow{\delta_2} & \bigwedge^2 F^* \end{array}$$

$$f_3(2) : (l_1, l_2, l_3) \mapsto \Delta(l_1, l_2) \wedge \Delta(l_1, l_3) - \Delta(l_2, l_1) \wedge \Delta(l_2, l_3) + \Delta(l_3, l_1) \wedge \Delta(l_3, l_2),$$
$$f_4(2) : (l_1, \ldots, l_4) \mapsto \{r(\bar{l}_1, \ldots, \bar{l}_4)\}.$$

Direct calculation using (3.4)–(3.5) shows that (3.8) is commutative. So

$$\delta_2 \left(\sum_{i=1}^{5} (-1)^i \right) \{r(\overline{m}_1, \ldots, \widehat{\overline{m}}_i, \ldots, \overline{m}_5)\} \equiv \delta_2 \circ f_4(2) \circ d = f_3(2) \circ d^2 = 0.$$

Now it is easy to complete the proof of Lemma 3.6 using specialization. □

PROPOSITION 3.7. $f_4(3) \circ d = d \circ f_5(3)$.

PROOF. Direct calculation using (3.4). □

Now put

$$(3.9)\qquad f_6(3) : (l_1, \ldots, l_6) \mapsto \operatorname{Alt} \left\{ \frac{\Delta(l_1, l_2, l_4)\Delta(l_2, l_3, l_5)\Delta(l_3, l_1, l_6)}{\Delta(l_1, l_2, l_5)\Delta(l_2, l_3, l_6)\Delta(l_3, l_1, l_4)} \right\}.$$

4. A geometric definition of the generalized cross-ratio (3.9). Let $(a_1, a_2, a_3, b_1, b_2, b_3)$ be a configuration of 6 distinct points in P^2 such that a_1, a_2, a_3 do not lie on a line and $b_i \in \overline{a_i a_{i+1}}$ (see Figure 3.3 on p. 64). Let $P^2 = P(V_3)$. Choose vectors in V_3 such that they are projected to points

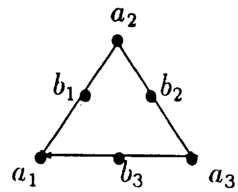


FIGURE 3.3

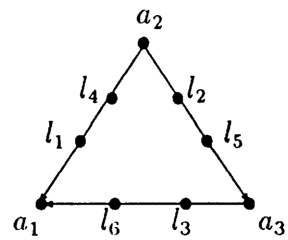


FIGURE 3.4

a_i, b_i. By an abuse of notation we shall denote them by the same letters. Choose $f_i \in V_3^*$ such that $f_i(a_i) = f_i(a_{i+1}) = 0$. Put

$$(3.10) \qquad r_3'(a_1, a_2, a_3, b_1, b_2, b_3) = \frac{f_1(b_2) \cdot f_2(b_3) \cdot f_3(b_1)}{f_1(b_3) \cdot f_2(b_1) \cdot f_3(b_2)}.$$

The right-hand side of (3.10) does not depend on the choice of vectors f_i, b_j.

LEMMA 3.8. $r(b_1 | a_2, a_3, b_2, b_3) = r_3'(a_1, a_2, a_3, b_1, b_2, b_3)$.

PROOF. Put

$$f_1(v) := \Delta(b_1, a_2, v); \; f_2(v) := \Delta(b_2, a_3, v); \; f_3(v) := \Delta(b_3, a_3, v).$$

Then the right-hand side of (3.10) is equal to

$$\frac{\Delta(b_1, a_2, b_2) \cdot \Delta(b_2, a_3, b_3) \cdot \Delta(b_3, a_3, b_1)}{\Delta(b_1, a_2, b_3) \cdot \Delta(b_2, a_3, b_1) \cdot \Delta(b_3, a_3, b_2)} = \frac{\Delta(b_1, a_2, b_2) \cdot \Delta(b_1, a_3, b_3)}{\Delta(b_1, a_2, b_3) \cdot \Delta(b_1, a_3, b_2)}$$
$$= r(b_1 | a_2, a_3, b_2, b_3). \quad \square$$

Now let $(l_1, \ldots, l_6)$ be a configuration of six distinct points in P^2 in generic position. Put $a_i := \overline{l_i l_{i+3}} \cap \overline{l_{i-1} l_{i+2}}$ $(1 \le i \le 3$, indices modulo 6; see Figure 3.4).

Then $l_i \in \overline{a_i a_{i+1}}$; so $(a_1, a_2, a_3, l_1, l_2, l_3)$ is a configuration of the type considered above. Let us define the generalized cross-ratio $r_3 \colon C_6(3) \to \mathbb{Z}[P_F^1 \backslash \{0, \infty\}]$ as follows:

$$(3.11) \qquad r_3(l_1, \ldots, l_6) := \mathrm{Alt}\{r_3'(a_1, a_2, a_3, l_1, l_2, l_3)\} \in \mathbb{Z}[P_F^1 \backslash \{0, \infty\}].$$

More precisely, the alternation here means the following. Let $s \in S_6$ be a permutation and

$$a_i^{(s)} := \overline{l_{s(i)} l_{s(i+3)}} \cap \overline{l_{s(i-1)} l_{s(i+2)}} \qquad (1 \le i \le 3).$$

Then

$$r_3(l_1, \dots, l_6) := \sum_{s\in S_6}(-1)^{|\sigma(s)|}\{r_3'(a_1^{(s)}, a_2^{(s)}, a_3^{(s)}, l_{s(1)}, l_{s(2)}, l_{s(3)})\}. \tag{3.12}$$

LEMMA 3.9. $r_3(l_1, \dots, l_6) = f_6(3)(l_1, \dots, l_6)$.

PROOF. It is sufficient to prove that

$$r_3'(a_1, a_2, a_3, l_1, l_2, l_3) = \frac{\Delta(l_1, l_2, l_4)\Delta(l_2, l_3, l_5)\Delta(l_3, l_1, l_6)}{\Delta(l_1, l_2, l_5)\Delta(l_2, l_3, l_6)\Delta(l_3, l_1, l_4)}.$$

But this follows immediately from the definition (3.10) if we put $f_i(v) := \Delta(l_i, l_{i+3}, v)$, $i = 1, 2, 3$. □

In the previous version of the proof (see **[G2]**) we used the same formulas for the homomorphisms $f_4(3)$ and $f_5(3)$ but a slightly different one for $f_6(3)$ that was not skew-symmetric. Formula 3.9 was discovered by Zagier by the skew-symmetrization of that formula and also by the author.

5.

THEOREM 3.10. $f_5(3) \circ d = \delta \circ f_6(3)$.

PROOF. Computing $\delta \circ f_6(3)$ using formula (3.9) and Lemma (3.8) we get

$$\begin{aligned}\delta \circ f_6(3)(l_1, \dots, l_6) &= \mathrm{Alt}(\{r(l_1 \mid l_2, l_3, l_4, a_3)\}_2 \otimes \Delta(l_1, l_2, l_4)) \\ &= \tfrac{1}{2}\mathrm{Alt}([\{r(l_1 \mid l_2, l_3, l_4, a_3)\}_2 - \{r(l_1 \mid l_2, l_6, l_4, a_3)\}_2] \\ &\qquad \otimes \Delta(l_1, l_2, l_4)).\end{aligned}$$

Here $a_3 = \overline{l_2 l_5} \cap \overline{l_3 l_6}$ and we understand alternation in the same way as in formula (3.11).

The five-term relation for the configuration $(l_1 \mid l_2, l_3, l_6, l_4, a_3)$ gives us

$$\begin{aligned}\delta \circ f_6(3)(l_1, \dots, l_6) = \tfrac{1}{2}\mathrm{Alt}\,[&-\{r(l_1 \mid l_3, l_6, l_4, a_3)\}_2 + \{r(l_1 \mid l_2, l_3, l_6, a_3)\}_2 \\ &- \{r(l_1 \mid l_2, l_3, l_6, l_4)\}_2] \otimes \Delta(l_1, l_2, l_4))\end{aligned} \tag{3.13}$$

Considering the projection onto the line $\overline{l_3 l_6}$ we see that (see Figure 3.4)

$$(l_1 \mid l_3, l_6, l_4, a_3) \equiv (l_4 \mid l_3, l_6, l_1, a_3),$$
$$(l_1 \mid l_2, l_3, l_6, a_3) \equiv (l_2 \mid l_1, l_3, l_6, a_3).$$

So the first two terms in the first factor in (3.13) disappear after alternation and we get

$$\begin{aligned}\delta \circ f_6(3)(l_1, \dots, l_6) &= -\tfrac{1}{2}\mathrm{Alt}(\{r(l_1 \mid l_2, l_3, l_6, l_4)\}_2 \otimes \Delta(l_1, l_2, l_4)) \\ &= -\tfrac{1}{2}\mathrm{Alt}(\{r(l_1 \mid l_2, l_3, l_4, l_5)\}_2 \otimes \Delta(l_1, l_2, l_3)).\end{aligned} \tag{3.14}$$

But this coincides with $f_5(3) \circ d(l_1, \dots, l_6)$ computed using formula (3.5). □

6. The "seven-term" functional equation for the trilogarithm.

THEOREM 3.11. *$f_6(3) \circ d = 0$. $\sum_{i=1}^{7}(-1)^{i-1}\mathscr{L}_3(r_3(l_1, \dots, \hat{l}_i, \dots, l_7)) = 0$ over $\mathbb{C}$.* (3.14)

PROOF. (a) According to Theorem 1.10 one has

$$\delta \circ f_6(3) \circ d = f_5(3) \circ d \circ d = 0,$$

i.e. (because $r_3 = f_6(3)$),

$$\delta \circ \left(\sum_{i=1}^{7}(-1)^{i-1} r_3(l_1, \dots, \hat{l}_i, \dots, l_7)\right) = 0 \quad \text{in } \mathscr{B}_2(F) \otimes F^*.$$

It remains to deform a generic configuration of seven vectors to such a special one that clearly satisfies condition (a) of the theorem. This can be done, for example, using a specialization $a = b = 1$, $c = o$ of explicit formula (3.17) below for the "seven-term" relation corresponding to a configuration represented on Figure 3.5. Applying Theorem 1.15 in the case $n = 3$ we get 3.15. □

REMARK. The "seven-term" functional equation has 840 summands. In order to get a shorter version we need to use a degenerate configuration $(l_1, \dots, l_7)$. For example, let homogeneous coordinates of points l_i be represented by columns of the following matrix (see also Figure 3.5)

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & c \\ 0 & 1 & 0 & 1 & a & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & b & 1 \end{bmatrix}. \tag{3.15}$$

Put

(3.16)

$$R_3(a, b, c) := \bigoplus_{\text{cycle}} \left(\{ca - a + 1\} + \left\{\frac{ca - a + 1}{ca}\right\} + \{c\} + \left\{\frac{(bc - c + 1)}{(ca - a + 1)b}\right\} - \left\{\frac{ca - a + 1}{c}\right\} + \left\{\frac{(bc - c + 1)a}{(ca - a + 1)}\right\} - \left\{\frac{(bc - c + 1)}{(ca - a + 1)bc}\right\} - \{1\} \right) + \{-abc\}.$$

Here $\bigoplus_{\text{cycle}} f(a, b, c) := f(a, b, c) + f(c, a, b) + f(b, c, a)$. The functional equation (3.15) for this special configuration (3.16) has the form

$$\mathscr{L}_3(R_3(a, b, c)) = 0.$$

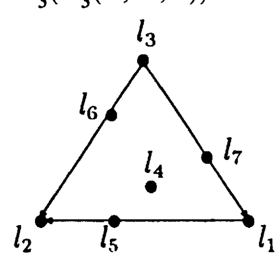


FIGURE 3.5

7. The Grassmannian bicomplex. This is the following bicomplex

$$(3.17)\qquad \begin{array}{ccccccc} & & \downarrow & & \downarrow & & \downarrow \\ \longrightarrow & & C_{n+5}(n+2) & \xrightarrow{d} & C_{n+4}(n+2) & \xrightarrow{d} & C_{n+3}(n+2) \\ & & \downarrow{d'} & & \downarrow{d'} & & \downarrow{d'} \\ \longrightarrow & & C_{n+4}(n+1) & \xrightarrow{d} & C_{n+3}(n+1) & \xrightarrow{d} & C_{n+2}(n+1) \\ & & \downarrow{d'} & & \downarrow{d'} & & \downarrow{d'} \\ \longrightarrow & & C_{n+3}(n) & \xrightarrow{d} & C_{n+2}(n) & \xrightarrow{d} & C_{n+1}(n) \end{array}$$

where

$$d' : (l_1, \dots, l_m) \mapsto \sum_{i=1}^{m} (-1)^{i-1} (l_i \mid l_1, \dots, \hat{l}_i, \dots, l_m).$$

Denote by $(T_*(n), \partial)$ the total complex associated with this bicomplex; $T_{n+1}(n) := C_{n+1}(n)$. Let us define a homomorphism $\psi_*(3)$

$$(3.19)\qquad \begin{array}{ccccc} \longrightarrow T_6(3) & \longrightarrow & T_5(3) & \longrightarrow & T_4(3) \\ \downarrow{\psi_6(3)} & & \downarrow{\psi_5(3)} & & \downarrow{\psi_4(3)} \\ \mathscr{B}_3(F) & \longrightarrow & \mathscr{B}_2(F) \otimes F^* & \longrightarrow & \bigwedge^3 F^* \end{array}$$

as follows. It coincides with the homomorphism (3.2) on the subcomplex $C_*(3) \hookrightarrow T_*(3)$ and is zero on all other groups $C_*(3+i)$.

THEOREM 3.12. *This is a correct definition, i.e.,*

$$\psi_{3+i}(3) \circ d' = 0 \quad \textit{for } i = 1, 2, 3.$$

PROOF. (a) $i = 1$. It is easy to see that

$$\psi_4(3) \circ d' : (l_1, \dots, l_5) \mapsto \operatorname{Alt} \Delta(l_1, l_2, l_3, l_4) \wedge \Delta(l_1, l_2, l_3, l_5) \wedge \Delta(l_1, l_2, l_4, l_5).$$

The right-hand side is zero because we alternate an expression that is symmetric with respect to permutation of l_1 and l_2.

(b) $i = 2$. The

$$\psi_5(3) \circ d' : (l_1, \dots, l_6) \mapsto \tfrac{1}{2} \operatorname{Alt}(\{r(l_1, l_2 \mid l_3, l_4, l_5, l_6)\} \otimes \Delta(l_1, l_2, l_3, l_4)).$$

This is zero for the same reason as above.

(c) $i = 3$. We have to prove

$$(3.20)\qquad \psi_6(3) \left(\sum_{i=1}^{7} (-1)^i (l_i \mid l_1, \dots, \hat{l}_i, \dots, l_7) \right) = 0.$$

This will be done in subsections 8–9. □

8. The duality of configurations (see §7 of [G2]). Let us denote by $\mathrm{Conf}_p(q)$ the set of all configurations of p vectors in a q-dimensional vector space V_q in generic position. There is a duality

$$*\colon \mathrm{Conf}_{m+n}(m) \to \mathrm{Conf}_{m+n}(n), \qquad *^2 = \mathrm{id}$$

that satisfies the following important properties:

(1) $*$ commutes with the action of the permutation group S_{m+n} on vectors of a configuration.

(2) If $*(l_1, \dots, l_{m+n}) = (l'_1, \dots, l'_{m+n})$, then

$$*(l_1, \dots, \hat{l}_i, \dots, l_{m+n}) = (l'_i \mid l'_1, \dots, \hat{l}'_i, \dots, l'_{m+n}),$$

i.e., forgetting of the ith vector of a configuration is dual to the projection along the ith vector.

(3) Let us choose volume forms in V^m and V^n; consider a partition

$$\{1, \dots, m+n\} = \{i_1 < \cdots < i_m\} \cup \{j_1 < \cdots < j_n\}.$$

Then $\Delta(l_{i_1}, \dots, l_{i_m})/\Delta(l'_{j_1}, \dots, l'_{j_n})$ does not depend on a partition.

Three definitions of $*$—the Grassmannian, the coordinate, and the geometric—were suggested in §7 of [G2]. We need only the first two.

(i) *The Grassmannian definition.* Let $(l_1, \dots, l_{m+n})$ be a coordinate frame in a vector space V. Let us denote by $\widehat{G}_m(V, \{l_i\})$ the set of all m-dimensional subspaces V that are in generic position to coordinate hyperplanes. MacPherson constructed in [M] an isomorphism $p\colon \widehat{G}_m(V, \{l_i\}) \xrightarrow{\sim} \mathrm{Conf}_{m+n}(n)$. Namely, $p(h)$ is a configuration formed by images of l_i in V/h. Let $(f^1, \dots, f^{m+n})$ be the dual basis in V^* and $h^\perp\colon \{f \in V^* \mid \langle f, v\rangle = 0 \text{ for any } v \in h\}$. Then the definition of $*$ is given by the diagram

$$\begin{array}{ccc} \widehat{G}_m(V, \{l_i\}) & \xrightarrow[\sim]{\perp} & \widehat{G}_n(V^*, \{f^j\}) \\ {\scriptstyle p\wr}\downarrow & & \downarrow{\scriptstyle \wr p} \\ \mathrm{Conf}_{m+n}(n) & \xrightarrow{*} & \mathrm{Conf}_{m+n}(m). \end{array}$$

(ii) *The coordinate definition.* A configuration of $(m+n)$ vectors in an m-dimensional coordinate space can be represented as columns of the $m \times (m+n)$ matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & a_{11} & \cdots & a_{1n} \\ 0 & 1 & \cdots & 0 & \vdots & & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & a_{m1} & \cdots & a_{mn} \end{pmatrix} = (I_m, A).$$

Then the dual configuration is represented by the $n \times (m+n)$ matrix $(-A^t, I_n)$. These definitions give the same duality. Indeed, the subspace h is generated by $l_{m+i} - \sum_{j=1}^m a_{ij} e_j$ and the subspace $h^\perp$ by $f^j + \sum_{i=1}^n a_{ij} f_{m+i}$.

Now properties (1), (2) follow immediately from the first definition, and (3) is easy to see from the second one.

9. The end of the proof of Theorem 3.12(c).

PROPOSITION 3.13. $\psi_6(3)((l_1, \dots, l_6) + *(l_1, \dots, l_6)) = 0$.

PROOF. If $*(l_1, \dots, l_6) = (l'_1, \dots, l'_6)$ then according to property (3) of $*$ we have

$$\frac{\Delta(l_1, l_2, l_4)\Delta(l_2, l_3, l_5)\Delta(l_3, l_1, l_6)}{\Delta(l_1, l_2, l_5)\Delta(l_2, l_3, l_6)\Delta(l_3, l_1, l_4)} = \frac{\Delta(l'_5, l'_6, l'_3)\Delta(l'_4, l'_6, l'_1)\Delta(l'_4, l'_5, l'_2)}{\Delta(l'_4, l'_6, l'_3)\Delta(l'_4, l'_5, l'_1)\Delta(l'_5, l'_6, l'_2)}$$
$$\equiv \frac{\Delta(l'_4, l'_5, l'_2)\Delta(l'_5, l'_6, l'_3)\Delta(l'_6, l'_4, l'_1)}{\Delta(l'_4, l'_5, l'_1)\Delta(l'_5, l'_6, l'_2)\Delta(l'_6, l'_4, l'_3)}.$$

But $\{x\} = \{x^{-1}\} \bmod \mathscr{R}_3(F)_{\mathbb{Q}}$ and $(1, 2, 3, 4, 5, 6) \mapsto (4, 5, 6, 1, 2, 3)$ is an odd permutation; so Proposition 3.13 is proved. □

Formula (3.20) and hence Theorem 3.12(c) follow immediately from Proposition 3.13 and property (2) of $*$.

10. The bicomplex $C_*^m(n)$. Let us define a differential $d^{(k)}: \widetilde{C}_p(n) \to \widetilde{C}_{p-1}(n)$ as follows: $d^{(k)}: (l_1, \dots, l_p) \mapsto \sum_{i=1}^{p-k}(-1)^{i-1}(l_1, \dots, \hat{l}_{k+i}, \dots, l_p)$. Note that $d^{(0)} \equiv d$; see subsection 1.

LEMMA 3.14. *The following complex is acyclic* $(k > 0)$:

$$\cdots \to \widetilde{C}_{k+2}(n) \xrightarrow{d^{(k)}} \widetilde{C}_{k+1}(n) \xrightarrow{d^{(k)}} \widetilde{C}_k(n).$$

The proof is in complete analogy with the one of Lemma 3.1.

Let $\mathrm{Sym}_k: \widetilde{C}_p(n) \to \widetilde{C}_p(n)$ be the symmetrization of the first k vectors:

$$\mathrm{Sym}_k: (l_1, \dots, l_p) \mapsto \sum_{\sigma \in S_k} \frac{1}{k!}(l_{\sigma(1)}, \dots, l_{\sigma(k)}, l_{k+1}, \dots, l_p).$$

Define a homomorphism $\lambda^{(k)}: \widetilde{C}_p(n) \to \widetilde{C}_p(n)$ as

$$\lambda^{(k)}: (l_1, \dots, l_p) \mapsto \sum_{i>k}(-1)^{i+k-1}\mathrm{Sym}_{k+1}(l_1, \dots, l_k, l_i, \dots, \hat{l}_i, \dots, l_p).$$

LEMMA 3.15. $d^{(k+1)} \circ \lambda^{(k)} = -\lambda^{(k)} \circ d^{(k)}$.

PROOF. This is obvious for the homomorphism $\tilde{\lambda}^{(k)}$ defined by the same formula as $\lambda^{(k)}$ but without symmetrization. It remains to symmetrize the first $k+1$ vectors. □

LEMMA 3.16. $\lambda^{(k+1)} \circ \lambda^{(k)} = 0$.

PROOF. Straightforward. (Note that $\tilde{\lambda}^{(k+1)} \circ \tilde{\lambda}^{(k)} \neq 0$.) □

Therefore, we get the following bicomplex $\widetilde{C}_*^m(n)$

$$
\begin{array}{ccccccc}
\cdots \xrightarrow{d} & \widetilde{C}_3(n) & \xrightarrow{d} & \widetilde{C}_2(n) & \xrightarrow{d} & \widetilde{C}_1(n) \\
& \downarrow \lambda^{(0)} & & \downarrow \lambda^{(0)} & & \downarrow \lambda^{(0)} \\
\cdots \xrightarrow{d^{(1)}} & \widetilde{C}_3(n) & \xrightarrow{d^{(1)}} & \widetilde{C}_2(n) & \xrightarrow{d^{(1)}} & \widetilde{C}_1(n) \\
& \downarrow \lambda^{(1)} & & \downarrow \lambda^{(1)} & & \\
\cdots \xrightarrow{d^{(2)}} & \widetilde{C}_3(n) & \xrightarrow{d^{(2)}} & \widetilde{C}_2(n) & & \\
& \downarrow \lambda^{(2)} & & & & \\
\cdots \xrightarrow{d^{(3)}} & \widetilde{C}_3(n) & & & & \\
& \vdots & & & & \\
& & & & & \downarrow \\
& & & \cdots & \longrightarrow & \widetilde{C}_{m-1}(n).
\end{array}
\tag{3.21}
$$

Remark. The bicomplex $C_*^2(3)$ was considered by Suslin in §3 of [S3].

Let $(\widetilde{\mathscr{D}}_*^m(n), \partial)$ be the simple complex, associated with the bicomplex $\widetilde{C}_*^m(n)$. More precisely, $\widetilde{\mathscr{D}}_k^m(n)$ is the direct sum of the terms $\widetilde{C}_{k+i}(n)$ at level i, for $i = 0, \ldots, m-1$ (the top line has level 0, the bottom line level $m-1$) and has degree $k-1$ in the complex $\widetilde{\mathscr{D}}_*^m(n)$. It is placed at degrees $-1, 0, +1, \ldots$ (∂ has degree -1).

Lemma 3.17. $H^i(\widetilde{\mathscr{D}}_*^m(n)) = \mathbb{Z}$, $i = 0$; 0, $i \neq 0$.

The proof follows immediately from Lemmas 3.14 and 3.15.

The group $\mathrm{GL}_n(F)$ acts naturally on the complex $\widetilde{\mathscr{D}}_*^m(n)$. Let us denote the complex $\widetilde{\mathscr{D}}_*^m(n)_{\mathrm{GL}_n(F)}$ as $\mathscr{D}_*^m(n)$. Lemma 3.17 implies that there is a canonical homomorphism

$$H_*(\mathrm{GL}_n(F), \mathbb{Z}) \to H_*(\mathscr{D}_*^m(n)).$$

Now let us define a homomorphism of complexes

$$
\begin{array}{ccccccccc}
\longrightarrow & \mathscr{D}_6^{(n-2)}(n) & \longrightarrow & \mathscr{D}_5^{(n-2)}(n) & \longrightarrow & \mathscr{D}_4^{(n-2)}(n) & \longrightarrow & & \\
& \downarrow f_3 & & \downarrow f_3 & & \downarrow f_3 & & & \\
\longrightarrow & T_6(3) & \longrightarrow & T_5(3) & \longrightarrow & T_4(3) & \longrightarrow & 0. &
\end{array}
\tag{3.22}
$$

More precisely, we shall define a homomorphism $\tilde{f}$ of the corresponding bicomplex $C_*^{(n-2)}(n)$ to the Grassmannian bicomplex (see 3.18)

$$\begin{array}{ccccccc} & & & & \downarrow & & \downarrow \\ & & & \longrightarrow & C_7(5) & \xrightarrow{d} & C_6(5) \\ & & \downarrow & & \downarrow{\scriptstyle d'} & & \downarrow{\scriptstyle d'} \\ & & C_7(4) & \xrightarrow{d} & C_6(4) & \xrightarrow{d} & C_5(4) \\ & & \downarrow{\scriptstyle d'} & & \downarrow{\scriptstyle d'} & & \downarrow{\scriptstyle d'} \\ & \longrightarrow & C_6(3) & \xrightarrow{d} & C_5(3) & \xrightarrow{d} & C_4(3). \end{array}$$

Namely, if $(l_1, \dots, l_m) \in C_m(p)$ is placed at the level k in the bicomplex $C_*^{n-2}(n)$, i.e., if we apply to $(l_1, \dots, l_m)$ the horizontal differential $d^{(k)}$ (see (3.21)), then we set

$$\tilde{f}\colon (l_1, \dots, l_m) \mapsto (l_1, \dots, l_k \mid l_{k+1}, \dots, l_m) \in C_{m-k}(p-k).$$

Here we use the following notation. Let $(l_1, \dots, l_k, \dots, l_m) \in C_m(V)$. Let us denote by $\langle l_1, \dots, l_n\rangle$ the subspace generated by $l_1, \dots, l_k$. Then

$$(l_1, \dots, l_k \mid l_{k+1}, \dots, l_m)$$

is the configuration of $m-k$ vectors in $V/\langle l_1, \dots, l_k\rangle$.

So we get a homomorphism f_3 of the corresponding total complexes (see (3.22)). The composition of this homomorphism with homomorphism ψ constructed above

$$\begin{array}{ccccccc} & \longrightarrow & T_6(3) & \longrightarrow & T_5(3) & \longrightarrow & T_4(3) \\ & & \downarrow{\scriptstyle \psi_6(3)} & & \downarrow{\scriptstyle \psi_5(3)} & & \downarrow{\scriptstyle \psi_4(3)} \\ 0 & \longrightarrow & \mathscr{B}_3(F) & \longrightarrow & \mathscr{B}_2(F)\otimes F^* & \longrightarrow & \bigwedge^3 F^* \end{array}$$

gives the desired homomorphism of complexes

$$\begin{array}{ccccccc} & \longrightarrow & \mathscr{D}_6^{(n-2)}(n) & \longrightarrow & \mathscr{D}_5^{(n-2)}(n) & \longrightarrow & \mathscr{D}_4^{(n-2)}(n) \\ & & \downarrow{\scriptstyle \psi\circ f} & & \downarrow{\scriptstyle \psi\circ f} & & \downarrow{\scriptstyle \psi\circ f} \\ 0 & \longrightarrow & \mathscr{B}_3(F) & \longrightarrow & \mathscr{B}_2(F)\otimes F^* & \longrightarrow & \bigwedge^3 F^*. \end{array}$$

Therefore, we get the canonical homomorphisms

$$H_i(\mathrm{GL}_n(F)) \to H^{6-i}(\Gamma(F, 3)). \tag{3.23}$$

LEMMA 3.18. *The restrictions of the homomorphisms* (3.23) *to the subgroup* $H_i(\mathrm{GL}_3(F))$ *coincide with those in* (3.3).

PROOF. Choose $n-3$ linearly independent vectors $v_1, \dots, v_{n-3}$ in an n-dimensional vector space V_n and a 3-dimensional complementary subspace

$V_3 : V_n = \langle v_1, \dots, v_{n-3}\rangle \oplus V_3$. Then there is a homomorphism of complexes $\xi\colon C_*(V_3) \to \mathscr{D}_*^{n-2}(V_n)$ where $\xi(C_*(V_3))$ lies in the lowest line of the bicomplex (3.20) and $\xi\colon (l_1, \dots, l_k) \mapsto (v_1, \dots, v_{n-3}, l_1, \dots, l_k)$.

It is clear from the definition that we get a commutative diagram

$$\begin{array}{ccc} C_*(3) & \xrightarrow{\ \xi\ } & \mathscr{D}_*^{(n-2)}(n) \\ & {\scriptstyle (3.2)}\searrow \quad \swarrow{\scriptstyle \psi\circ f} & \\ & \Gamma(F, 3). & \end{array} \quad \square$$

Finally, the restriction of the homomorphisms

$$c_i(3)\colon H_i(\mathrm{GL}_3(F)) \to H^{6-i}(\Gamma(F; 3))$$

to the image of the subgroup $H_i(\mathrm{GL}_2(F))$ is equal to zero, because the resolution $\widetilde{D}_*(3)$ of the trivial $\mathrm{GL}_3(F)$-module $\mathbb{Z}$ has a $\mathrm{GL}_2(F)$-invariant section

$$\begin{array}{ccccc} & & & & \mathbb{Z} \\ & & & & \downarrow \\ \cdots & \longrightarrow & \widetilde{C}_2(3) & \longrightarrow & \widetilde{C}_1(3). \end{array}$$

Namely, if $V_3 = V_2 \oplus \langle v\rangle$, then the formula $n \mapsto n\cdot(v) \in \widetilde{C}_1(3)$ defines a $\mathrm{GL}_2(V_2)$-invariant section $\mathbb{Z} \to \widetilde{C}_*(V_3)$.

So we have constructed homomorphisms

$$C_5^{[2]}\colon K_5^{[2]}(F)_{\mathbb{Q}} \to H^1(\Gamma(F; 3)_{\mathbb{Q}}),$$
$$C_4^{[1]}\colon K_4^{[1]}(F)_{\mathbb{Q}} \to H^2(\Gamma(F; 3)_{\mathbb{Q}}).$$

CONJECTURE 3.19. *Homomorphisms $C_4^{[1]}$, $C_5^{[2]}$ are isomorphisms.*

11. Explicit formula for a five-cocycle representing a class of continuous cohomology of $\mathrm{GL}_3(\mathbb{C})$. Choose a point $x \in \mathbb{C}P^2$. Then there is a measurable cocycle

$$f^{(x)} : \underbrace{\mathrm{GL}_3(\mathbb{C}) \times \cdots \times \mathrm{GL}_3(\mathbb{C})}_{6 \text{ times}} \to \mathbb{R}, \tag{3.24}$$
$$f^{(x)}(g_1, \dots, g_6) := \mathscr{L}_3(r_3(g_1x, \dots, g_6x))$$

where r_3 is the generalized cross-ratio of six points in P^2 (see subsection 4). It is certainly invariant under the left action of $\mathrm{GL}_3(\mathbb{C})$. So the seven-term relation (3.15) for the trilogarithm just means that $f^{(x)}$ is a measurable cocycle of $\mathrm{GL}_3(\mathbb{C})$. Different points x give cohomologous cocycles.

The function $\mathscr{L}_3(z)$ is continuous on $\mathbb{C}P^1$ and hence bounded. So the function $f^{(x)}$ is also bounded. Applying Proposition 1.14 from Chapter III of **[Gu]** we see that the cohomology class of the cocycle (3.24) lies in

$$\mathrm{Im}(H^5_{\mathrm{cts}}(\mathrm{GL}_3(\mathbb{C}), \mathbb{R}) \to H^5(\mathrm{GL}_3(\mathbb{C}), \mathbb{R})).$$

It remains to be proved that the constructed class is a nonzero rational multiple of the Borel class in $H^5_{\mathrm{cts}}(\mathrm{GL}_3(\mathbb{C}), \mathbb{R})$. Several possible proofs were outlined in **[G2]**. In §5 a different proof will be given. It is based on an explicit formula for indecomposable elements in $H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet, \mathbb{R}(3))$.

4. Some arguments for the main conjecture

1. We have already seen before that $L(F)^\vee_{-1}$ must be isomorphic to $F^*_{\mathbb{Q}}$.

2. The Bloch-Suslin complex. Let us define a subgroup $R_2(F) \subset \mathbb{Z}[P^1_F\backslash 0, 1, \infty]$ as

$$R_2(F) := \left\{ \sum_{i=0}^{4} (-1)^i \{r(x_0, \dots, \hat{x}_i, \dots, x_4)\}, \qquad x_i \in P^1_F, \ x_i \neq x_j \right\}.$$

Then $\delta_2(R_2(F)) = 0$ according to Lemma 3.6 ($\delta_2 \colon \{x\} \mapsto (1-x) \wedge x$). So we get a complex $B_F(2)$ (the Bloch-Suslin complex)

$$B_2(F) \xrightarrow{\delta} \bigwedge^2 F^*, \qquad B_2(F) := \frac{\mathbb{Z}[P^1_F\backslash 0, 1, \infty]}{R_2(F)} \tag{4.1}$$

where the group $B_2(F)$ placed in degree 1 and δ has degree $+1$. Let $K_3^{\mathrm{ind}}(F) := \mathrm{Coker}(K_3^M(F) \to K_3(F))$. Using some ideas of Bloch, Suslin proved the following remarkable theorem (see also closely related results of Dupont and Sah **[DS]**, **[Sa]**, who work with homology group of $\mathrm{SL}_2(F)$).

THEOREM 4.1 **[S2]**. *There is an exact sequence*

$$0 \to \mathrm{Tor}(F^*, F^*)^\sim \to K_3^{\mathrm{ind}}(F) \to H^1(B_F(2)) \to 0$$

where $\mathrm{Tor}(F^*, F^*)^\sim$ *is the unique nontrivial extension of* $\mathbb{Z}/2\mathbb{Z}$ *by* $\mathrm{Tor}(F^*, F^*)$.

In particular,

$$H^1(B_F(2)_{\mathbb{Q}}) \cong K_3^{\mathrm{ind}}(F)_{\mathbb{Q}} \cong K_3^{[1]}(F)_{\mathbb{Q}} \cong \mathrm{gr}^2_j K_3(F)_{\mathbb{Q}}.$$

So the complex $B_F(2)$ has the same homology as the complex $L(F)^\vee_{-2} \xrightarrow{\partial} \Lambda^2 L(F)^\vee_{-1}$. Assume that there is a homomorphism of complexes

$$\begin{array}{ccc} B_2(F) & \xrightarrow{\delta} & \bigwedge^2 F^* \\ {\scriptstyle\varphi_2}\downarrow & & \| \\ L(F)^\vee_{-2} & \xrightarrow{\partial} & \bigwedge^2 F^* \end{array} \tag{4.2}$$

that induces isomorphism on cohomologies modulo torsion. Then $\varphi_2 \colon B_2(F) \to L(F)^\vee_{-2}$ must be an isomorphism.

In fact, the existence of a homomorphism of complexes (4.2) can be deduced from results of **[BGSV, BMS]** and standard assumptions about the category $\mathscr{M}_T(F)$. After this, using the Borel theorem, one can prove that the induced homomorphism $H^1(B_F(2)_{\mathbb{Q}}) \to H^1_{(2)}(L(F)_\bullet)$ must be an isomorphism for number fields. Finally, Rigidity Conjecture 1.7 tells us that the same is true for an arbitrary field F (see subsection 12 of §1 in **[Go2]**).

Note that Theorem 4.1 and isomorphism $K_3^{\text{ind}}(F) \cong K_3^{\text{ind}}(F(t))$ imply that the canonical map $B_2(F) \to \mathcal{B}_2(F)$, $(\{x\} \mapsto \{x\})$ is an isomorphism.

3. Weight 3 motivic complexes. Recall that the generalized cross-ratio $r_3\colon C_6(P^2) \to \mathbb{Z}[P^1_F]$ is defined by the formula

$$r_3(l_1, \dots, l_6) = \text{Alt}\left\{\frac{\Delta(l_1, l_2, l_4)\Delta(l_2, l_3, l_5)\Delta(l_3, l_1, l_6)}{\Delta(l_1, l_2, l_5)\Delta(l_2, l_3, l_6)\Delta(l_3, l_1, l_4)}\right\}.$$

Set

$$R_3(F) := \left\{\sum_{i=0}^{6}(-1)^i r_3(l_0, \dots, \hat{l}_i, \dots, l_6),\ \text{ where } (l_0, \dots, l_6) \in C_7(P^2)\right\},$$

$$B_3(F) := \mathbb{Z}[P^1_F]/R_3(F),\ \{0\},\ \{\infty\}.$$

Theorem 3.11 implies that $\delta_3(R_3(F)) = 0$; so we get a complex $B_F(3)$:

$$B_3(F) \xrightarrow{\delta} B_2(F) \otimes F^* \xrightarrow{\delta} \Lambda^3 F^*$$

where $B_3(F)$ placed in degree 1 and δ has degree $+1$.

Let us assume that there is a homomorphism $\varphi_3\colon B_3(F) \to L(F)^\vee_{-3}$ making the following diagram commutative (we have assumed $L(F)^\vee_{-2} \cong B_2(F)_{\mathbb{Q}}$, $L(F)^\vee_{-1} \cong F^*_{\mathbb{Q}}$):

$$\begin{array}{ccc} B_3(F) & \xrightarrow{\delta} & B_2(F) \otimes F^* \\ {\scriptstyle\varphi_3}\downarrow & & \wr\downarrow{\scriptstyle\varphi_2\otimes\varphi_1} \\ L(F)^\vee_{-3} & \xrightarrow{\partial} & L(F)^\vee_{-2} \otimes L(F)^\vee_{-1}. \end{array}$$

Then we get a morphism of complexes

$$\begin{array}{ccccc} B_3(F) & \xrightarrow{\delta} & B_2(F) \otimes F^* & \xrightarrow{\delta} & \bigwedge^3 F^* \\ {\scriptstyle\varphi_3}\downarrow & & \wr\downarrow{\scriptstyle\varphi_2\otimes\varphi_1} & & \wr\downarrow{\scriptstyle\Lambda^3\varphi_1} \\ L(F)^\vee_{-3} & \xrightarrow{\partial} & L(F)^\vee_{-2} \otimes L(F)^\vee_{-1} & \xrightarrow{\partial} & \bigwedge^3 L(F)^\vee_{-1}. \end{array}$$

The bottom complex is just $(\bigwedge^\bullet_{(3)}(L(F)_\bullet), \partial)$, the part of grading three of the cochain complex of the Lie algebra $L(F)_\bullet$.

The results of §3 give considerable evidence for the expected isomorphisms

$$H^i(B_F(3)_{\mathbb{Q}}) \cong H^i\left(\bigwedge\nolimits^\bullet_{(3)}(L(F)_\bullet)\right). \tag{4.3}$$

(According to Conjecture 3.19 and (1.3) both sides are expected to be isomorphic to $K^{[3-i]}_{6-i}(F)_{\mathbb{Q}}$.) (4.3) implies that $\varphi_3\colon B_3(F)_{\mathbb{Q}} \to L(F)^\vee_{-3}$ is an isomorphism. I expect, of course, that $B_3(F)_{\mathbb{Q}} \cong \mathcal{B}_3(F)_{\mathbb{Q}}$.

In any case the complexes $(\bigwedge^{\bullet}_{(n)}(L(F)_{\bullet}), \partial)$ for $n = 1, 2, 3$ look like the complexes $\Gamma(F; n)$. But already the weight four part of the cochain complex of $L(F)_{\bullet}$, that is

$$L(F)^{\vee}_{-4} \xrightarrow{\partial} \bigoplus \begin{matrix} L(F)^{\vee}_{-3} \otimes L(F)^{\vee}_{-1} \\ \bigwedge^2 L(F)^{\vee}_{-2} \end{matrix} \xrightarrow{\partial} L(F)^{\vee}_{-2} \otimes \bigwedge^2 L(F)^{\vee}_{-1} \xrightarrow{\partial} \bigwedge^4 L(F)^{\vee}_{-1} \tag{4.4}$$

looks quite different from $\Gamma(F; 4)$, because we have an extra term $\bigwedge^2 L(F)^{\vee}_{-2}$ $(4 = 2 + 2)$ that has no analog in $\Gamma(F; 4)$. So assuming a homomorphism $\varphi_4: \mathscr{B}_4(F)_{\mathbb{Q}} \to L(F)^{\vee}_{-4}$ making (2.1b) commutative we get a morphism of complexes $\tilde{\varphi}_4: \Gamma(F; 4) \to (\bigwedge^{\bullet}_{(4)}(L(F)_{\bullet}), \partial)$, but it cannot be an isomorphism. However,

Theorem 4.2. *$\tilde{\varphi}_4: H^3\Gamma(F; 4) \to H^3(\bigwedge^{\bullet}_{(4)}(L(F)_{\bullet}), \partial)$ is an isomorphism.*

Proof. Set

(4.5)

$$\kappa(x, y) := \varphi_3\left[-\{1-x\} - \{1-y\} + \left\{\frac{1-x}{1-y}\right\} - \left\{\frac{1-x^{-1}}{1-y^{-1}}\right\}\right] \otimes \frac{x}{y}\varphi_3\{x\} \otimes (1-y) - \varphi_3\{y\} \otimes (1-x) + \varphi_3\left\{\frac{x}{y}\right\} \otimes \frac{1-x}{1-y} - \varphi_2\{x\} \wedge \varphi_2\{y\}$$

that lies in

$$L(F)^{\vee}_{-3} \otimes L(F)^{\vee}_{-1} \oplus \bigwedge^2 L(F)^{\vee}_{-2} = B_3(F) \otimes F^* \oplus \bigwedge^2 B_2(F).$$

Lemma 4.3. $\partial(\kappa(x, y)) = 0$.

Proof. Direct calculation. □

Note that

$$\kappa(x, y) + \varphi_2\{x\} \wedge \varphi_2\{y\} \subset (\varphi_3 \otimes \varphi_1)(B_3(F) \otimes F^*) = L(F)^{\vee}_{-3} \otimes L(F)^{\vee}_{-1}.$$

So it follows from Lemma 4.3 that

$$\partial\left(\bigwedge^2 L(F)^{\vee}_{-2}\right) \subset \partial(L(F)^{\vee}_{-3} \otimes L(F)^{\vee}_{-1}).$$

But this is the only fact that we need in order to prove Theorem 4.2. □

Corollary 4.4. *Assume that for $n = 1, 2, 3$ we have isomorphisms $\varphi_n: \mathscr{B}_n(F)_{\mathbb{Q}} \xrightarrow{\sim} L(F)^{\vee}_{-n}$ making diagram (2.1b) commutative. Then*

$$H^{n-1}_{(n)}(L(F)_{\bullet}) \cong \frac{\operatorname{Ker}(B_2(F)_{\mathbb{Q}} \otimes \bigwedge^{n-2} F^*_{\mathbb{Q}} \to \bigwedge^n F^*_{\mathbb{Q}})}{\{x\}_2 \otimes x \wedge \bigwedge^{n-3} F^*_{\mathbb{Q}}}.$$

Proof. The left-hand side is just the cohomology of the following complex

$$\bigoplus \begin{matrix} L^{\vee}_{-3} \otimes \bigwedge^{n-3} L^{\vee}_{-1} \\ \bigwedge^2 L^{\vee}_{-2} \otimes \bigwedge^{n-4} L^{\vee}_{-1} \end{matrix} \xrightarrow{\partial} L^{\vee}_{-2} \otimes \bigwedge^{n-2} L^{\vee}_{-1} \xrightarrow{\partial} \bigwedge^n L^{\vee}_{-1}.$$

It remains to apply Theorem 4.2. □

Lemma 4.3 tells us that an element $\varphi_4(x, y) \in L(F)^\vee_{-4}$ should exist such that

$$\partial \varphi_4(x, y) = \kappa(x, y).$$

(The reason is that $\Gamma(F, n)_{\mathbb{Q}}$ should be a "resolution" for $K_n^M(F)$. See appendix in **[G2]**.) Let us assume that such $\varphi_4(x, y)$ exists.

5. Weight five motivic complexes. The part of grading five of the cochain complex of $L(F)_\bullet$ looks as follows

$$L^\vee_{-5} \xrightarrow{\partial} \bigoplus \begin{matrix} L^\vee_{-4} \otimes L^\vee_{-1} \\ L^\vee_{-3} \otimes L^\vee_{-2} \end{matrix} \xrightarrow{\partial} \oplus \begin{matrix} L^\vee_{-3} \otimes \Lambda^2 L^\vee_{-1} \\ \bigwedge^2 L^\vee_{-2} \otimes L^\vee_{-1} \end{matrix} \xrightarrow{\partial} L^\vee_{-2} \otimes \bigwedge^3 L^\vee_{-1} \xrightarrow{\partial} \bigwedge^5 L^\vee_{-1}.$$

We would like to prove that the component $\partial_{3,2}\colon L^\vee_{-5} \to L^\vee_{-3} \otimes L^\vee_{-2}$ of the coboundary ∂ is an epimorphism. Unfortunately it is not quite clear how to construct an element in $L^\vee_{-5}$ because $L^\vee_{-5}$ itself is a quite mysterious object. However, assuming the existence of $\phi_4(x, y)$ we can find an element in $L^\vee_{-4} \otimes L^\vee_{-1} \oplus L^\vee_{-3} \otimes L^\vee_{-2}$ with zero coboundary, whose component in $L^\vee_{-3} \otimes L^\vee_{-2}$ is $\varphi_3\{x\} \otimes \varphi_2\{y\}$. We expect that such a cycle should be in $\varphi(L^\vee_{-5})$.

Let us do this. We assume that a $\varphi_4\colon \mathscr{B}_4(F) \to L(F)^\vee_{-4}$ making (2.1b) commutative exists. Consider the element

$$(4.6) \quad \begin{aligned} \phi_5(x, y) := \phi_4(x, y) \otimes \frac{x}{y} + \varphi_4\left\{\frac{x}{y}\right\} \otimes \frac{1-x}{1-y} + \varphi_4\{x\} \otimes (1-y) \\ + \varphi_4(y) \otimes (1-x) - \varphi_3\{x\} \otimes \varphi_2\{y\} - \varphi_3\{y\} \otimes \varphi_2\{x\}. \end{aligned}$$

LEMMA 4.5. $\partial \phi_5(x, y) = 0$.

PROOF. Direct calculations using formula (4.5) for $\partial \phi_4(x, y) = \kappa(x, y)$.

The $L^\vee_{-3} \otimes L^\vee_{-2}$ component of $-\frac{1}{2}(\phi_5(x, y) + \phi_5(x, y^{-1}))$ is equal to $\varphi_3\{x\} \otimes \varphi_2\{y\}$ because $\{y\}_2 + \{y^{-1}\}_2 = 0$ in $B_2(F)_{\mathbb{Q}}$ and $\{y\}_3 = \{y^{-1}\}_3$ in $B_3(F)_{\mathbb{Q}}$.

We can pursue this idea further and "construct" by induction elements $\phi_n(x, y) \in L(F)^\vee_{-n}$ (using the same assumptions as above) such that

$(4.7)_n$

$$\begin{aligned} \partial \phi_n(x, y) = \phi_{n-1}(x, y) \otimes \frac{x}{y} + \varphi_{n-1}\left\{\frac{x}{y}\right\} \otimes \frac{1-x}{1-y} \\ + \sum_{k=1}^{[n/2]} (-1)^{k-1} (\varphi_{n-k}\{x\} \otimes \varphi_k\{y\} + (-1)^{n-k} \varphi_{n-k}\{y\} \otimes \varphi_k\{x\}) \end{aligned}$$

for n odd; for n even we have the same formula, but the last term will be $(-1)^{n/2-1} \varphi_{n/2}\{x\} \wedge \varphi_{n/2}\{y\}$. (Here $\varphi_1(a) := 1 - a \in F^*$.)

PROPOSITION 4.6. *Suppose that $\partial \phi_{n-1}(x, y)$ is given by formula $(4.7)_{(n-1)}$. Then the coboundary of the right-hand side of $(4.7)_n$ is equal to 0.*

PROOF. Direct calculation using the formula

$$\partial\left(\phi_{n-1}(x, y)\otimes\frac{x}{y}+\varphi_{n-1}\left\{\frac{x}{y}\right\}\otimes\frac{1-x}{1-y}\right)$$
$$=\sum_{k=1}^{[n/2]}(-1)^{k-1}(\varphi_{n-k}\{x\}\otimes\varphi_k\{y\}+(-1)^{n-k}\varphi_{n-k}\{y\}\otimes\varphi_k\{x\})$$

(for n odd the last term in this sum should be $(-1)^{(n-1)/2-1}\varphi_{(n-1)/2}\{x\}\wedge\varphi_{(n-1)/2}\{y\}$.

6. Nonexistence of natural generators for $L(F)_{\leq -2}$ inside $L(F)_\bullet$. Let us choose a splitting $s\colon \mathscr{B}_4^\vee\to L_{-4}$ of the exact sequence

$$0\to[L_{-2}, L_{-2}]\to L_{-4}\overset{s}{\leftrightarrows}\mathscr{B}_4^\vee\to 0.$$

This means that we make a choice of degree -4 generators for $L(F)_{\leq -2}$. Then the composition of the commutator map $L_{-3}\otimes L_{-1}\to L_{-4}$ with the projection of L_{-4} along $s(\mathscr{B}_4^\vee)$ gives us a homomorphism

$$L_{-3}\otimes L_{-1}\to\bigwedge^2 L_{-2}.$$

Assume that $L(F)_{-i}=B_i(F)^\vee$ for $i=1, 2, 3$. Then dualizing we get a homomorphism

$$p\colon B_2(F)\wedge B_2(F)\to B_3(F)\otimes F^*. \tag{4.8}$$

The following result, proved in collaboration with Zagier, shows that there is no such reasonable nonzero map! More precisely, let us call a map p *natural* if it is given by the formula

$$p\colon\{x\}_2\wedge\{x\}_2\mapsto\sum_i\{\varphi_i(x, y)\}_3\otimes\psi_i(x, y) \tag{4.9}$$

where $\varphi_i(x, y)$ and $\psi_i(x, y)$ are rational functions with coefficients in $\mathbb{Q}$.

THEOREM 4.7. *There is no natural nonzero homomorphism* (4.8).

PROOF. In the case $F=\mathbb{C}$ there is a homomorphism

$$l\colon B_3(\mathbb{C})\otimes\mathbb{C}^*\to B_2(\mathbb{C})\otimes\mathbb{C}^*\otimes\mathbb{C}^*\to R,$$
$$l\colon\{z_1\}_3\otimes z_2\mapsto\mathscr{L}_2(z_1)\cdot\log|z_1|\cdot\log|z_2|.$$

Consider the composition

$$\begin{gathered}B_2(\mathbb{C})\wedge B_2(\mathbb{C})\xrightarrow{p}B_3(\mathbb{C})\otimes\mathbb{C}^*\xrightarrow{l}R,\\ l\circ p\colon\{x\}_2\wedge\{y\}_2\mapsto\sum_i\mathscr{L}_2(\varphi_i(x, y))\cdot\log|\varphi_i(x, y)|\cdot\log|\psi_i(x, y)|.\end{gathered} \tag{4.10}$$

The right-hand side of (4.10) satisfies the five-term functional equation in variable x (as well as in y) because both p and l are homomorphisms and so $l\circ p(R_2(\mathbb{C})\wedge\{y\}_2)=0$. From the other hand we have the following beautiful result of Bloch **[Bl1]**):

THEOREM 4.8. *Let $f(z)$ be a measurable function satisfying the five-term relation* $\sum_{i=1}^{5}(-1)^i\mathscr{L}_2(r(x_1, \dots, \hat{x}_i, \dots, x_5)) = 0$. *Then* $f(z) = \lambda \cdot \mathscr{L}_2(z)$ *for some* $\lambda \in \mathbb{C}$.

Applying this theorem to the right-hand side of (4.10) considered as a function in x and then as a function in y we get

$$(4.11)\qquad \sum_i \mathscr{L}_2(\varphi_i(x, y)) \cdot \log|\varphi_i(x, y)| \cdot \log|\psi_i(x, y)| = \lambda \cdot \mathscr{L}_2(x) \cdot \mathscr{L}_2(y).$$

The left expression is skewsymmetric in x, y because of its definition (4.10), while $\lambda \cdot \mathscr{L}_2(x) \cdot \mathscr{L}_2(y)$ is obviously symmetric. So $\lambda = 0$.

There is another argument: the right-hand side of (4.11) is invariant under the involution $x \mapsto \bar{x}$, $y \mapsto \bar{y}$, while the left one is skew invariant. (It works for a homomorphism $\tilde{p}\colon B_2 \otimes B_2 \to B_3 \otimes F^*$.) Therefore, $\lambda = 0$.

This is the crucial point and now it becomes absolutely clear that Theorem 4.7 is true. However, we shall present a rigorous proof.

Let us choose a generic number $y_0 \in \mathbb{C}$. There is a natural basis $(x - a)$, $a \in \mathbb{C}$, in the $\mathbb{Q}$-vector space $\mathbb{C}(x)^*/\mathbb{C}^* \otimes \mathbb{Q}$. Using this basis we can rewrite (4.9) as follows $(\alpha \in \mathbb{C}^*)$:

$$\begin{aligned}\sum_i \{\varphi_i(x, y_0)\}_3 \otimes \psi_i(x, y_0)\\ = \sum_{i,j} n_j^i\{f_i^j(x)\}_3 \otimes (x - a_i) + \sum_j n_j^0\{f_0^j(x)\}_3 \otimes \alpha.\end{aligned}$$

Then (4.11) is of the form $\sum_i A_i(x) + A_0(x)$ where

$$(4.12)\qquad A_i(x) := \sum_{i,j} n_j^i \mathscr{L}_2(f_i^j(x)) \cdot \log|f_i^j(x)| \cdot \log|x - a_i|.$$

The function $\mathscr{L}_2(z)$ is real-analytic on $\mathbb{C}P^1\backslash\{0, 1, \infty\}$, continuous on $\mathbb{C}P^1$, and has a singularity of type $r \cdot \log r$ at $z = 0, 1, \infty$. Therefore, for any $k > 0$ the functions $A_k(x)$ and $A_{\neq k}(x) := \sum_{i \neq k} A_i(x) + A_0(x)$ have the following singularity near $x = a_k$:

$$\begin{aligned} A_k(x)&\colon r^{2m}\log^{m+1} r \quad \text{or} \quad r^m \log^{m+2} r \qquad (m \geq 0),\\ A_{\neq k}(x)&\colon r^{2m}\log^m r \quad \text{or} \quad r^m \log^{m+1} r \qquad (m \geq 1).\end{aligned}$$

For example, if $f_k^j(x) = 1 - c\cdot(x - a_k)^m + \cdots$ then $A_k(x)$ has a singularity of type $r^{2m}\log^{m+1} r$. Fortunately all pairs $(2m, m+1)$, $m \geq 0$; $(m, m+2)$, $m \geq 0$; $(2m, m)$, $m \geq 1$; $(m, m+1)$, $m \geq 1$ are different. (For example $(2m, m+1) = (m, m+1)$ only if $m = 0$, but in our situation $m \geq 1$ for $(m, m+1)$.) This means that the singularities of $A_k(x)$ never coincide with the one of $A_{\neq k}(x)$ and hence $A_k(x) + A_{\neq k}(x) = 0$ implies

$$(4.13)\qquad \sum_j n_j^k \mathscr{L}_2(f_k^j(x)) \cdot \log|f_k^j(x)| \equiv 0.$$

Now let us prove that

$$\sum_j n_j^k \{f_k^j(x)\}_2 \otimes f_k^j(x) = 0 \quad \text{in } B_2(\mathbb{C}(x)) \otimes \mathbb{C}(x)^*.$$

Let us decompose this element using our basis in $\mathbb{C}(x)^*/\mathbb{C}^*$:

$$\begin{aligned}\sum n_j^k \{f_k^j(x)\}_2 \otimes f_k^j(x) \\ = \sum_{m,n} c_n^m \{g_m^n(x)\}_2 \otimes (x - b_m) + \sum c_n^0 \{g_0^n(x)\}_2 \otimes \beta.\end{aligned}$$

Then (4.13) is of the form

$$\sum_{m,n} c_n^m \mathscr{L}_2(f_k^j(x)) \cdot \log|x - b_m| + \sum c_n^0 \mathscr{L}_2(g_0^n(x)) \cdot \log|\beta| = 0.$$

Looking on the type of singularities of this expression near $x = b_m$ it is easy to see that for any m

$$\sum_n c_n^m \mathscr{L}_2(g_m^n(x)) \equiv 0.$$

Proposition 4.9. *If $\sum_n c_n \mathscr{L}_2(f_n(x)) \equiv 0$ for some $f_n(x) \in \mathbb{C}(x)$ then for all complex numbers z*

$$\sum_n c_n \{f_n(z)\}_2 - \sum_n c_n \{f_n(0)\}_2 = 0 \quad \textit{in } B_2(\mathbb{C}).$$

Proof. Let

$$\delta_2\left(\sum_n c_n \{f_n(x)\}_2\right) = \sum_i (x - \alpha_i) \wedge (x - \beta_i) + \sum_j \delta_j \wedge (x - \gamma_j) + \sum_i \varepsilon_i \otimes \xi_i.$$

Then

$$\begin{aligned} 0 &= d\left(\sum_n c_n \mathscr{L}_2(f_n(x))\right) \\ &= \sum_i -(\log|x - \alpha_i| \cdot d\arg(x - \beta_i) \\ &\quad + \log|x - \beta_i| d\arg(x - \alpha_i)) - \sum_j \log|\delta_j| d\arg(x - \gamma_i). \end{aligned}$$

Now look at the singularities of the right-hand side at $x = \alpha_i$. The first term has a singularity of type $\log r$, while $d\arg(x - \alpha_i)$ has a different type of singularity because

$$d\arg z = \frac{-y\,dx + x\,dy}{x^2 + y^2}, \qquad z = x + iy.$$

Therefore, $\delta_2(\sum_n c_n \{f_n(x)\}_2) = 0$ and so by definition

$$\sum_n c_n \{f_n(z)\}_2 - \sum_n c_n \{f_n(0)\}_2 \in R_2(\mathbb{C}). \quad \square \tag{4.14}$$

Let us decompose the element $(\delta_3 \otimes \mathrm{id}) \circ p(\{x\}_2 \wedge \{y_0\}_2)$ using the basis $(x-b_j)\otimes(x-a_i)$, $(x-b_j)\otimes\alpha_i$, $\beta_j\otimes(x-\alpha_i)$, $\beta_j\otimes\alpha_i$ in $\mathbb{C}(x)^*_{\mathbb{Q}}\otimes\mathbb{C}(x)^*_{\mathbb{Q}}$:

$$(\delta_3 \otimes \mathrm{id}) \circ p(\{x\}_2 \wedge \{y_0\}_2) = \sum_{i,j} (\alpha_{ij})_2 \otimes (x-b_j)\otimes(x-a_i) + \cdots$$

where $(\alpha_{ij})_2 \in B_2(\mathbb{C})$. Insert into this formula the five-term relation

$$\{x\}_2 - \{z\}_2 + \{z/x\}_2 - \left\{\frac{1-x^{-1}}{1-z^{-1}}\right\}_2 + \left\{\frac{1-x}{1-z}\right\}_2$$

instead of $\{x\}_2$. It is easy to see that for generic $z \in \mathbb{C}$ $(x-b_j)\otimes(x-a_i)$ will appear with coefficient $(\alpha_{ij})_2$. Hence $(\alpha_{ij})_2 = 0$. Pursuing further this argument we get

$$(\delta_3 \otimes \mathrm{id}) \circ p(\{x\}_2 \wedge \{y_0\}) = 0 \quad \text{in } B_2(\mathbb{C}(x)) \otimes \mathbb{C}(x)^* \otimes \mathbb{C}(x)^*.$$

So for any $x_0 \in \mathbb{C}$

$$p(\{x\}_2 \wedge \{y_0\}_2) - p(\{x_0\}_2) = 0 \quad \text{in } B_3(\mathbb{C}) \otimes \mathbb{C}^*.$$

The same argument with the five-term relation as above shows that in fact $p(\{x\}_2 \wedge \{y_0\}_2) = 0$. Using this it is easy to complete the proof of Theorem 4.7. □

7. Recall that one of the Beilinson-Lichtenbaum axioms predicts existence of the tensor product of motivic complexes $\Gamma(n) \overset{L}{\otimes} \Gamma(m) \to \Gamma(n+m)$ defined in the derived category. Theorem 4.7 implies that, for our complexes $\Gamma(F; n)_{\mathbb{Q}}$, natural tensor product exists as a morphism in the derived category *only* and cannot be defined at the level of complexes even for $m = n = 2$.

Indeed, an essential ingredient of construction of a natural morphism of complexes

$$\begin{array}{c} \left[\left(B_2 \xrightarrow{\delta} \bigwedge^2 F^*\right) \otimes \left(B_2 \xrightarrow{\delta} \bigwedge^2 F^*\right)\right] \\ \big\downarrow m_{2,2} \\ \left[\mathscr{B}_4 \xrightarrow{\delta} \mathscr{B}_3 \otimes F^* \xrightarrow{\delta} \mathscr{B}_2 \otimes \bigwedge^2 F^* \xrightarrow{\delta} \bigwedge^4 F^*\right] \end{array}$$

is the existence of the following commutative diagram

$$\begin{array}{ccc} B_2 \otimes B_2 & \xrightarrow{\delta\otimes\mathrm{id} - \mathrm{id}\otimes\delta} & B_2 \otimes \bigwedge^2 F^* \oplus \bigwedge^2 F^* \otimes B_2 \\ \big\downarrow m^{(2)}_{2,2} & & \big\downarrow m^{(3)}_{2,2} \\ \mathscr{B}_3 \otimes F^* & \xrightarrow{\delta} & \mathscr{B}_2 \otimes \bigwedge^2 F^*. \end{array} \tag{4.15}$$

But $m^{(2)}_{2,2}$ must be zero by Theorem 4.7 and $m^{(3)}_{2,2}$ should be equal to $(\mathrm{id}, \mathrm{id}\circ s)$ where s is the switch, so (4.15) cannot be commutative.

I am completely sure there is the same situation with tensor products of complexes $\Gamma(F, *)$ for any $m \geq 2$, $n \geq 2$.

Notice that we have a natural homomorphism

$$\delta(k)\colon \mathscr{B}_n \to \mathscr{B}_{n-k} \otimes \underbrace{F^* \otimes \cdots \otimes F^*}_{k \text{ times}}$$

$$\delta(k) := (\delta \otimes \mathrm{id}) \circ \delta(k-1); \qquad \delta(1) := \delta.$$

CONJECTURE 4.10. *The only nontrivial natural homomorphisms* $\bigotimes_i \mathscr{B}_i \to \bigotimes_j \mathscr{B}_j$ *are (up to a permutation) tensor products of the homomorphisms* $\delta(k)$.

Finally, look at the tensor product $\Gamma(1) \otimes \Gamma(1) \to \Gamma(2)$, i.e., $F^* \otimes F^* \to \Gamma(2)$. Theorem 4.1 suggests that it should be defined in the derived category $F^* \overset{L}{\otimes} F^* \to \Gamma(2)$, providing $\mathrm{Tor}(F^*, F^*) \subset H^1(\Gamma(2))$.

8. The structure of the Lie algebra $L(F)_\bullet$: Conclusions. For each $n \geq 1$ there is a canonical monomorphism

$$\varphi_n\colon \mathscr{B}_n(F) \to L(F)^{\vee}_{-n}.$$

It is isomorphism for $n = 1, 2, 3$. However, *it is not surjective for* $n \geq 4$. Let us suppose $B_n(F) = \mathscr{B}_n(F)$ for $n = 1, 2, 3$. Let us define groups $L(F)^{\vee}_{-n}$ for $n = 4, 5$. Set

$$L(F)^{\vee}_{-4} := \frac{\mathbb{Z}[P^1_F] \oplus \bigwedge^2 \mathbb{Z}[P^1_F]}{\mathscr{R}l_4(F)}.$$

Here $\mathscr{R}l_4(F)$ is the "connected component of zero" of the kernel of the homomorphism

$$\partial_4\colon \mathbb{Z}[P^1_F] \oplus \bigwedge^2 \mathbb{Z}[P^1_F] \to B_3(F) \otimes F^* \oplus \bigwedge^2 B_2(F),$$
$$\{a\} + \{x\} \wedge \{y\} \mapsto \{a\} \otimes a + \kappa(x, y)$$

where $\kappa(x, y)$ is given by formula (4.5). More precisely, this means the following (compare with Definition 1.10). Let $\mathscr{A}l_4(F) := \operatorname{Ker} \partial_4$. Any element

$$\alpha(t) = \{a_i(t)\} + \{x_i(t)\} \wedge \{y_i(t)\} \in \mathbb{Z}[P^1_{F(t)}] + \bigwedge^2 \mathbb{Z}[P^1_{F(t)}]$$

has a specialization

$$\alpha(t_0) = \{\alpha_i(t_0)\} + \{x_i(t_0)\} \wedge \{y_i(t_0)\} \in \mathbb{Z}[P^1_F] + \bigwedge^2 \mathbb{Z}[P^1_F].$$

DEFINITION 4.11. $\mathscr{R}l_4(F)$ is generated by elements $\alpha(0)-\alpha(1)$ where $\alpha(t)$ runs through all elements of $\mathscr{A}l_4(F(t))$ and also $\{0\}$, $\{\infty\}$, $\{0\}\wedge\{x\}$, $\{\infty\}\wedge\{x\}$.

The natural map

$$\mathbb{Z}[P^1_F] \to \mathbb{Z}[P^1_F] \oplus \bigwedge^2 \mathbb{Z}[P^1_F],$$
$$\{a\} \mapsto \{a\}$$

induces an inclusion $\mathscr{R}_4(F) \hookrightarrow \mathscr{R}l_4(F)$. So we get a homomorphism

$$\varphi_4\colon \mathscr{B}_4(F) \to L(F)^{\vee}_{-4}.$$

It is clear from the definitions of subgroups $\mathscr{R}_4(F)$ and $\mathscr{R}l_4(F)$ that φ_4 is a monomorphism. Further, there is a natural projection

$$\tilde{p}_4 : \mathbb{Z}[P^1_F] \oplus \bigwedge^2 \mathbb{Z}[P^1_F] \to \bigwedge^2 B_2(F).$$

It follows from the definition that $\tilde{p}_4(\mathscr{R}l_4(F)) = 0$; so we get an epimorphism

$$p_r : L(F)^{\vee}_{-4} \to \bigwedge^2 B_2(F).$$

Therefore, we obtain a sequence

$$0 \to \mathscr{B}_4(F) \xrightarrow{\varphi_4} L(F)^{\vee}_{-4} \xrightarrow{p_4} \bigwedge^2 B_2(F) \to 0.$$

It should be exact. To prove this, one should find an element $S(x, y, z) \in \mathscr{B}_4(F)$ such that

$$\delta_4(S_4(x, y; z)) + \tilde{k}(r_2(x, y), z) = 0, \qquad S_4(0, 0; 0) = 0,$$

where

$$r_2(x, y) := \{x\} - \{y\} + \left\{\frac{x}{y}\right\} - \left\{\frac{1 - x^{-1}}{1 - y^{-1}}\right\} + \left\{\frac{1 - x}{1 - y}\right\}$$

and $\tilde{k}(\cdot, \cdot)$ is a homomorphism

$$\tilde{k}(\cdot, \cdot) : \bigwedge^2 \mathbb{Z}[P^1_F] \to B_3(F) \otimes F^* \oplus \bigwedge^2 B_2(F)$$

given on generators by the formula $\tilde{k}(x, y) := k(x, y)$. Let us suppose that such an element exists. Then $S(x, y; z) + r_2(x, y) \wedge \{z\} \in \mathscr{R}l_4(F)$. On the other hand, $\operatorname{Ker} \tilde{p} = \mathbb{Z}[P^1_F] \oplus R_2(F) \wedge \mathbb{Z}[P^1_F]$ and the subgroup $R_2(F)$ is generated by elements $r_2(x, y)$. So $\operatorname{Ker} p_4 = \mathscr{B}_4(F)$. The homomorphism ∂_4 provides us with a coboundary

$$\partial_4 : L(F)^{\vee}_{-4} \to B_3(F) \otimes F^* \oplus \bigwedge^2 B_2(F).$$

According to Lemma 4.3 the composition

$$L(F)^{\vee}_{-4} \xrightarrow{\partial_4} B_3 \otimes F^* \oplus \bigwedge^2 B_2 \xrightarrow{\partial_4} B_2 \otimes \bigwedge^2 F^*$$

is equal to zero, where the second coboundary is $\delta_3 + \delta_2 \wedge \mathrm{id}$. Theorem 4.9 just means that there is no natural section $s_4 : \bigwedge^2 B_2(F) \to L(F)^{\vee}_{-4}$ $(p_4 \circ s_4 = \mathrm{id})$. Indeed, the composition

$$s_4 \circ \partial_4 : \bigwedge^2 B_2(F) \to B_3(F) \otimes F^* \oplus \bigwedge^2 B_2(F)$$

must have a nonzero $(B_3 \otimes F^*)$-component. But according to Theorem 4.9 there is no such natural nonzero homomorphism. So the sequence has no natural splitting. Let us denote by $\phi_4(x, y)$ the image of $\{x\} \wedge \{y\}$ to $L(F)^{\vee}_{-4}$. Set

$$L(F)^{\vee}_{-5} := \frac{\mathbb{Z}[P^1_F] \oplus \mathbb{Z}[P^1_F] \otimes \mathbb{Z}[P^1_F]}{\mathscr{R}l_5(F)}$$

where $\mathscr{R}l_5(F)$ is the "connected component of zero" of the kernel of the following homomorphism

$$\partial_5\colon \{a\} + \{x\} \otimes \{y\} \mapsto \phi_4\{a\} \otimes a + \Phi_5(x, y)$$

where (compare with (4.6))

$$\Phi_5(x, y) = \tfrac{1}{2}(\tilde{\phi}_5(x, y) - \tilde{\phi}_5(x, y^{-1}))$$

and

$$\begin{aligned}\tilde{\phi}_5(x, y) := \phi_4(x, y) \otimes \frac{x}{y} + \phi_4\left\{\frac{x}{y}\right\} \otimes \frac{1-x}{1-y} + \phi_4\{x\} \\ \otimes (1-y) + \phi_4\{y\} \otimes (1-x) \\ - \phi_3\{x\} \otimes \phi_2\{x\} - \phi_3\{y\} \otimes \phi_2\{x\}.\end{aligned}$$

There is a sequence

$$0 \to \mathscr{B}_5(F) \xrightarrow{\phi_5} L(F)^\vee_{-5} \xrightarrow{p_5} B_3(F) \otimes B_2(F) \to 0$$

where ϕ_5 is a monomorphism and p_5 is an epimorphism. This sequence should be exact. To prove this one should find some elements $S_5(x, y; z)$ and $S_5((l_0, \dots, l_6); z)$ in $\mathscr{B}_5(F)$ such that

$$\delta_5(S_5(x, y; z)) + \widetilde{\Phi}_5(\{z\}, r_2(x, y)) = 0,$$

$$\delta_5(S_5(l_0, \dots, l_6)) + \widetilde{\Phi}_5\left(\sum_{i=1}^{6}(-1)^i r_3(l_0, \dots, \hat{l}_i, \dots, l_6), \{z\}\right) = 0.$$

Here $\widetilde{\Phi}_5(\cdot, \cdot)$ is considered as a homomorphism

$$\widetilde{\Phi}_5(\cdot, \cdot)\colon B_3(F) \otimes B_2(F) \to \mathscr{B}_4(F) \otimes F^* \oplus B_3(F) \otimes B_2(F)$$

and r_3 is the generalized cross-ratio (3.9). This sequence has no natural splitting. In any case we have a well-defined coboundary

$$\partial_5\colon L(F)^\vee_5 \to L(F)^\vee_{-4} \otimes F^* \oplus B_3(F) \otimes B_2(F).$$

The composition

$$\partial_5\colon L(F)^\vee_{-5} \to L(F)^\vee_{-4} \otimes F^* \oplus B_3(F) \otimes B_2(F) \to \bigoplus \bigwedge^2 B_2(F)^*$$

is equal to zero (the second differential comes from the standard cochain complex of the Lie algebra $L(F)_\bullet$). So we have constructed the weight five part of the standard cochain complex of $L(F)_\bullet$. I believe that there is a similar construction of the groups $L(F)^\vee_{-n}$ for all n. This would imply a precise construction of the motivic Lie algebra $L(F)_\bullet$. Explicit formulas for $S_4(x, y; z)$, $S_5(z; x, y)$, and $S_5(l_0, \dots, l_6; z)$ would give us functional equations for 4- and 5-logarithms and an explicit definition of tensor product of motivic complexes

$$\Gamma(F, n)_{\mathbb{Q}} \overset{L}{\otimes} \Gamma(F, m)_{\mathbb{Q}} \to \Gamma(F, n+m)_{\mathbb{Q}}$$

for $n = 2$, $m = 2$ and $n = 2$, $m = 3$. Let us consider in more detail the case $n = 2$, $m = 2$. Let $\widetilde{R}_2(F)$ be a free abelian group generated by configurations $(l_0, \dots, l_4)$ of five distinct points in P^1_F. There is a complex $\widetilde{B}_F(2)$:

$$0 \to \widetilde{R}_2(F) \xrightarrow{\delta} \mathbb{Z}[P^1_F] \xrightarrow{\delta} \bigwedge^2 F^*$$

where $\delta(l_0, \dots, l_4) := \sum_{i=0}^4 \{r(l_0, \dots, \hat{l}_i, \dots, l_4)\}$ and $\widetilde{R}_2(F)$ placed in degree 0. It is canonically quasi-isomorphism to the Bloch-Suslin complex. Let us define a tensor product

$$\widetilde{m}_{2,2}\colon \widetilde{B}_F(2) \otimes \widetilde{B}_F(2) \to \Gamma(F, 4)_{\mathbb{Q}}$$

as

$$\begin{aligned}
\widetilde{m}_{2,2}(x_1 \wedge x_2 \otimes y_1 \wedge y_2) &= x_1 \wedge x_2 \wedge y_1 \wedge y_2 \in \bigwedge^4 F^*, \\
\widetilde{m}_{2,2}(\{x\} \otimes y_1 \wedge y_2) &= \widetilde{m}_{2,2}(y_1 \wedge y_2 \otimes \{x\}) \\
&= \{x\} \otimes y_1 \wedge y_2 \in B_2(F) \otimes F^*, \\
\widetilde{m}_{2,2}(\{x\} \otimes \{y\}) &= \kappa(x, y) + \phi_2(x) \wedge \phi_2(y), \\
\widetilde{m}_{2,2}((0, \infty, 1, x, y); z) &= S_4(x, y; z), \\
\widetilde{m}_{2,2}((0, \infty, 1, x, y), (0, \infty, 1, z, t)) &= S_4(x, y; r_2(z, t)) \\
&\quad + S_4(z, t; r_2(x, y)).
\end{aligned}$$

It is easy to see that

$$\delta_4(S_4(x, y; r_2(z, t)) + S_4(z, t; r_2(x, y))) = 0.$$

So

$$S_4(x, y; r_2(z, t)) + S_4(z, t; r_2(x, y)) \in \mathscr{R}_4(F)$$

and, according to Theorem 1.15, we would get a functional equation for tetralogarithm. Therefore, the elements $S_4(x, y; z)$ would provide us with a tensor product of Bloch-Suslin complexes and a functional equation for tetralogarithm. The case $n = 2$, $m = 3$ can be considered in complete analogy with the first one.

An explicit construction of the motivic Lie algebra $L(F)_\bullet$ was suggested in **[G5]**.

5. Explicit formulas for the universal Chern class $c_3 \in H^6_?(\mathrm{BGL}_{3\bullet,\mathbb{Q}}(3))$ in motivic and Deligne cohomology

1. The third motivic complex $\Gamma(X; 3)$ for a regular scheme (see subsection 14 of §1 in **[G2]**). Let F be a field with a discrete valuation v and the residue class F_v $(= \overline{F})$. The group of units U has a natural homomorphism $U \to \overline{F}^*$, $u \mapsto \bar{u}$. An element $\pi \in F^*$ is prime if $\mathrm{ord}_v \pi = 1$. Let us construct a canonical homomorphism of complexes

$$\partial_v\colon \Gamma(F, n) \to \Gamma(\bar{F}_v, n-1)[-1] \tag{5.1}$$

such that the induced homomorphism

$$H^n(\Gamma(F, n)) = K_n^M(F) \to H^{n-1}(\Gamma(\overline{F}_v, n-1)) = K_{n-1}^M(\overline{F}_v)$$

coincides with Milnor's tame symbol on $K_n^M(F)$.

There is a homomorphism $\theta\colon \bigwedge^n F^* \to \bigwedge^{n-1} \overline{F}_v^*$ uniquely defined by the following properties $(u_i \in U)$:

(1) $\theta(\pi \wedge u_1 \wedge \cdots \wedge u_{n-1}) = \bar{u}_1 \wedge \cdots \wedge \bar{u}_{n-1}$.
(2) $\theta(u_1 \wedge \cdots \wedge u_n) = 0$.

It clearly does not depend on the choice of π.

Let us define a homomorphism $s_v\colon \mathbb{Z}[P_F^1] \to \mathbb{Z}[P_{\overline{F}_v}^1]$ as

$$s_v\{x\} = \begin{cases} \{\bar{x}\} & \text{if } x \text{ is a unit,} \\ 0 & \text{otherwise.} \end{cases}$$

Then it induces a homomorphism (see subsection 9 of §1 of **[G2]**)

$$s_v\colon \mathscr{B}_k(F) \to \mathscr{B}_k(\overline{F}_v)).$$

Set

$$\partial_v := s_v \otimes \theta : \mathscr{B}_k(F) \otimes \bigwedge^{n-k} F^* \to \mathscr{B}_k(\overline{F}_v) \otimes \bigwedge^{n-k-1} \overline{F}_v^*.$$

Lemma 5.1. *The homomorphism ∂_v commutes with the coboundary δ and hence defines a homomorphism of complexes* (5.1).

Proof. See subsection 14 of §1 in **[G2]**. □

Now let X be an arbitrary regular scheme, $X^{(i)}$ the set of all codimension i points of X, and $F(x)$ the field of functions corresponding to a point $x \in X^{(i)}$. We define the third motivic complex $\Gamma(X; 3)$ as the total complex associated with the following bicomplex:

$$\begin{array}{ccccccc}
\bigwedge^3 F(x)^* & \xrightarrow{\partial_1} & \coprod_{x\in X^{(1)}} \bigwedge^2 F(x)^* & \xrightarrow{\partial_2} & \coprod_{x\in X^{(2)}} F(x)^* & \xrightarrow{\partial_3} & \coprod_{x\in X^{(3)}} \mathbb{Z} \\
\uparrow \delta & & \uparrow \delta & & & & \\
\mathscr{B}_2(F(X)) \otimes F(X)^* & \xrightarrow{\partial_1} & \coprod_{x\in X^{(1)}} \mathscr{B}_2(F(X)) & & & & \\
\uparrow \delta & & & & & & \\
\mathscr{B}_3(F(X)) & & & & \partial = \bigoplus \partial v_x & &
\end{array} \tag{5.2}$$

where $\mathscr{B}_3(F(X))$ placed in degree 1 and coboundaries have degree $+1$.

The coboundaries ∂_i are defined as follows: $\partial_1 := \coprod_{x\in X^{(1)}} \partial_{v_x}$. The others are a little bit more complicated. Let $x \in X^{(k)}$ and $v_1(y), \dots, v_m(y)$ be all discrete valuations of the field $F(x)$ over a point $y \in X^{(k+1)}$, $y \in \overline{x}$. Then $\overline{F(x)}_i := \overline{F(x)}_{v_i(y)} \supset F(y)$. (If $\overline{x}$ is nonsingular at the point y, then

$\overline{F(x)_i} = F(y)$ and $m = 1$.) Let us define a homomorphism $\partial_2 \colon \bigwedge^2 F(x) \to F(y)^*$ as the composition

$$\bigwedge^2 F(x)^* \xrightarrow{\bigoplus \partial_{v_i(y)}} \bigoplus_{i=1}^{m} \overline{F(x)_i}^* \xrightarrow{\bigoplus N_{\overline{F(x)_i}/F(y)}} F(y)^*$$

and $\partial_3 \colon F(x)^* \to \coprod_{y \in X^{(3)}} \mathbb{Z}$ as the composition

$$F(x)^* \xrightarrow{\bigoplus \partial_{v_i}} \bigoplus_{i=1}^{m} \mathbb{Z} \xrightarrow{\sum} \mathbb{Z}.$$

2. Explicit formula for the motivic Chern class $c_3 \in H^6_{\mathscr{M}}(\mathrm{BGL}_3(F)_\bullet, \mathbb{Z}(3))$. Set $G^n := \underbrace{G \times \cdots \times G}_{n \text{ times}}$. Recall that

$$BG_\bullet := \mathrm{pt} \overset{s_0}{\underset{s_1}{\leftleftarrows}} G \overset{s_1}{\underset{s_2}{\Lleftarrow}} G^2 \overset{s_0}{\underset{s_3}{\Lleftarrow}} G^3 \overset{s_0}{\underset{s_4}{\Lleftarrow}} \cdots$$

is the simplicial scheme representing the classifying space of the group G. There is a canonical G-bundle over $BG_\bullet$ (G acts on the left on $EG_\bullet$).

(5.3)
$$\begin{array}{ccccccccc} EG_\bullet & & G & \leftleftarrows & G^2 & \Lleftarrow & G^3 & \cdots & G^4 \cdots \\ {\scriptstyle G}\big\downarrow{\scriptstyle \pi} & & \big\downarrow & & \big\downarrow & & \big\downarrow & & \big\downarrow \\ BG_\bullet & & \mathrm{pt} & \leftleftarrows & G^1 & \Lleftarrow & G^2 & \cdots & G^3 \cdots . \end{array}$$

The cochain we have to construct lives in the following bicomplex (we shall show a part on Diagram 5.4 and the remaining one on (5.5)):
(5.4)

$$\begin{array}{ccccc} \vdots & & & & \\ \cdots & & & & \\ \uparrow{\scriptstyle \delta} & & & & \\ \bigwedge^3 F(G^3)^* \oplus (\coprod) & \xrightarrow{s^*} & \bigwedge^3 F(G^4)^* \oplus \coprod_{x \in (G^4)^{(1)}} \mathscr{B}_2(F(x)) & & \\ & & \uparrow{\scriptstyle \delta} & & \\ & & \mathscr{B}_2(F(G^4)) \otimes F(G^4)^* & \xrightarrow{s^*} & \mathscr{B}_2(F(G^5)) \otimes F(G^5)^* \\ & & & & \uparrow{\scriptstyle \delta} \\ & & & & \mathscr{B}_3(F(G^5)) \xrightarrow{s^*} \mathscr{B}_3(F(G^6)). \end{array}$$

Here $s^* := \sum (-1)^i s_i^*$ and $(\coprod) := \coprod_{x \in (G^3)^{(1)}} \mathscr{B}_2(F(x))$.

Let v be a nonzero vector in a three-dimensional vector space V^3 over F. Put for a generic 5-tuple $(g_1, \dots, g_5)$ of elements of the group $\mathrm{GL}(3)$

(see subsection 3 of §3)

$$m_0(g_1, \dots, g_5) := r_3(v, g_1 v, \dots, g_5 v) \in \mathscr{B}_3(F(G^5)),$$
$$m_1(g_1, \dots, g_4) := -f_5(3)(v, g_1 v, \dots, g_4 v) \in \mathscr{B}_2(F(G^4)) \otimes F(G^4)^*,$$
$$m_2(g_1, g_2, g_3) := f_4(3)(v, g_1 v, g_2 v, g_3 v) \in \bigwedge^3 F(G^3)^*.$$

THEOREM 5.2. (a) $s^* m_0 = 0$.
(b) $s^* m_1 + \delta m_0 = 0$.
(c) $s^* m_2 + \delta m_1 = 0$.

PROOF. (a) follows from the definition of $B_3(F)$ and existence of the homomorphism $B_3(F) \to \mathscr{B}_3(F)$.

(b) is equivalent to Theorem 3.10.

(c) follows from Proposition 3.7 and the following simple but important remark: $\Delta(l_1, l_2, l_3)$ appears in formula (3.7) with factor $\{r(l_4 \mid l_1, l_2, l_3, l_5)\}_2 - \{r(l_5 \mid l_1, l_2, l_3, l_4)\}_2$ that is zero if $\Delta(l_1, l_2, l_3) = 0$. (This implies that the $\mathscr{B}_2(F(x))$-component of δm_1 is zero for any $x \in (G^4)^{(1)}$.) □

We see that this part of the construction of the cocycle c_3 is essentially equivalent to a construction of a homomorphism of complexes (3.2). The remaining part of bicomplex (5.4) looks as follows:

$$\begin{array}{ccccc}
\coprod_{x \in G^{(3)}} \mathbb{Z} & & & & \\
\uparrow{\scriptstyle \partial_3} & & & & \\
\coprod_{x \in G^{(2)}} F(x)^* & \xrightarrow{s^*} & \coprod_{x \in (G^2)^{(2)}} F(x)^* & & \\
& & \uparrow{\scriptstyle \partial_2} & & \\
& & \coprod_{x \in (G^2)^{(1)}} F(x)^* & \xrightarrow{s^*} & \coprod_{x \in (G^3)^{(1)}} \bigwedge^2 F(x)^* \\
& & & & \uparrow{\scriptstyle \partial_1} \\
& & & & \bigwedge^3 F(G^3)^* \oplus (\coprod) \vdots
\end{array} \tag{5.5}$$

Let us describe the corresponding components of the cocycle c_3. Put

$$\mathscr{D}_{v,1} = \{(g_1, g_2) \in G \times G \mid \Delta(v, g_1 v, g_2 v) = 0\}.$$

For generic $(g_1, g_2) \in \mathscr{D}_{v,1}$ we have $\dim\langle v, g_1 v, g_2 v\rangle = 2$; so we can set

$$m_3(g_1, g_2) := -6(\Delta_2(v, g_1 v) \wedge \Delta_2(v, g_2 v) - \Delta_2(g_1 v, v) \wedge \Delta_2(g_1 v, g_2 v) + \Delta_2(g_2 v, v) \wedge \Delta_2(g_2 v, g_1 v)) \in \bigwedge^2 F(\mathscr{D}_{v,1})^*.$$

(Δ_2 is defined using a volume form in $\langle v, g_1 v, g_2 v\rangle$.)

LEMMA 5.3. $s^* m_3 + \partial_1 m_2 = 0$.

PROOF. This is equivalent to the following: $\Delta(l_0, l_1, l_2)$ appears in the formula

$$f_4(3)(l_0, l_1, l_2, l_3) = \operatorname{Alt}\Delta(l_0, l_1, l_2) \wedge \Delta(l_0, l_1, l_3) \wedge \Delta(l_0, l_2, l_3)$$

with factor

$$\begin{aligned}3f_3(2)(l_3 \mid l_0, l_1, l_2) := 6(&\Delta(l_0, l_1, l_3) \wedge \Delta(l_0, l_2, l_3)\\ &- \Delta(l_1, l_0, l_3) \wedge \Delta(l_1, l_2, l_3)\\ &+ \Delta(l_2, l_0, l_3) \wedge \Delta(l_2, l_1, l_3)). \quad \square\end{aligned}$$

Set $\mathscr{D}_{v,2} = \{g \in G \mid gv = \lambda v \text{ for some } \lambda \in F^*\}$. We have a canonical invertible function $\lambda(g) := gv/v$ on $\mathscr{D}_{v,2}$. Put $m_4(g) := 6 \cdot \lambda(g)$.

LEMMA 5.4. $s^* m_4 + \partial_2 m_3 = 0$; $\partial_3 m_4 = 0$.

PROOF. In complete analogy with the previous one. $\square$

So we have constructed the cocycle $(m_0(g_1, \dots, g_5), \dots, m_4(g))$ representing a class $c_3 \in H^6_{\mathscr{M}}(\mathrm{BGL}_3(F)_\bullet, \mathbb{Z}(3))$. In the next section for any complex algebraic manifold X a regulator

$$r_3 : H^\bullet_{\mathscr{M}}(X, \mathbb{Z}(3)) \to H^\bullet_{\mathscr{D}}(X, \mathbb{R}(3))$$

will be constructed. We shall apply it to c_3.

3. Explicit construction of the regulator r_3. Recall that a (real-valued) p-current on X is by definition a linear continuous functional on the space of $(\dim_R X - p)$-forms with compact support. Let us denote by $\mathscr{A}^p_X$ the space of all p-currents on X. There is a differential $d : \mathscr{A}^p_X \to \mathscr{A}^{p+1}_X$, and the de Rham complex $(A^\bullet_X, d)$ is a resolution of the constant sheaf $\mathbb{R}$.

The third Deligne complex $\underline{\widetilde{\mathbb{R}}}(3)_X$ can be defined as a total complex associated with the following bicomplex (see **[B3]**):

$$\begin{array}{ccccccccccc}\mathscr{A}^0_X & \xrightarrow{d} & \mathscr{A}^1_X & \xrightarrow{d} & \mathscr{A}^2_X & \xrightarrow{d} & \mathscr{A}^3_X & \xrightarrow{d} & \mathscr{A}^4_X & \xrightarrow{d} & \cdots \\ & & & & & & \uparrow{\scriptstyle \mathrm{Re}} & & \uparrow{\scriptstyle -\mathrm{Re}} & & \\ & & & & & & \Omega^3_X & \xrightarrow{\partial} & \Omega^4_X & \xrightarrow{\partial} & \cdots .\end{array}$$

Here $\mathscr{A}^0_X$ placed in degree 1 and $(\Omega^\bullet_X, \partial)$ is the de Rham complex of holomorphic forms with logarithmic singularities at infinity.

The Deligne complex $\underline{\widetilde{\mathbb{R}}}(n)_X$ is defined as follows:

$$\underline{\widetilde{\mathbb{R}}}(n)_X := \operatorname{Cone}(\Omega^{\geq n}_X \xrightarrow{\alpha_n} \mathscr{A}^\bullet_X)[-1]$$

where $\alpha_n = (-1)^{n-1}\mathrm{Re}$ for odd n and $(-1)^n \mathrm{Im}$ for even.

To compute $H^*(X, \underline{\widetilde{\mathbb{R}}}(n)_X)$ we shall use the Dolbeaux resolution $(\mathscr{D}^{\geq p, q}_X)$ for the complex of sheaves $(\Omega^{\geq n}_X, \partial)$ where $\mathscr{D}^{p,q}_X$ is the space of complex-valued (p, q)-currents.

EXAMPLE 5.5. $dz/z \in \mathscr{A}_{\mathbb{C}}^{1,0}$ and $\bar\partial(dz/z) = 2\pi i\delta(0)\, dz\, \overline{dz}$. So $\bar\partial d\log f = 2\pi i\delta(f)\, df\, \overline{df}$ for $f \in \mathscr{O}_X$.

EXAMPLE 5.6. $d\arg f \in \mathscr{A}_X^1$ and $d(d\arg f) = 2\pi i\delta(f)\, df\, d\overline{f}$ for $f \in \mathscr{O}_X$.

In order to produce the regulator r_3 we shall construct maps (that are not homomorphisms of complexes; see, however, Proposition 5.7)

$$s_n\colon \Gamma(\operatorname{Spec}\mathbb{C}(X);\, n) \to \widetilde{\mathbb{R}}(n)_X \qquad (n \le 3).$$

Namely, the map $s_3(\cdot)$

$$(5.6)\qquad \begin{array}{ccccccc} \mathscr{B}_3(\mathbb{C}(X)) & \xrightarrow{\delta} & \mathscr{B}_2(\mathbb{C}(X)) \otimes \mathbb{C}(X)^* & \xrightarrow{\delta} & \bigwedge^3 \mathbb{C}(X)^* & \longrightarrow & 0 \\ \big\downarrow {\scriptstyle s_3(1)} & & \big\downarrow {\scriptstyle s_3(2)} & & \big\downarrow {\scriptstyle s_3(3)\oplus d\log\wedge d\log\wedge d\log} & & \\ \mathscr{A}_X^0 & \xrightarrow{d} & \mathscr{A}_X^1 & \xrightarrow{(d,0)} & \mathscr{A}_X^2 \oplus \Omega_X^3 & \longrightarrow & \cdots. \end{array}$$

is defined as follows:

$$\begin{aligned} s_3(1)&\colon \{f(x)\}_3 \mapsto \mathscr{L}_3(f(x)),\\ s_3(2)&\colon \{f(x)\}_2 \otimes g(x) \mapsto -\mathscr{L}_2(f(x))d\arg g(x)\\ &\qquad + \tfrac{1}{3}\log|g(x)|\\ &\qquad \cdot (\log|1-f(x)|d\log|f(x)| - \log|f(x)|d\log|1-f(x)|),\\ s_3(3)&\colon f_1\wedge f_2\wedge f_3 \mapsto \operatorname{Alt}(\tfrac{1}{2}\cdot\log|f_1|d\arg f_2 \wedge d\arg f_3\\ &\qquad\qquad - \tfrac{1}{6}\cdot|f_1|d\log|f_2|d\log|f_3|) \in \mathscr{A}_X^2;\end{aligned}$$

$$d\log\wedge d\log\wedge d\log\colon f_1\wedge f_2\wedge f_3 \mapsto d\log f_1\wedge d\log f_2\wedge d\log f_3 \in \Omega_X^3[A].$$

PROPOSITION 5.7. *Then maps $s_3(\cdot)$ define a homomorphism of complexes*

$$(5.7)\qquad \begin{array}{ccccc} \mathscr{B}_3(\mathbb{C}(X)) & \longrightarrow & \mathscr{B}_2(\mathbb{C}(X)) \otimes \mathbb{C}(X)^* & \longrightarrow & \bigwedge^3 \mathbb{C}(X)^* \\ \big\downarrow {\scriptstyle s_3(1)} & & \big\downarrow {\scriptstyle s_3(2)} & & \big\downarrow {\scriptstyle s_3(3)} \\ S^0_{\eta(X)} & \xrightarrow{d} & S^1_{\eta(X)} & \xrightarrow{d} & S^2_{\eta(X)} \end{array}$$

where $S^p_{\eta(X)}$ is the space of p-forms at the generic point $\eta(X)$ of X.

PROOF. Direct calculation using (1.14). □

This proposition means that $s_3(\cdot)$ is a homomorphism of complexes modulo currents supported on subvarieties of nonzero codimension of X.

The map

$$(5.8)\qquad \begin{array}{ccc} \mathscr{B}_2(\mathbb{C}(X)) & \xrightarrow{\delta} & \bigwedge^2 \mathbb{C}(X)^* \\ \big\downarrow {\scriptstyle s_2(1)} & & \big\downarrow {\scriptstyle s_2(2)\oplus d\log\wedge d\log} \\ \mathscr{A}_X^0 & \xrightarrow{(d,0)} & \mathscr{A}_X^1 \oplus \Omega_X^2 \end{array}$$

is defined as follows:

$$s_2(1)\colon \{f(x)\}_2 \mapsto \mathscr{L}_2(f(x))$$
$$s_2(2)\colon f\wedge g \mapsto -\log|f|d\arg g + \log|g|d\arg f \in \mathscr{A}^1_X.$$

Finally, $s_1\colon f(x)\mapsto [\log|f(x)|, -df/f] \in \mathscr{A}^0_X \oplus \Omega^1_X$.

If $i\colon Y \hookrightarrow X$ is a complex algebraic subvariety of codimension d then there is a canonical homomorphism of complexes

$$i_*\colon \widetilde{\mathbb{R}}(m)_Y \to \widetilde{\mathbb{R}}(m+d)_X[2d]$$

provided by natural maps $i_*\colon \mathscr{D}^{p,q}_Y \hookrightarrow \mathscr{D}^{p+d,q+d}_X$. Therefore, there is a collection of maps

$$i_*\circ s_{n-d}\colon \coprod_{x\in X^{(d)}} \Gamma(\operatorname{Spec}\mathbb{C}(X), n-d) \to \widetilde{\mathbb{R}}(n)_X[2d]. \tag{5.9}$$

Recall that by definition $\Gamma(X, 3)$ is the total complex associated with the bicomplex

$$\begin{aligned}\Gamma(\operatorname{Spec}\mathbb{C}(X), 3) &\overset{\partial_1}{\to} \coprod_{x\in X^{(1)}} \Gamma(\operatorname{Spec}\mathbb{C}(x), 2)[-1] \overset{\partial_2}{\to} \coprod_{x\in X^{(2)}} \mathbb{C}(x)^*[-2]\\ &\overset{\partial_3}{\to} \coprod_{x\in X^{(3)}} \mathbb{Z}[-3].\end{aligned} \tag{5.10}$$

So applying (5.9) to this complex we get the desired map

$$r_3\colon \Gamma(X, 3) \to \widetilde{\mathbb{R}}(3)_X. \tag{5.11}$$

THEOREM 5.8. (5.11) *is a homomorphism of complexes.*

PROOF. Follows immediately from the construction and Proposition 5.7 together with an analogous claim for s_2 and Examples (5.5), (5.6). □

REMARK 5.9. We can define regulators $r_n\colon \Gamma(X, n) \to \widetilde{R}(n)_X$ in complete analogy with this definition of r_3. The only thing that we need is an explicit formula for $s_n(\cdot)$. See **[G3]** for details and formulas.

4. Formula for a cocycle representing $c_3 \in H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet, R(3))$. Let v be a nonzero vector in a 3-dimensional vector space V^3, $G = \mathrm{GL}(V^3)$. The cocycle $c_3^{(v)}$ we have to construct will depend on v and look as follows:

$$
(5.12)\qquad
\begin{array}{ccccccccccc}
(f^4_{(v)}, w^{3,2})_{(v)} & \xrightarrow{s^*} & & & & & & & & & \\
 & \uparrow{\scriptstyle\delta} & & & & & & & & & \\
 & (f^3_{(v)}, w^{3,1}_{(v)}) & \xrightarrow{s^*} & & & & & & & & \\
 & & \uparrow{\scriptstyle\delta} & & & & & & & & \\
 & & (f^2_{(v)}, w^{3,0}_{(v)}) & \xrightarrow{s^*} & & & & & & & \\
 & & & \uparrow{\scriptstyle d} & & & & & & & \\
 & & & f^1_{(v)} & \xrightarrow{s^*} & & & & & & \\
 & & & & \uparrow{\scriptstyle d} & & & & & & \\
 & & & & f^0_{(v)} & \xrightarrow{s^*} & 0 & & & &
\end{array}
$$

$$\mathrm{pt} \leftleftarrows G^1 \leftleftarrows G^2 \leftleftarrows G^3 \leftleftarrows G^4 \leftleftarrows G^5 \leftleftarrows \cdots .$$

Set (see (5.4), (5.6))

$$
\begin{aligned}
f^0_{(v)}(g_1, \dots, g_5) &:= \mathscr{L}_3(m^0_{(v)}(g_1, \dots, g_5)) := \mathscr{L}_3(r_3(v, g_1 v, \dots, g_5 v)),\\
f^1_{(v)}(g_1, \dots, g_4) &:= s_3(2)(m^1_{(v)}(g_1, \dots, g_4),\\
f^2_{(v)}(g_1, g_2, g_3) &:= s_3(3)(m^2_{(v)}(g_1, g_2, g_3)),\\
w^{(3,0)}_{(v)}(g_1, g_2, g_3) &:= d\log \wedge d\log \wedge d\log(m^2_{(v)}(g_1, g_2, g_3)),\\
f^3_{(v)}(g_1, g_2) &:= i_{1*s_2}(2)(m^3_{(v)}(g_1, g_2)),\\
w^{3,1}_{(v)}(g_1, g_2) &:= i_{1*} d\log \wedge d\log(m^3_{(v)}(g_1, g_2)),\\
f^4_{(v)}(g) &:= i_{2*}(-\log|m^4_{(v)}(g)|),\\
w^{3,2}_{(v)}(g) &:= i_{2*} d\log(m^4_{(v)}(g)).
\end{aligned}
$$

Here $i_1 : \mathscr{D}_{v,1} \hookrightarrow G \times G$, $i_2 : \mathscr{D}_{v,2} \hookrightarrow G$ and

$$
\begin{aligned}
i_{1*} &: \Omega^2_{\mathscr{D}_{v,1}} \hookrightarrow \mathscr{D}^{3,1}_{G\times G}, \quad \mathscr{A}^1_{\mathscr{D}_{v,1}} \hookrightarrow \mathscr{D}^3_{G\times G};\\
i_{2*} &: \Omega^1_{\mathscr{D}_{v,2}} \hookrightarrow \mathscr{D}^{3,2}_{G}, \quad \mathscr{A}^0_{\mathscr{D}_{v,2}} \hookrightarrow \mathscr{D}^4_{G}.
\end{aligned}
$$

Theorem 5.10. (a) *$c_3^{(v)}$ is a cocycle.*
(b) *It represents a nontrivial nondecomposable class in* $H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet, R(3))$.
(c) *This class lies in*

$$\mathrm{Im}(H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet, \mathbb{Q}(3)) \to H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet, R(3))).$$

PROOF. (a) follows from Theorem 5.2, Lemmas 5.3 and 5.4 and Theorem 5.8.

(b) Let π: $\mathrm{EG}_\bullet \to \mathrm{BG}_\bullet$ be the universal G-bundle realized as in (5.3). Then $EG_{(p)} = BG_{(p+1)}$ and so any i-cochain $c_{(\bullet)}$ for $BG_\bullet$ defines an $(i-1)$-cochain $\tilde{c}_{(\bullet)}$ for $EG_\bullet : \tilde{c}_{(p)} := c_{(p+1)}$. Moreover, if $c_{(0)} = 0$ and $c_{(\bullet)}$ is a cocycle then $d\tilde{c}_{(\bullet)} = c_{(\bullet)}$. Therefore, $c_{(1)} = \tilde{c}|_G$ is the transgression of the cocycle $c_{(\bullet)}$.

Applying this to the cocycle $c_3^{(v)}$ constructed above we get a cocycle

$$(f^4_{(v)}, w^{3,2}_{(v)}) \in H^5_{\mathscr{D}}(\mathrm{GL}_3(\mathbb{C}), R(3)).$$

The usual exact sequence for Deligne cohomology gives us

$$\cdots \to H^5_{\mathscr{D}}(\mathrm{GL}_3(\mathbb{C}), \mathbb{R}(3)) \xrightarrow{\alpha} H^5(\mathrm{GL}_3(\mathbb{C}), R(3)) \cap H^5(\mathrm{GL}_3(\mathbb{C}), \Omega^{\geq 3}) \to \cdots .$$

By definition

$$\alpha(f^4_{(v)}, w^{3,2}_{(v)}) = w^{3,2}_{(v)} \in \mathscr{D}^{3,2}(\mathrm{GL}_3(\mathbb{C})).$$

Now let us check that the cohomology class defined by $w^{3,2}_{(v)}$ is nonzero and belongs to

$$H^5(\mathrm{GL}_3(\mathbb{C}), \mathbb{Q}(3)) \cap H^5(\mathrm{GL}_3(\mathbb{C}), \Omega^{\geq 3}).$$

There is the standard map

$$p_v\colon \mathrm{GL}_3(\mathbb{C}) \to \mathbb{C}^3\backslash 0, \qquad g \to gv.$$

A vector v defines a canonical invertible function on the line $(\mathbb{C}^* \cdot v) \in \mathbb{C}^3\backslash 0$, namely, $\widetilde{m}_4(\lambda \cdot v) := \lambda$. So there is a $(3, 2)$-current $\widetilde{w}^{3,2}_{(v)}$ represented by the 1-form $d \log \widetilde{m}_4$ on the line generated by v. By definition

$$w^{3,2}_{(v)} = p_v^* \widetilde{w}^{3,2}_{(v)}.$$

On the other hand, let u be the canonical class in $H^5_B(\mathrm{GL}_3(\mathbb{C}), \mathbb{Z})$. (It is dual to the homology class defined by the 5-sphere in $\mathbb{C}^3\backslash 0$.

It is easy to see that $\widetilde{w}^{3,2}_{(v)}$ represents the class $(2\pi i)^3 u$.

So the class defined by $w^{3,2}_{(v)}$ coincides with the canonical generator in $H^5_B(\mathrm{GL}_3(\mathbb{C}), \mathbb{Z}(3))$. Theorem 5.10 is proved.

THEOREM 5.11. *The 5-cocycle $\mathscr{L}_3(r_3(v, g_1 v, \ldots, g_5 v))$ defines a class in $H^5_{\mathrm{cts}}(\mathrm{GL}_3(\mathbb{C}), R(2))$ that coincides with a nonzero rational multiple of the Borel class.*

PROOF. Let G^δ be the Lie group made discrete. The morphism of groups $\mathrm{GL}_3(\mathbb{C})^\delta \to \mathrm{GL}_3(\mathbb{C})$ provides a morphism

$$e\colon \mathrm{BGL}_3(\mathbb{C})^\delta_\bullet \to \mathrm{BGL}_3(\mathbb{C})_\bullet.$$

Therefore,

$$\begin{aligned} e^*\colon H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet\,, R(3)) &\to H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})^\delta\,, R(3)) \\ &= H^5(\mathrm{BGL}_3(\mathbb{C})_\bullet\,, S^0) \equiv H^5_{\mathrm{cts}}(\mathrm{GL}_3(\mathbb{C})\,, R(2)) \end{aligned}$$

(S^0 is the sheaf of C^∞-functions). It is known that e^* maps the indecomposable class in $H^6_{\mathscr{D}}(\mathrm{BGL}_3(\mathbb{C})_\bullet\,, \mathbb{Z}(3))$ just to a nonzero rational of the Borel class in $H^5_{\mathrm{cts}}(\mathrm{GL}_3(\mathbb{C})\,, R(2))$. (This is a particular case of Beilinson's theorem comparing his regulator with the Borel one; see **[B3]** or **[RSS]**.) In our case $e^*(c_3^{(v)}) = \mathscr{L}_3(r_3(v\,, g_1 v\,, \dots\,, g_5 v))$ by construction. □

5. Possible generalizations. Recall that $(T_*(n)\,, \partial)$ is the total complex associated with the Grassmannian bicomplex (3.18) and $T_{n+1}(n) = C_{n+1}(n)$.

OPTIMISTIC CONJECTURE 5.12. *There exists a homomorphism of complexes* $\psi_*(n)$:

$$\begin{array}{ccccccccc} \xrightarrow{\partial} & T_{2n}(n) & \xrightarrow{\partial} & \cdots & \longrightarrow & T_{n+2}(n) & \xrightarrow{\partial} & T_{n+1}(n) \\ & \downarrow{\scriptstyle \psi_{2n}(n)} & & & & \downarrow{\scriptstyle \psi_{n+2}(n)} & & \downarrow{\scriptstyle \psi_{n+1}(n)} \\ 0 \longrightarrow & \mathscr{B}_n(F) & \xrightarrow{\delta} & \cdots & \xrightarrow{\delta} & \mathscr{B}_2(F)\otimes\bigwedge^{n-2}F^* & \xrightarrow{\delta} & \bigwedge^n F^* \end{array}$$

such that

$$\psi_{n+1}(n)\colon (l_0\,, \dots\,, l_n) \in C_{n+1}(n) \mapsto \mathrm{Alt}\bigwedge\nolimits_{i=1}^n \Delta(l_0\,, \dots\,, \hat{l}_i\,, \dots\,, l_n) \in \bigwedge\nolimits^n F^*.$$

This conjecture together with formulas for $\psi_*(n)$ would imply all explicit formulas for characteristic classes that I can imagine. Let me illustrate this by the following examples.

COROLLARY 5.13. *Conjecture* 5.12 *implies a construction of the Chern classes*

$$C_{i,n}\colon K^{[n-i]}_{2n-i}(F)_{\mathbb{Q}} \to H^i(\Gamma_F(n)_{\mathbb{Q}}).$$

(*I use the rank filtration instead of the Adams one.*)

PROOF. See subsections 7 and 10 in §3. □

COROLLARY 5.14. *Zagier's conjecture about* $\zeta_F(n)$ *follows from Conjecture* 5.12.

PROOF. The function $P_n := \widetilde{\mathscr{L}}_n \circ \psi_{2n}(n)$ on $C_{2n}(n)$

$$P_n\colon (l_0\,, \dots\,, l_{2n-1}) \xrightarrow{\psi_{2n}(n)} \mathscr{B}_n(\mathbb{C}) \xrightarrow{\widetilde{\mathscr{L}}_n} R$$

satisfies the functional equations

$$\sum_{i=0}^{2n}(-1)^i P_n(l_0\,, \dots\,, \hat{l}_i\,, \dots\,, l_{2n}) = 0 \quad \forall (l_0\,, \dots\,, l_{2n}) \in C_{2n+1}(n)\,,$$

$$\sum_{i=0}^{2n}(-1)^i P_n(l_i \,|\, l_0\,, \dots\,, \hat{l}_i\,, \dots\,, l_{2n}) = 0 \quad \forall (l_0\,, \dots\,, l_{2n}) \in C_{2n+1}(n+1).$$

Therefore for a nonzero vector $v \in \mathbb{C}^n$ the function $P_n(v, g_1 v, \ \ , g_{2n} v)$ is a measurable $(2n-1)$-cocycle of $\mathrm{GL}_n(\mathbb{C})$. The formula for a homomorphism $\psi_{n+1}(n)$ together with Conjecture 5.12 would imply that the cohomology class of this cocycle is a nonzero rational multiple of the Borel class. In fact, using the main result of **[DHZ]**, one can see that it coincides with the Borel class. This, together with Borel's Theorem 2.3, would then prove Zagier's conjecture. For $n = 3$ this was explained in subsections 7 and 10 in §3 and in §5. See **[G4]** for the general case and a construction of a cocycle representing the Borel class in $H^{2n-1}_{\mathrm{cts}}(\mathrm{GL}_N(\mathbb{C}), R(n-1))$ for $N > n$. □

Formulas for $\psi_*(n)$ provide an explicit construction of the universal Chern class $c_n \in H^{2n}_{\mathcal{M}}(\mathrm{BGL}_N(F)_\bullet, \mathbb{Q}(n))$ $(N \geq n)$ together with their realization in Deligne cohomology. In particular, we shall get an explicit construction of the Chern classes of vector bundles with values in motivic cohomology (see **[G4]**). I would like to emphasize that all this is closely related to the work of Gabrielov, Gelfand, and Losik about combinatorial formulas for the first Pontryagin class **[GGL, You]**.

The Grassmannian complex $(C_*(n), d)$ is a subcomplex in $(T_*(n), \partial)$. Therefore, homomorphism $\psi_*(n)$ provides a formula for the Grassmannian n-cocycle in Deligne cohomology conjectured in **[BMS, HM]**.

It is interesting that for applications (to characteristic classes for instance) it is *not* sufficient to have such formulas for the Grassmannian complex only: we have to extend them to the whole Grassmannian *bicomplex.* This problem becomes nontrivial already for $n = 4$.

Another important application of formulas for $\psi_*(n)$ is a very explicit construction using the classical polylogarithms for Beilinson's regulator for curves and, moreover, arbitrary regular schemes X. Together with Beilinson's conjecture about regulators this will give us a (hypothetical) explicit formula for $\zeta_X(n)$. Note that such formulas can be written without mentioning Conjecture 5.11; see **[G3]**.

I am sure that homomorphisms $\psi_*(n)$ exist for any n. It is canonical for $n < 4$ and in this case formulas for $\psi_*(n)$ were written in **[G2]**. I still hope that there exist nice formulas for all n (that is why I call Conjecture 5.12 "optimistic"). But I have no idea whether there should be canonical in some sense homomorphisms for $n \geq 4$. This is related very much to the discussion in the last subsection of §4. Today I know an explicit formula (for arbitrary n) for $\psi_{n+2}(n)$ and $\psi_{n+1}(n)$ only. (Formulas for $\psi_{n+2}(n)$ can be found in the proof of Proposition 3.3 in **[G4]**.) I think that the computation of $\psi_*(n)$ should be a very interesting problem. Theorem 4.7 indicates that new phenomenas could appear for $n \geq 4$. The case $n = 4$ is crucial for the understanding of Conjecture 5.12 and should be quite different from $n = 2, 3$.

REFERENCES

[B1] A. A. Beilinson, *Height pairings between algebraic cycles*, Lecture Notes in Math., vol. 1289, Springer-Verlag, New York and Berlin, 1987, pp. 1–26.

[B2] ——, *Polylogarithms and cyclotomic elements*, preprint, 1989.

[B3] ——, *Higher regulators and values of L-functions*, VINITI **24** (1984), 181–238; English transl., J. Soviet Math. **30** (1985), 2036–2070.

[BD] A. A. Beilinson and P. Deligne, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs*, these Proceedings, vol. 2, pp. 97–121.

[BMS] A. A. Beilinson, R. MacPherson, and V. V. Schechtman, *Notes on motivic cohomology*, Duke Math. J. **54** (1987), 679–710.

[BGSV] A. A. Beilinson, A. B. Goncharov, V. V. Schechtman, and A. N. Varchenko, *Projective geometry and algebraic K-theory*, Algebra i Analiz **2** (1990), 78–130, Translation in Leningrad Math. J. **2** (1991), 523–576.

[Bl1] S. Bloch, *Higher regulators, algebraic K-theory and zeta functions of elliptic curves*, Lecture Notes, U. C. Irvine, 1977.

[Bl2] ——, *Application of the dilogarithm function in algebraic K-theory and algebraic geometry*, Proc. Internat. Sympos. Alg. Geometry, Kyoto, 1977, pp. 1–14.

[Bo1] A. Borel, *Cohomology des espaces fibrés principaux*, Ann. of Math. (2) **57** (1953), 115–207.

[Bo2] ——, *Cohomologie de* SL_n *et valeurs de fonctions zêta aux points entiers*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **4** (1977), 613–636.

[D1] P. Deligne, *Le groupe fondamental de la droite projective moins trois points*, Galois groups over $\mathbb{Q}$ (Y. Ihara, K. Ribet, and J.-P. Serre, eds.), Publ. Math. Sci. Res. Inst. **16** (1989), 80–290.

[D2] ——, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs*, preprint 1990.

[Du1] J. Dupont, *The dilogarithm as a characteristic class for flat bundles*, Proceedings of the Northwestern conference on cohomology of groups (Evanston, Ill, 1985), J. Pure Appl. Algebra **44** (1987), 137–164.

[DS] J. Dupont and C.-H. Sah, *Scissors congruences*. II, J. Pure Appl. Algebra **25** (1982), 159–196; Characteristic classes of flat bundles, preprint, 1992.

[GGL] A. M. Gabrielov, I. M. Gelfand, and M. V. Losic, *Combinatorial computation of characteristic classes*, Functional Anal. Applications **9** (1975), 103–115.

[GM] I. M. Gelfand and R. MacPherson, *Geometry in Grassmannians and a generalisation of the dilogarithm*, Adv. Math. **44** (1982), 279–312.

[G1] A. B. Goncharov, *The classical trilogarithm, algebraic K-theory of fields, and Dedekind zeta-functions*, Bull. Amer. Math. Soc. (N.S.) **29** (1991), 155–161.

[G2] ——, *Geometry of configurations, polylogarithms and motivic cohomology*, Preprint of the Max-Planck-Institut fur Mathematik, 1991; Adv. Math. (to appear).

[G3] ——, *Explicit formulas for regulators*, in preparation.

[G4] ——, *Explicit construction of characteristic classes*, preprint of the Max-Planck-Institut fur Mathematik, 1992.

[G5] ——, *Hyperlogarithms, mixed Tate motives and multiple* ζ*-numbers*, MSRI, Preprint, 1993, 35 pp.

[Gu] A. Guichardet, *Cohomologie des groupes topologiques et des algèbres de Lie*, Textes Mathématiques **2**, Paris, 1980.

[Gil] H. Gillet, *Riemann-Roch Theorem in higher K-theory*, Adv. Math. **40** (1981), 203–289.

[HZ] D. Hain and S. Zucker, *Unipotent variations of mixed Hodge structures*, Invent. Math. **88** (1987), 83–125.

[H-M] R. Hain and R. MacPherson, *Higher logarithms*, Illinois J. Math. **34**, 392–475.

[K] E. E. Kummer, *Über die Transcendeten, welche aus wiederholten Integrationen rationaler Formeln entstehen*, J. Reine Angew. Math. **21** (1840), 74–90, 193–225, 328–371, (=C.P., vol. 11, N 15).

[L] L. Lewin, *Dilogarithms and Associated Functions*, North-Holland, Amsterdam, 1981.

[Le] ——, *Structural properties of polylogarithms*, Math. Surveys Monographs, vol. **37**, Amer. Math. Soc., Providence, RI, 1991.

[Lei] C. I. Gerhardt, ed., *G. W. Leibniz Mathematische Schriften* III/1, Georg Olms Verlag, Hildesheim, New York, 1971, pp. 336–339.

[L1] S. Lichtenbaum, *Values of zeta functions at non-negative integers*, Journées Arithmétiques, Noordwijkerhovt, Netherlands, Springer-Verlag, New York, 1983.

[L2] ——, *The construction of weight two arithmetic cohomology*, Invent. Math. **88** (1987), 183–215.

[Ma] H. Matsumoto, *Sur les sous-groupes arithmétiques des groupes semisimples deployes*, Ann. Sci. École Norm. Sup. (4) **2** (1969), 1–62.

[Mo] C. Moore, *Group extensions of p-adic and adelic linear groups*, Inst. Hautes Études Sci. Publ. Math. **35** (1969), 5–74.

[M] R. MacPherson, *The combinatorial formula of Gabrielov, Gelfand and Losik for the first Pontriagin class*, Sém. Bourbaki 497, Fev., 1977, Lecture Notes in Math., vol. 677, Springer-Verlag, New York, 1978.

[M2] J. Milnor, *Algebraic K-theory and quadratic forms*, Invent. Math. **9** (1970), 318–340.

[MM] J. Milnor and J. Moore, *On the structure of Hopf algebras*, Ann. of Math. (2) **81** (1965), 211–264.

[MS] A. S. Mercuriev and A. A. Suslin, *On the* K_3 *of a field*, LOMI preprint, Leningrad, 1987.

[NS] Yu. P. Nesterenko and A. A. Suslin, *Homology of the full linear group over a local ring and Milnor K-theory*, Izv. Acad. Sci. USSR **553** (1989), 121–146. (Russian)

[Q1] D. Quillen, *Higher algebraic K-theory*. I, Lecture Notes in Math., vol. 341, Springer-Verlag, Berlin and New York, 1973, pp. 85–197.

[RSS] M. Rapoport, N. Schappacher, and M. Schneider, *Beilinson's conjectures on values of L-functions*, Progr. Math, vol. 4, Birkhäuser, Boston, MA, 1989.

[S] W. Spence, *An essay on logarithmic transcendents*, London and Edinburgh, 1809, pp. 26–34.

[Sa] C.-H. Sah, *Homology of classical groups made discrete*. III, J. Pure Appl. Algebra (1989), 269–312.

[So] C. Soulé, *Opérations en K-théorie algébrique*, Canad. J. Math. **27** (1985), 488–550.

[S1] A. A. Suslin, *Homology of* GL_n, *characteristic classes and Milnor's K-theory*, Proceedings of the Steklov Institute of Mathematics 1985, Issue 3, pp. 207–226; Lecture Notes in Math., vol. 1046, Springer-Verlag, New York, 1989, pp. 357–375 .

[S2] ——, *Algebraic K-theory of fields*, Proceedings of the International Congress of Mathematicians, Berkeley, CA, 1986, pp. 222–243.

[S3] ——, K_3 *of a field and Bloch's group*, Proceedings of the Steklov Institute of Mathematics, 1991, Issue 4.

[E] W. T. van Est, *Group cohomology and Lie algebra cohomology in Lie groups*, Indag. Math. **15** (1953), 484–504.

[Y1] J. Yang, *On the real cohomology of arithmetic groups and the rank conjecture for number fields*.

[Y2] ——, *The Hain-Macpherson's trilogarithm, the Borel regulators and the value of Dedekind zeta function at* 3, in preparation.

[You] B. V. Youssin, *Sur les formes* $S^{p,q}$ *apparaissant dans le calcul combinatoire de la deuxième classe de Pontrjagine par la méthode de Gabrielov, Gelfand, et Losik*, C. R. Acad. Sci. Paris Ser. I Math. **292** (1981), 641–649.

[Z1] D. Zagier, *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields* (G. v. d. Geer, F. Oort, and J. Steenbrink, eds.), Arithmetic Algebraic Geometry, Progr. Math., vol. 89, Birkhäuser, Boston, MA, 1991, pp. 392–430; Proceedings of the Texel conference on Arithmetical Algebraic Geometry.

[Z2] ——, *Hyperbolic manifolds and special values of Dedekind zeta functions*, Invent. Math. **83** (1986), 285–301.

Massachusetts Institute of Technology, Cambridge, MA
E-mail address: `sashagon@math.mit.edu`

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs

A. BEILINSON ET P. DELIGNE

Cet article est une version remaniée d'un preprint du même titre, dont les idées ont motivé certaines des constructions de l'article "motivic polylogarithm and Zagier conjecture", cité **[BD]**, à paraître. Nous espérons qu'il peut encore rester utile comme introduction à **[BD]**. Nous remercions D. Zagier de conversations où se sont dégagées les idées essentielles.

Au §1, après avoir énoncé la conjecture de D. Zagier, nous donnons un formalisme motivique qui l'implique. La preuve est donnée au §2. Son principe est expliqué plus concrètement au §3. Enfin, au §4, nous montrons que la conjecture de D. Zagier est la partie réelle d'une conjecture complexe, elle aussi impliquée par le formalisme motivique.

1. Enoncés

1.1. Rappelons quelques propriétés des fonctions polylogarithme, définies pour $|z| < 1$ par

$$\mathrm{Li}_k(z) = \sum_1^\infty \frac{z^n}{n^k} \qquad (k \geqslant 1). \tag{1.1.1}$$

Ce sont des intégrales itérées de formes différentielles à pôles logarithmiques à l'infini sur $\mathbb{P}^1(\mathbb{C}) - \{0, 1, \infty\}$:

$$\begin{aligned} \mathrm{Li}_1(z) &= -\log(1-z) = \int_0^z \frac{dt}{1-t}, \\ \mathrm{Li}_{k+1}(z) &= \int_0^z \frac{dt}{t}\,\mathrm{Li}_k(t) \qquad (k \geqslant 1). \end{aligned} \tag{1.1.2}$$

Cette expression comme intégrale itérée montre que $\mathrm{Li}_k(z)$, défini par (1.1.1) pour $|z| < 1$, se prolonge comme une fonction multivaluée sur $\mathbb{P}^1 - \{0, 1, \infty\}$.

1991 *Mathematics Subject Classification.* Primary 19F27; Secondary 33E20, 14A20.
This paper is in final form and no version of it will be submitted for publication elsewhere.

Fixons $N \geqslant k$. Les propriétés de monodromie de Li_k **[BD,** (0.7.1)] s'expriment au mieux en considérant cette fonction comme un coefficient de la matrice $L(z)$ suivante, de lignes et colonnes indexées par l'ensemble $[0, N]$ des entier de 0 à N:

$$L(z) := \begin{pmatrix} 1 & & & & \\ -\mathrm{Li}_1(z) & 1 & & \bigcirc & \\ -\mathrm{Li}_2(z) & \log z & 1 & & \\ \vdots & \frac{(\log z)^2}{2!} & \log z & 1 & \\ & \vdots & & \ddots & \ddots \end{pmatrix}. \tag{1.1.3}$$

Posons

(1.1.4)

$$e_0 := \begin{pmatrix} 0 & & & & & \\ 0 & 0 & & & \bigcirc & \\ & 1 & 0 & & & \\ & & 1 & \ddots & & \\ & & & 1 & 0 & \\ & \bigcirc & & & 1 & 0 \end{pmatrix}, \qquad e_1 := \begin{pmatrix} 0 & & & \bigcirc \\ 1 & 0 & & \\ & & \ddots & \\ \bigcirc & & & 0 \end{pmatrix}.$$

L'algèbre de Lie engendrée par e_0 et e_1 a pour base e_0 et les $(\mathrm{ad}\, e_0)^\ell(e_1)$ (1 en position $(\ell+1, 0)$, 0 ailleurs) $(0 \leqslant \ell < N)$. Les $(\mathrm{ad}\, e_0)^\ell(e_1)$ commutent entre eux et $\mathrm{ad}\,(e_0)^N(e_1) = 0$. L'algèbre de Lie engendrée par e_0 et e_1 est définie sur $\mathbb{Q}$ et correspond à un sous-groupe unipotent définie sur $\mathbb{Q}$ V de GL_{N+1}. La matrice $L(z)$ fait partie du sous-groupe $V(\mathbb{C})$ de $\mathrm{GL}_{N+1}(\mathbb{C}) = \mathrm{GL}(\mathbb{C}^{[0,N]})$. Le mineur $[1, N] \times [1, N]$ de $L(z)$ est $\exp(\log z \cdot e_0)$ et

$$L(z) = \exp\left(\sum -\mathrm{Li}_{\ell+1}(z)(\mathrm{ad} e_0)^\ell(e_1)\right) \cdot \exp(\log z \cdot e_0) \tag{1.1.5}$$

Soit $\tau(\lambda)$ la matrice diagonale de diagonale $(1, \lambda, \lambda^2, \ldots, \lambda^N)$. Elle normalise V. Nous poserons

$$A(z) := L(z)\tau(2\pi i).$$

Une détermination de $A(z)$ ou $L(z)$ dépend du choix d'un chemin γ, dans $\mathbb{P}^1(\mathbb{C}) - \{0, 1, \infty\}$, d'un point de $]0, 1[$ à z. Si ce chemin est préfixé par un tour positif γ_0 (resp. γ_1) autour de 0 (resp. 1), la détermination $A_\gamma(z)$ est changée en

$$A_{\gamma\gamma_0}(z) = A_\gamma(z)\exp(e_0), \tag{1.1.6}$$

$$A_{\gamma\gamma_1}(z) = A_\gamma(z)\exp(e_1). \tag{1.1.7}$$

En conséquence, deux déterminations de $A(z)$ diffèrent par multiplication à droite par un élément de $V(\mathbb{Q}) \subset \mathrm{GL}_{N+1}(\mathbb{Q})$ et le $\mathbb{Q}$-sous-espace vectoriel $A(z)\mathbb{Q}^{[0,N]} \subset \mathbb{C}^{[0,N]}$ ne dépend pas de la détermination choisie de $A(z)$. Il ne dépend pas non plus du choix de i: changer i en $-i$ remplace $\tau(2\pi i)$ par $\tau(-2\pi i) = \tau(2\pi i)\tau(-1)$ et $\tau(-1) \in \mathrm{GL}_{N+1}(\mathbb{Q})$.

La matrice $L(z)$ vérifie

$$dL(z) = \left(\frac{dz}{z}e_0 + \frac{dz}{z-1}e_1\right) L(z). \tag{1.1.8}$$

1.2. Appelons $\mathbb{Q}$- (resp. $\mathbb{R}$-) *variation de Tate mixte* sur une variété analytique S une $\mathbb{Q}$- (resp. $\mathbb{R}$-) variation de structures de Hodge mixtes sur S, de seuls nombres de Hodge non nuls les h^{pp} ($p \in \mathbb{Z}$), et de gradué Gr^W constant sur chaque composante connexe de S. Une $\mathbb{Q}$ (resp. $\mathbb{R}$) variation de Tate mixte sur S est simplement la donnée de (a), (b), (c):

(a) Un système local $H_\mathbb{Q}$ (resp. $H_\mathbb{R}$) de $\mathbb{Q}$ (resp. $\mathbb{R}$) espaces vectoriels de dimension finie;

(b) Une filtration finie croissante W de $H_\mathbb{Q}$ (resp. $H_\mathbb{R}$) par des sous-systèmes locaux, indexée par les entiers pairs, avec $\mathrm{Gr}^W(H_\mathbb{Q})$ (resp. $\mathrm{Gr}^W H_\mathbb{R}$) constant sur chaque composante connexe de S. On note encore W la filtration du système local complexifié $H_\mathbb{C}$ qui s'en déduit.

(c) Chaque fibre $(H_\mathbb{C})_s$ est munie d'une filtration finie décroissante F, opposée à W en ce sens que

$$(H_\mathbb{C})_s = \bigoplus (H_\mathbb{C})_{s,k} \quad \text{avec} \tag{1.2.1}$$
$$W_{-2k} = \bigoplus_{\ell \geqslant k} (H_\mathbb{C})_{s,\ell},$$
$$F^{-k} = \bigoplus_{\ell \leqslant k} (H_\mathbb{C})_{s,\ell}.$$

Cette filtration dépend holomorphiquement de s et vérifie l'axiome de transversalité: si $x(s) \in (H_\mathbb{C})_s$ dépend différentiablement de s, et est pour tout s dans F^k, sa dérivée par rapport à un champ de vecteurs est dans F^{k-1}.

Soit $H_\mathscr{O}$ le fibré vectoriel holomorphe défini par $H_\mathbb{C}$. Si on identifie $H_\mathscr{O}$ (resp. $H_\mathbb{C}$) au faisceau de ses sections holomorphes (resp. localement constantes), on a

$$H_\mathscr{O} = H_\mathbb{C} \otimes_\mathbb{C} \mathscr{O}$$

et $H_\mathbb{C}$ est le faisceau des sections horizontales de $H_\mathscr{O}$ pour une connexion intégrable holomorphe ∇. On peut regarder W comme une filtration horizontale de $H_\mathscr{O}$, F comme une filtration holomorphe, et on a une décomposition

$$H_\mathscr{O} = \bigoplus_k (H_\mathscr{O})_k \tag{1.2.2}$$

comme en (c). L'axiome de transversalité équivaut à

$$\nabla H_{\mathscr{O}k} \subset \Omega^1 \otimes (H_{\mathscr{O},k} \oplus H_{\mathscr{O},k+1}). \tag{1.2.3}$$

Pour S réduit à un point, on parlera de structure (de Hodge mixte) de Tate mixte. On a alors $H_{\mathbb{C}} = \bigoplus_k (H_{\mathbb{C}})_k$.

1.3. Les propriétés (1.1.6) à (1.1.8) permettent d'attacher à la matrice A la $\mathbb{Q}$-variation de Tate mixte suivante $\mathscr{M}^{(N)}$ sur $\mathbb{P}^1 - \{0, 1, \infty\}$:

- fibré vectoriel holomorphe sous-jacent: $\mathscr{M}_{\mathscr{O}}^{(N)} := \mathscr{O}^{[0,N]}$;
- filtrations par le poids et de Hodge: dans la décomposition (1.2.2) $(\mathscr{M}_{\mathscr{O}}^{(N)})_k$ est le facteur d'indice k de $\mathscr{O}^{[0,N]}$:

$$W_{-2k} = \mathscr{O}^{[k,N]} \qquad (k \geqslant 0),$$
$$F^{-k} = \mathscr{O}^{[0,k]} \qquad (k \leqslant N);$$

- réseau rationnel: $\mathscr{M}_{\mathbb{Q}}^{(N)} := A(z)\mathbb{Q}^{[0,N]}$.

D'après 1.1.8, la connection pour laquelle le réseau rationnel est horizontal est

$$\nabla = \frac{d}{dz} - \left(\frac{dz}{z}e_0 + \frac{dz}{z-1}e_1\right).$$

L'axiome de transversalité (1.2.3) exprime que e_0 et e_1 envoient $\mathbb{C}^{[0,k]}$ dans $\mathbb{C}^{[0,k+1]}$.

La définition de $\mathscr{M}^{(N)}$ ne dépend pas du choix de $i \in \mathbb{C}$: elle garde un sens sur $\mathbb{P}^1(C) - \{0, 1, \infty\}$ pour C une clôture algébrique de $\mathbb{R}$, et est fonctorielle en C.

Les $\mathscr{M}^{(N)}$ forment un système projectif en N. Quand N est fixé, nous nous permettrons d'écrire $\mathscr{M}$ pour $\mathscr{M}^{(N)}$.

Notre philosophie est qu'il y a intérêt à exprimer les propriétés de la fonction Li_k en terme de la variation $\mathscr{M}^{(k)}$.

1.4. La variation $\mathscr{M}^{(N)}$ est la composante en théorie de Hodge d'un système de réalisations au sens de **[D]** H, qui admet une description géométrique en terme du groupoïde fondamental de la droite projective moins trois points.

Fixons un nombre premier ℓ et décrivons la composante ℓ-adique. Soit $X = \mathbb{P}^1 - \{0, 1, \infty\} = \mathbb{G}_m - \{1\}$. Pour tout point base b, soit $\pi_1(X, b)_\ell$ le plus grand pro-ℓ quotient de $\pi_1(X, b)$. C'est la fibre en b d'un pro-système local sur la droite projective, privée de $0, 1, \infty$, au-dessus de Spec $(\mathbb{Z}[1/\ell])$. Soit $\pi_1(X, b)_\ell^{(N)}$ le quotient de $\pi_1(X, b)_\ell$ par le sous-groupe d'indice $N+1$ de sa série centrale descendante. C'est un groupe de Lie ℓ-adique, d'algèbre de Lie le quotient d'une algèbre de Lie libre à 2 générateurs par la sous-algèbre d'indice $N+1$ de sa série centrale descendante.

Pour $N \geq 1$, $\pi_1(X, b)_\ell^{(N)}$ s'envoie sur $\pi_1(\mathbb{G}_m, b)_\ell = \mathbb{Z}_\ell(1)$. Soit $K^{(N)}(b)$ le noyau et soit $V(b)^{(N)}$ le plus grand quotient de $\pi_1(X, b)_\ell^{(N)}$ dans lequel l'image de $K^{(N)}(b)$ soit abélienne. C'est une extension

$$1 \to K^{(N)}(b)^{\mathrm{ab}} \to V(b)^{(N)} \to \mathbb{Z}_\ell(1) \to 1.$$

Soit $t(z)$ le vecteur tangent non nul z en 0. Prenons le comme point base **[D,** §15]. Dans $\mathbb{G}_m$, il existe une classe d'homotopie de chemins canonique du vecteur tangent $t(z)$ au point z de même coordonnée.

Pour $z \neq 1$, le $\pi_1(X, t(z))_\ell$-torseur des (classes d'homotopie de) chemins de $t(z)$ à z se trivialise donc en poussant à $\pi_1(\mathbb{G}_m, t(z))_\ell$, et fournit un torseur sous

$$\operatorname{Ker}(\pi_1(X, t(z))_\ell \to \pi_1(\mathbb{G}_m, t(z))_\ell):$$

le torseur des (classes de) chemins de projection sur $\mathbb{G}_m$ le chemin canonique. Le noyau ci-dessus s'envoie dans $K^{(N)}(t(z))$, d'où enfin un torseur $P^{(N)}$ sous $K^{(N)}(t(z))^{\mathrm{ab}}$. A ce torseur correspond une extension $E^{(N)}$ de $\mathbb{Z}_\ell$ par $K^{(N)}(t(z))^{\mathrm{ab}}$.

Parce que $t(z)$ est un vecteur tangent en 0, on dispose de la monodromie autour de 0:

$$\varphi_0 : \mathbb{Z}_\ell(1) \to \pi_1(X, t(z))_\ell .$$

Le $\mathbb{Z}_\ell(1)$ à la source est à voir comme le $\pi_{1\ell}$ de T_0^*, l'espace tangent épointé en 0. Faisons $z = 1$. Soit t_1 un vecteur tangent non nul en 1. Le chemin canonique, dans $\mathbb{G}_m$, de $t(1)$ à 1 fournit dans X une classe mod Ker $(\pi_1(X, t(1))_\ell \to \mathbb{Z}_\ell(1))$ de chemins de $t(1)$ à t_1, et une classe de conjugasion, sous ce noyau, de monodromie en 1:

$$\varphi_1 : \mathbb{Z}_\ell(1) \to \pi_1(X, t(z))_\ell.$$

Ces morphismes φ_1 sont à valeurs dans Ker $(\pi_1(X, t(1))_\ell \to \mathbb{Z}_\ell(1))$. Si on les compose avec la projection sur $K^{(N)}(t(1))^{\mathrm{ab}}$, ils deviennent tous égaux: on a obtenu

$$\varphi_1 : \mathbb{Z}_\ell(1) \to K^{(N)}(t(1))^{\mathrm{ab}} \subset V(t(1))^{\mathrm{ab}}.$$

Via φ_0, $\mathbb{Z}_\ell(1)$ agit par conjugaison sur $K^{(N)}(t(1))^{\mathrm{ab}}$. On vérifie l'existence et l'unicité d'une décomposition

$$K^{(N)}(t(1))^{\mathrm{ab}} \otimes \mathbb{Q}_\ell \simeq \bigoplus_1^N \mathbb{Q}_\ell(i)$$

telle que (a) φ_1 est l'inclusion du facteur d'indice 1 et que (b) l'action de Lie de $\operatorname{Lie}(\mathbb{Z}_\ell(1)) = \mathbb{Q}_\ell(1)$ est la somme des multiplications $\mathbb{Q}_\ell(1) \otimes \mathbb{Q}_\ell(i) \to \mathbb{Q}_\ell(i+1)$.

Pour $t(z)$ un autre vecteur tangent non nul en $0 : t(z) \in T_0^*$, les chemins dans T_0^* de $t(1)$ à $t(z)$ formant un $\pi_1(T_0^*, t(1)) = \mathbb{Z}_\ell(1)$-torseur. C'est le torseur de Kummer $K(z)$. Le groupe fondamental $\pi_1(X, t(z))_\ell$ est déduit de $\pi_1(X, t(1))_\ell$ par torsion par ce torseur, $\mathbb{Z}_\ell(1)$ agissant sur $\pi_1(X, t(1))_\ell$ par automorphismes intérieurs, via φ_0. Il en résulte que $K^{(N)}(t(z))^{\mathrm{ab}}$ est déduit de $K^{(N)}(t(1))^{\mathrm{ab}}$ par torsion kummerienne, et

$$K^{(N)}(t(z))^{\mathrm{ab}} \oplus \mathbb{Q}_\ell \simeq \left[\bigoplus_1^N \mathbb{Q}_\ell(i)\right]^{K(z)} .$$

Tensorisée avec $\mathbb{Q}_\ell$, l'extension $E^{(N)}$ est donc une extension de $\mathbb{Q}_\ell$ par $[\bigoplus_1^N \mathbb{Q}_\ell(i)]^{K(z)}$. C'est l'analogue ℓ-adique de la variation $\mathscr{M}^{(N)}$. Une fois acquis le formalisme motivique du groupe fondamental [**D,** §13], la construction, présentée ci-dessus dans le cadre ℓ-adique, vaut uniformément dans les diverses théorie de cohomologie.

Que, en théorie de Hodge, on retrouve bien $\mathscr{M}^{(N)}$, peut soit se vérifier par un calcul direct, soit se déduire du résultat d'unicité [**BD**].

1.5. La structure de Hodge mixte réelle sous-jacente à $\mathscr{M}_z$ ne dépend que de l'image de $A(z)$ dans $\mathrm{GL}_{N+1}(\mathbb{C})/\mathrm{GL}_{N+1}(\mathbb{R})$ soit,ce qui revient au même, que de $A(z)\overline{A(z)}^{-1}$. On a

$$\begin{aligned} A(z)\overline{A(z)}^{-1} &= (L(z)\tau(2\pi i))(\tau(-2\pi i)^{-1}\overline{L(z)}^{-1}) \\ &= L(z)\cdot \operatorname{int}\tau(-1)(\overline{L(z)}^{-1})\cdot\tau(-1) \end{aligned}$$

Posons

$$B(z) = L(z) \operatorname{int} \tau(-1)(\overline{L(z)})^{-1}. \tag{1.5.1}$$

Cette matrice appartient à $V \subset \mathrm{GL}_{N+1}(\mathbb{C})$ et vérifie

$$B(z) = \operatorname{int} \tau(-1)(\overline{B}(z))^{-1}. \tag{1.5.2}$$

Elle est caractérisée par cette propriété et le fait que sa racine carrée $B^{1/2}$ dans V vérifie

$$A(z)\tau(2\pi i) = B^{1/2}(z)\tau(2\pi i) \quad \text{dans } \mathrm{GL}_{N+1}(\mathbb{C})/\mathrm{GL}_{N+1}(\mathbb{R}). \tag{1.5.3}$$

Prenant le logarithme de (1.5.2), on trouve que le logarithme de $B(z)$ vérifie

$$\log B(z) = -\mathrm{ad}\tau(-1)(\log \bar{B}(z)) :$$

le coefficient $(\log B(z))_{ij}$ est purement imaginaire pour $i-j$ pair et réel pour $i-j$ impair.

Parce que les $(\mathrm{ad}\, e_0)^\ell(e_1)$ commutent entre eux, on a (Bourbaki, Lie II, §5 $n^0 5$, Proposition 5)

$$\exp\left(\sum \lambda_{\ell+1}(\mathrm{ad}\, e_0)^\ell(e_1) + ae_0\right) = \exp\left(\sum \mu_{\ell+1}(\mathrm{ad}\, e_0)^\ell(e_1)\right)\exp(ae_0)$$

$$\text{avec} \quad \sum \mu_\ell T^\ell = \frac{e^{aT}-1}{aT}\sum \lambda_\ell T^\ell. \tag{1.5.4}$$

Le mineur $[1, N]\times[1, N]$ de $\log B(z)$ est $\log(z\bar{z})\cdot e_0$. Soit $(0, D_1(z), \dots, D_N(z))$ la première colonne de $-\log(B(z)^{1/2}$. Utilisant (1.5.4), on vérifie facilement que

$$D_k(z) = \begin{cases} i\sum b_\ell \frac{\log(z\bar{z})^\ell}{\ell!}\mathscr{I}m(\mathrm{Li}_{k-\ell}(z)) & \text{pour } k \text{ pair}, \\ \sum b_\ell \frac{\log(z\bar{z})^\ell}{\ell!}\mathscr{R}e(\mathrm{Li}_{k-\ell}(z)) & \text{pour } k \text{ impair}, \end{cases}$$

où les b_ℓ sont les nombres de Bernouilli $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \dots$.

1.6. Rappelons que la $\mathbb{Q}$-structure de Hodge de Tate $\mathbb{Q}(k)$ est celle de type $(-k, -k)$, d'espace vectoriel complexe sous-jacent $\mathbb{C}$ et de réseau rationnel $(2\pi i)^k\mathbb{Q} \subset \mathbb{C}$. On a $\mathbb{Q}(k) = \mathbb{Q}(1)^{\otimes k}$ et on note (k) un produit tensoriel avec $\mathbb{Q}(k)$.

Dans la catégorie abélienne des $\mathbb{Q}$-structures de Hodge mixtes, on a pour $k \geqslant 1$

$$\mathrm{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(k)) = \mathbb{C}/(2\pi i)^k\mathbb{Q}.$$

Si E est une extension de $\mathbb{Q}(0)$ par $\mathbb{Q}(1)$, la décomposition (1.2.1) de $E_{\mathbb{C}}$ est

$$E_{\mathbb{C}} = (\mathbb{C} \text{ en degré } 0) \oplus (\mathbb{C} \text{ en degré } 1)$$

et $E_{\mathbb{Q}}$ est de la forme

$$\begin{pmatrix} 1 & 0 \\ a & (2\pi i)^k \end{pmatrix} \mathbb{Q}^2 \, ;$$

à E on attache a modulo $(2\pi i)^k\mathbb{Q}$.

Soit $\mathbb{R}(k)$ la $\mathbb{R}$-structure de Hodge sous-jacente à $\mathbb{Q}(k)$. Dans la catégorie abélienne des $\mathbb{R}$-structures de Hodge mixtes, on a de même pour $k \geqslant 1$

$$\mathrm{Ext}^1(\mathbb{R}(0), \mathbb{R}(k)) = \mathbb{C}/(2\pi i)^k\mathbb{R}.$$

On identifiera le membre de droite à $i^{k-1}\mathbb{R}$ par l'inclusion de $i^{k-1}\mathbb{R}$ dans $\mathbb{C}$.

Pour S un schéma lisse sur $\mathbb{Q}$, notons $(K_n(S) \otimes \mathbb{Q})^{(m)}$ le facteur direct de $K_n(S) \otimes \mathbb{Q}$ sur lequel chaque opération d'Adams ψ_ℓ agit par ℓ^m. On dispose d'une application régulateur à valeur dans un Ext^1 dans la catégorie abélienne des $\mathbb{Q}$-structures de Hodge mixtes sur $S(\mathbb{C})$:

$$\mathrm{reg} : (K_{2k-1}(S) \otimes \mathbb{Q})^{(k)} \to \mathrm{Ext}^1_{S(\mathbb{C})}(\mathbb{Q}(0), \mathbb{Q}(k)).$$

Par exemple, pour $k = 1$, on attache à $f \in \mathcal{O}^*(S)$ l'extension $[f]$ de $\mathbb{Q}(0)$ par $\mathbb{Q}(1)$ donnée par:

- décomposition (1.2.2) du fibré holomorphe sous-jacent: $\mathcal{O}$ en degré $0 \oplus \mathcal{O}$ en degré 1;
- réseau rationnel : $\begin{pmatrix} 1 & 0 \\ \log f & 2\pi i \end{pmatrix} \mathbb{Q}^2$ (indépendant de la détermination choisie localement de $\log f$).

Pour S le spectre d'un corps de nombre F, $S(\mathbb{C})$ est l'ensemble des plongements complexes de F et le régulateur devient la donnée, pour chaque $\sigma \in S(\mathbb{C})$ de

$$\mathrm{reg}^\sigma : K_{2k-1}(F) \to \mathbb{C}/(2\pi i)^k\mathbb{Q}.$$

Ces constructions ne dépendent pas du choix de $i \in \mathbb{C}$ et on a donc $\mathrm{reg}^{\bar\sigma}(x) = \mathrm{reg}^\sigma(x)^-$.

Si on passe aux $\mathbb{R}$-structures de Hodge mixtes, on obtient

$$\mathrm{reg}_{\mathbb{R}} : (K_{2k-1}(S) \otimes \mathbb{Q})^{(k)} \to \mathrm{Ext}^1_{S(\mathbb{C})}(\mathbb{R}(0), \mathbb{R}(k))$$

qui, pour $S = \operatorname{Spec} F$, fournit, pour chaque plongement complexe σ de F,

$$\operatorname{reg}^{\sigma}_{\mathbb{R}} : K_{2k-1}(F) \to \mathbb{C}/(2\pi i)^k\mathbb{R} \xleftarrow{\sim} i^{k-1}\mathbb{R}.$$

Notre convention: $\operatorname{reg}^{\sigma}_{\mathbb{R}}$ à valeurs dans $i^{k-1}\mathbb{R}$, a l'avantage que

$$\operatorname{reg}^{\bar\sigma}_{\mathbb{R}}(x) = \operatorname{reg}_{\mathbb{R}}(x)^{-}.$$

On sait que les $\operatorname{reg}^{\sigma}_{\mathbb{R}}$ induisent un isomorphisme de $K_{2k-1}(F) \otimes \mathbb{R}$ avec le sous-espace de $i^{k-1}\mathbb{R}^{S(\mathbb{C})}$ formé des (x_σ) vérifiant $x_{\bar\sigma} = \bar{x}_\sigma$.

1.7. Nous pouvons maintenant énoncer la conjecture de D. Zagier. Soit F un corps de nombre (= une extension finie de $\mathbb{Q}$). Il y a une conjecture pour chaque entier k, et la $k^{\text{ième}}$ conjecture n'a de sens que si les précédentes sont vraies.

Dans la même induction, on énonce la $k^{\text{ième}}$ conjecture, on définit un $\mathbb{Q}$-espace vectoriel $\mathscr{L}^k$, une application

$$\{\}_k : F - \{0,\ 1\} \to \mathscr{L}^k,$$

un homomorphisme

$$d_k : \mathscr{L}^k \to \bigwedge^2 \left(\bigoplus_1^{k-1} \mathscr{L}^\ell \right)$$

et

$$\varphi_k :\ \operatorname{Ker}(d_k) \hookrightarrow K_{2k-1}(F) \otimes \mathbb{Q}.$$

(A) Pour $k = 1$, $\mathscr{L}^1$ est $F^* \otimes_{\mathbb{Z}} \mathbb{Q}$, on note (x) l'image dans $\mathscr{L}^1$ de $x \in F^*$, $\{\ \}_1$ est

$$\{\}_1 : x \mapsto (1-x),$$

$d_1 = 0$ et φ_1 est l'application identique de $F^* \otimes \mathbb{Q}$.

(B) Pour $k \geqslant 2$, soit $\widetilde{\mathscr{L}}^k$ le $\mathbb{Q}$-espace vectoriel librement engendré par des $\{x\}_k^{\sim}$, $x \in F - \{0,\ 1\}$. Définissons

$$\tilde{d}_k : \widetilde{\mathscr{L}}^k \to \mathscr{L}^{k-1} \otimes \mathscr{L}^1 \to \bigwedge^2 \left(\bigotimes_1^{k-1} \mathscr{L}^\ell \right):$$

$$\{x\}_k^{\sim} \mapsto \{x\}_{k-1} \otimes (x) \to \{x\}_{k-1} \wedge (x).$$

Pour $k = 2$, $\tilde{d}_2$ est $\{x\}_2^{\sim} \to (1-x) \wedge (x)$. Pour $k > 2$, le noyau de $\tilde{d}_k$ est celui de $\{x\}_k \mapsto \{x\}_{k-1} \otimes (x) : \widetilde{\mathscr{L}}^k \to \mathscr{L}^{k-1} \otimes F^*$.

D. Zagier conjecture l'existence d'une application

$$\varphi_k :\ \operatorname{Ker} \left(\tilde{d}_k : \widetilde{\mathscr{L}}^k \to \bigwedge^2 \left(\bigotimes_1^{k-1} \mathscr{L}_\ell \right) \right) \to K_{2n-1}(F) \otimes \mathbb{Q}$$

tel que, pour tout plongement complexe σ de F et tout $x = \sum \lambda_\alpha \{x_\alpha\}_k^{\sim}$ dans le noyau de $\tilde{d}_k$,

$$D_k(\sigma x) := \sum \lambda_\alpha D_k(\sigma x_\alpha)$$

vérifie

(1.7.1) $$D_k(x) = -\operatorname{reg}^{\sigma}_{\mathbb{R}}(\varphi_k(x)).$$

Noter que cette formule est vraie pour $k = 1$, où $D_k(x) = -\mathscr{R}\log(1-x)$. Noter aussi que φ_k est uniquement déterminé par (1.7.1).

Si la conjecture est vraie, on pose

$$\mathscr{L}^k := \widetilde{\mathscr{L}}^k / \operatorname{Ker}(\tilde{\varphi}^k)$$

et on définit $\{\}_k$, d_k, et φ_k par passage au quotient.

Si $\mathscr{L}$ est la somme des $\mathscr{L}_k$, $(\mathscr{L}, d)$ est une co-algèbre de Lie graduée. D. Zagier conjecture de plus que les φ_k sont surjectifs, i.e. que

$$(\operatorname{Ker} d)_k \tilde{\to} K_{2k-1}(F) \otimes \mathbb{Q}.$$

Sur cette conjecture additionnelle, nous n'avons rien à dire.

Remarque 1.8. Inductivement, la condition que $\sum \lambda_\alpha \{x_\alpha\}^{\sim}_k$ soit dans $\operatorname{Ker}(\tilde{d}_k)$ est que

(1.8.1) Quel que soit l'homomorphisme $v : F^* \to \mathbb{Q}$, on a

$$\sum \lambda_\alpha v(x_\alpha)^{k-2}(1-x_\alpha) \wedge (x_\alpha) = 0 \qquad \text{dans } \overset{2}{\wedge} F^* \otimes \mathbb{Q}.$$

Forme polarisée: quels que soient $v_1, \dots, v_{k-2} : F^* \to \mathbb{Q}$, on a

$$\sum \lambda_\alpha v_1(x_\alpha) \cdots v_{k-2}(x_\alpha)(1-x_\alpha) \wedge (x_\alpha) = 0.$$

(1.8.2) Pour $2 \leqslant \ell < k$ et pour chaque plongement complexe σ, quel que soit $v : F^* \to \mathbb{Q}$, on a

$$\sum \lambda_\alpha v(x_\alpha)^{k-l} D_\ell(\sigma x_\alpha) = 0.$$

Forme polarisée: pour $v_1, \dots, v_{k-\ell} : F^* \to \mathbb{Q}$,

$$\sum \lambda_\alpha v_1(x_\alpha) \cdots v_{k-\ell}(x_\alpha) D_\ell(\sigma x_\alpha) = 0.$$

Dans (1.8.2), on peut aussi bien prendre v ou les v_i à valeurs complexes.

1.9. Avant de donner l'interprétation motivique de 1.7, quelques remarques sur le système projectif de variations $\mathscr{M}^{(N)}$ de 1.3.

On a

(1.9.1) $$\mathscr{M}^{(N)} = \mathscr{M}^{(M)} / W_{-2N-2}\mathscr{M}^{(M)} \quad \text{pour } N \leqslant M,$$

(1.9.2) $$\operatorname{Gr}^W_{-2k}\mathscr{M}^{(N)} = \mathbb{Q}(k) \quad \text{pour } k \leqslant N,$$

(1.9.3) $$\mathscr{M}^{(1)} = \text{ extension } [1-z] \text{ de } \mathbb{Q}(0) \text{ par } \mathbb{Q}(1)$$

et $W_{-2}\mathscr{M}^{(N)}$ admet une description simple en terme de l'extension $[z]$. On a

$$W_{-2}\mathscr{M}^{(N)} = \operatorname{Sym}^{N-1}([z])(1),$$

avec des isomorphismes (1.9.2) donnés par

$$\begin{aligned}\mathrm{Gr}^W \operatorname{Sym}^{N-1}([z])(1) &= \operatorname{Sym}^{N-1}\mathrm{Gr}^W([z])(1)\\ &= \operatorname{Sym}^{N-1}\left(Q(0)\bigoplus Q(1)\right)(1)\\ &= \bigoplus_1^N \mathbb{Q}(k) \underset{(N-k)!\text{ sur }\mathbb{Q}(k)}{\longrightarrow} \bigoplus_1^N \mathbb{Q}(k).\end{aligned} \tag{1.9.4}$$

Les morphismes de transition du système projectif des $W_{-2}\mathscr{M}^{(N)}$ sont donnés par la dérivation de degré -1 de $\operatorname{Sym}^*([z])$ qui sur Sym^1 est la projection $[z] \to \mathbb{Q}(0)$: $\operatorname{Sym}^1 \to \operatorname{Sym}^0$.

Dans le langage de **[D]**, dont nous n'aurons pas besoin, la description est peut-être plus limpide: $[z]$ définit un $\mathbb{Q}(1)$-torseur, l'algèbre de Lie $\mathbb{Q}(1)$ agit de sur $\bigoplus Q(k)$ par les $\mathbb{Q}(1)\otimes\mathbb{Q}(k) \to Q(k+1)$ et on tord $\bigoplus \mathbb{Q}(k)$ par le $\mathbb{Q}(1)$-torseur définit par $[z]$.

1.10. Nous nous proposons de montrer que la conjecture de D. Zagier (surjectivité des φ_k exclue) résulte des conjectures suivantes.

(A) Pour S un schéma lisse et connexe sur $\mathbb{Q}$, on dispose d'une catégorie tannakienne sur $\mathbb{Q}$ $\mathscr{T}(S)$. Pour S variable, elles forment un champ: on dispose de foncteurs image inverse, avec des propriétés convenables. On appellera $\mathscr{T}(S)$ la catégorie des *motifs de Tate mixtes* sur S.

(B) On dispose d'un objet de rang un $\mathbb{Q}(1)$ dans $\mathscr{T}(\operatorname{Spec}(\mathbb{Q}))$, avec les propriétés (B1), (B2) ci-dessous.

Notations: Pour tout S, on note encore $\mathbb{Q}(1)$ l'image inverse de $\mathbb{Q}(1)$ dans $\mathscr{T}(S)$. On pose $\mathbb{Q}(k) := \mathbb{Q}(1)^{\otimes k} (k \in \mathbb{Z})$ et on note (k) une tensorisation par $\mathbb{Q}(k)$.

(B1) Etant tannakienne, la catégorie $\mathscr{T}(S)$ est en particulier une catégorie abélienne dont tous les objets sont de longueur finie. On suppose que ses objets simples sont les $\mathbb{Q}(n)$, que les $\mathbb{Q}(n)$ sont deux à deux non isomorphes, et que

$$\text{pour } n \leqslant 0,\ \operatorname{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(n)) = 0$$

(d'où $\operatorname{Ext}^1(\mathbb{Q}(n), \mathbb{Q}(m)) = 0$ pour $n \geqslant m$).

Cette hypothèse équivaut à ce que chaque objet M de $\mathscr{T}(S)$ ait une unique filtration croissante finie W, qu'il est commode d'indexer par les entiers pairs, avec $\mathrm{Gr}^W_{-2k}(M)$ somme de copies de $\mathbb{Q}(k)$, et que tout morphisme $f : M_1 \to M_2$ soit strictement compatible à W: le foncteur $M \to \mathrm{Gr}^W(M)$ est exact.

(B2) Notons $(K_m(S)\otimes\mathbb{Q})^{(n)}$ le sous-espace de $K_m(S)\otimes\mathbb{Q}$ où les opérations d'Adams ψ_ℓ agissent par ℓ^n. Pour $k \geqslant 1$, on suppose que

$$\operatorname{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(k)) = (K_{2k-1}(S)\otimes\mathbb{Q})^{(k)}$$

de façon compatible aux changements de base $S' \to S$. Pour $f \in \mathscr{O}^*(S)$, on notera $[f]$ l'extension correspondante de $\mathbb{Q}(0)$ par $\mathbb{Q}(1)$.

(C) On dispose d'un $\otimes$-foncteur "réalisation", compatible aux changements de base $S' \to S$, de $\mathcal{T}(S)$ dans les $\mathbb{Q}$-variations de structures de Hodge mixtes sur $S(\mathbb{C})$. Plus précisément, on veut une réalisation sur $S(C)$, pour C une clôture algébrique de $\mathbb{R}$, fonctorielle en C. On suppose donné un isomorphisme entre la réalisation de $\mathbb{Q}(1)$ et la structure de Hodge de Tate $\mathbb{Q}(1)$.

(D) On suppose la commutativité du diagramme

$$\begin{array}{ccc} (K_{2k-1}(S)\otimes\mathbb{Q})^{(k)} & \xrightarrow{\text{(B2)}} & \mathrm{Ext}^1_{\mathcal{T}(S)}\quad(\mathbb{Q}(0)\,,\,\mathbb{Q}(k)) \\ & {}_{\text{reg}}\searrow \qquad \swarrow_{\text{real}} & \\ & \mathrm{Ext}^1_{S(\mathbb{C})}(\mathbb{Q}(0)\,,\,\mathbb{Q}(k)). & \end{array}$$

(E) Sur $\mathbb{P}^1-\{0\,,\,1\,,\,\infty\}$ identifions $\mathrm{Gr}^W\ \mathrm{Sym}^{N-1}([z])(1)$ à $\bigoplus_1^N \mathbb{Q}(k)$ par la formule (1.9.4). Les $\mathrm{Sym}^{N-1}([z])(1)$ forment un système projectif en N, comme en 1.9, et on demande l'existence d'un système projectif ${}_0\mathcal{M}^{(N)}$ d'extension de $\mathbb{Q}(0)$ par les $\mathrm{Sym}^{N-1}([z])(1)$, avec ${}_0\mathcal{M}^{(1)}$ l'extension $[1-z]$ de $\mathbb{Q}(0)$ par $\mathbb{Q}(1)$, de réalisation le système projectif d'extensions analogue $\mathcal{M}$ de 1.3, 1.9.

1.11. Nous n'avons pas besoin de toute la force de la conjecture ci-dessus. Le cas où $S=\mathbb{P}^1-\{0\,,\,1\,,\,\infty\}$ ou le spectre d'un corps de nombres nous suffit. Pour prouver la conjecture de Zagier pour un corps de nombres F, il suffirait même de disposer de:

($\mathrm{A_F}$) Une catégorie tannakienne $\mathcal{T}(F)$ sur $\mathbb{Q}$.

($\mathrm{B_F}$) Un objet $\mathbb{Q}(1)$ de $\mathcal{T}(F)$, de rang un, vérifiant (B1) de 1.10 et des isomorphismes

$$\mathrm{Ext}^1(\mathbb{Q}(0)\,,\,\mathbb{Q}(k)) = K_{2k-1}(F)\otimes\mathbb{Q} \qquad \text{pour } k\geqslant 1.$$

($\mathrm{C_F}$) Pour chaque plongement complexe σ de F, un $\otimes$ foncteur exact: "réalisation" de $\mathcal{T}(F)$ dans les $\mathbb{Q}$-structures de Hodge mixtes. On suppose donné un isomorphisme real $(\mathbb{Q}(1)) = \mathbb{Q}(1)$ et vérifiée une compatibilité (D) de 1.10.

($\mathrm{E_F}$) Enfin, pour $x\in\mathbb{P}^1(F)-\{0\,,\,1\,,\,\infty\}=F-\{0\,,\,1\}$ on suppose donné un système projectif ${}_0\mathcal{M}_x^{(N)}$ d'extensions de $\mathbb{Q}(0)$ par $\mathrm{Sym}^{N-1}([x])(1)$, avec ${}_0\mathcal{M}_x^{(1)}=[1-x]$, de réalisation rel σ le système projectif analogue $\mathcal{M}_{\sigma x}^{(N)}$ fibre de $\mathcal{M}^{(N)}$ en σx.

2. Preuve

Dans ce paragraphe, nous prouvons que la conjecture 1.11 implique la conjecture de D. Zagier 1.7 (surjectivité des φ_k exclue).

2.1. Supposons donnés une catégorie tannakienne $\mathcal{T}$ sur $\mathbb{Q}$, et un objet de rang un $\mathbb{Q}(1)$, vérifiant la condition (B1) de 1.10. Dans la terminologie de **[BD]**, c'est une "mixed Tate $\mathbb{Q}$-category". Les $\mathbb{Q}(k) := \mathbb{Q}(1)^{\otimes k}$ $(k \in \mathbb{Z})$ étant de rang un, on a $\mathrm{End}(\mathbb{Q}(k)) = \mathbb{Q}$.

Pour tout objet X d'une catégorie additive $\mathbb{Q}$-linéaire, le produit tensoriel avec un $\mathbb{Q}$-espace vectoriel de dimension finie V est défini par

$$\mathrm{Hom}\,(X \otimes V, Y) = \mathrm{Hom}\,(V, \mathrm{Hom}\,(X, Y)).$$

On a $X \otimes \mathbb{Q}^n = X^n$.

Pour tout M dans $\mathscr{T}$, $\mathrm{Gr}^W_{-2k}(M)$ est somme de copies de $\mathbb{Q}(k)$:

$$\mathrm{Gr}^W_{-2k}(M) = \mathbb{Q}(k) \otimes \omega(M)_k \quad \text{avec} \quad \omega(M)_k = \mathrm{Hom}\,(\mathbb{Q}(k), \mathrm{Gr}^W_{-2k}(M)). \tag{2.1.1}$$

Soit $\omega(M)$ l'espace vectoriel gradué somme des $\omega(M)_k$, $\omega(M)_k$ en degré k. La filtration W est compatible au produit tensoriel, le foncteur $M \mapsto \mathrm{Gr}^W(M)$ est exact et compatible au produit tensoriel et $M \mapsto \omega(M)$ est donc un $\otimes$-functeur exact: c'est un foncteur fibre sur $\mathbb{Q}$ de la catégorie tannakienne $\mathscr{T}$.

Par la théorie générale des catégories tannakiennes, ce foncteur induit une équivalence de $\mathscr{T}$ avec la catégorie des représentations du schéma en groupes

$$G := \underline{\mathrm{Aut}}^{\otimes}(\omega). \tag{2.1.2}$$

Pour M dans $\mathscr{T}$, la graduation de $\omega(M)$ correspond à une action τ du groupe multiplicatif $\mathbb{G}_m$: λ agit par multiplication par λ^k sur $\omega(M)_k$. Cette construction, compatible au produit tensoriel, définit

$$\tau : \mathbb{G}_m \to G.$$

Tout automorphisme du $\otimes$-foncteur ω respecte la filtration de $\omega(M)$ par les

$$\omega(W_{-2k}M) = \bigoplus_{\ell \geqslant k} \omega(M)_\ell .$$

Soit U le schéma en groupe prounipotent des $\otimes$-automorphismes de ω qui induisent l'identité sur le gradué pour cette filtration. Si g dans G agit sur $\omega\mathbb{Q}(1)$ par multiplication par λ, g agit sur $\omega\mathrm{Gr}^W M$ par $\tau(\lambda)$ et $g\tau(\lambda)^{-1}$ est dans U : G est un produit semi-direct

$$G = \mathbb{G}_m \ltimes U.$$

Le schéma en groupes G est limite projective de quotients

$$G_\alpha = \mathbb{G}_m \ltimes U_\alpha$$

avec U_α groupe algébrique unipotent. L'algèbre de Lie $\mathrm{Lie}(U_\alpha)$ est graduée, avec $\mathrm{int}\,\tau(\lambda) = \lambda^k$ sur $\mathrm{Lie}(U_\alpha)_k$. On a $\mathrm{Lie}(U_\alpha)_k = 0$ pour $k \leqslant 0$.

Les représentations de G_α s'identifient aux espaces vectoriels gradués V munis d'une action homogène de $\mathrm{Lie}\,(U_\alpha)$: la graduation définit l'action de $\mathbb{G}_m$, et l'action de $\mathrm{Lie}(U_\alpha)$ est automatiquement nilpotente.

En particulier, une extension E de $\mathbb{Q}$ par $\mathbb{Q}(k)$, telle que la représentation $\omega(E)$ de G se factorise par G_α, s'identifie à une action homogène de $\mathrm{Lie}(U_\alpha)$ sur

$$\mathbb{Q} \text{ en degré } 0 \oplus \mathbb{Q} \text{ en degré } k\,,$$

i.e. à une forme linéaire e sur $\mathrm{Lie}(U_\alpha)_k$, nulle sur $[\mathrm{Lie}\ U_\alpha\,,\ \mathrm{Lie}\ U_\alpha]_k$.

Pour passer à la limite sur α, il est commode de se dualiser et de considérer plutôt $\mathrm{Lie}(U_\alpha)^\vee$, gradué par les $(\mathrm{Lie}\ U_\alpha)^{\vee k} := (\mathrm{Lie}\ U_{\alpha k})^\vee$ et l'opposé du transposé de $[\ \ ]$:

$$d : (\mathrm{Lie}\ U_\alpha)^\vee \to \bigwedge^2 (\mathrm{Lie}\ U_\alpha)^\vee.$$

C'est la différentielle extérieure du complexe de de Rham des formes différentielles invariantes à gauche sur U_α.

Passant à la limite inductive, on obtient une coalgèbre de Lie graduée à degrés > 0

$$(\mathrm{Lie}\ U)^\vee := \varinjlim (\mathrm{Lie}\ U_\alpha)^\vee\,,$$

et une extension de $\mathbb{Q}$ par $\mathbb{Q}(k)$ s'identifie à

$$e \in (\mathrm{Lie}\ U^\vee)^k \text{ tel que } de = 0.$$

Soient M dans $\mathcal{T}$, $x \in \mathrm{Hom}(\mathbb{Q}(i), \mathrm{Gr}_{-2i}(M)) = \omega(M)_i$, $y \in \mathrm{Hom}(\mathrm{Gr}_{-2j}(M), \mathbb{Q}(j)) = \omega(M)_j^\vee$ et $k = j - i$.

Si l'action de G sur $\omega(M)$ se factorise par $G_\alpha = \mathbb{G}_m \ltimes U_\alpha$, le coefficient

$$C_{y,x}(u) = \langle y\,,\ ux\rangle \quad (y \in U_\alpha)$$

et une fonction sur U_α, de différentielle à l'origine la forme linéaire

$$c_{y,x}(u) = \langle y\,,\ ux\rangle \quad (u \in \mathrm{Lie}\ U_\alpha).$$

Cette forme linéaire est dans $(\mathrm{Lie}\ U_{\alpha k})^\vee$ et définit

$$c_{y,x} \in (\mathrm{Lie}\ U^\vee)^k.$$

Notre idée essentielle est que si une combinaison linéaire de tels coefficients est annulée par d, elle fournit une extension de $\mathbb{Q}(0)$ par $\mathbb{Q}(k)$.

Admettons 1.11. Pour la conjecture de Zagier, les coefficients considérés sont ceux définis par les ${}_0\mathcal{M}_z^{(N)}$, l'isomorphisme $\mathrm{Gr}_0^W({}_0\mathcal{M}_z^{(N)}) = \mathbb{Q}(0)$ et l'isomorphisme $\mathrm{Gr}_{-2N}^W({}_0\mathcal{M}_z^{(N)}) = \mathbb{Q}(N)$ (comme en 1.9). L'action de Lie U sur $W_0^{-2}\mathcal{M}_z^{(N)} = \mathrm{Sym}^{N-1}([z])(1)$ se factorise par le quotient $(\mathrm{Lie}\ U)_1$. C'est ce qui permet d'étudier les coefficients considérés dans le formalisme des groupes de Bloch (cf. 3.4). Cela implique aussi que l'action de Lie U sur ${}_0\mathcal{M}_z^{(N)}$ est triviale sur $[\mathrm{Lie}\ U^{\geq 2}\,,\ \mathrm{Lie}\ U^{\geq 2}]$.

2.2. Soient F un corps de nombres, admettons 1.11 et faisons dans 2.1 $\mathcal{T} = \mathcal{T}(\mathrm{Spec}\ F)$. Pour $z \in F^*$, soit $[z]$ dans $\mathcal{T}(F) := \mathcal{T}(\mathrm{Spec}\ (F))$ l'extension de $\mathbb{Q}$ par $\mathbb{Q}(1)$ de classe z $(1.11\ (\mathrm{B_F}))$. On identifie $\mathrm{Gr}^W \mathrm{Sym}^{N-1}([z])(1)$ à $\bigoplus_1^N \mathbb{Q}(k)$ par la formule 1.9.4. Comme en 1.9, les

$\mathrm{Sym}^{N-1}([z])(1)$ forment un système projectif en N. Par l'hypothèse 1.11 (E) on dispose pour chaque $z \in F - \{0, 1\}$ d'un système projectif d'extensions ${}_0\mathscr{M}_z^{(N)}$ de $\mathbb{Q}(0)$ par $\mathrm{Sym}^{N-1}([z])(1)$, avec ${}_0\mathscr{M}_z^{(1)} = [1-z]$, de réalisation dans le plongement complexe σ le système projectif analogue d'extensions $\mathscr{M}_{\sigma z}^{(N)}$.

Soient $k \geqslant 1$ et $N \geqslant k$. Ecrivons simplement $\mathscr{M}_z$ pour ${}_0\mathscr{M}_z^{(N)}$. Soit x l'isomorphisme $\mathbb{Q}(0) \tilde{\to} \mathrm{Gr}_0^W(\mathscr{M}_z)$, y l'isomorphisme $\mathrm{Gr}_{-2k}^W(\mathscr{M}_z) \tilde{\to} \mathbb{Q}(k)$ et posons

$$\{z\}_k := c_{y,x} \in (\mathrm{Lie}\ U^\vee)^k$$

(indépendant de N).

Parce que Lie $U^\vee$ est à degré > 0, on a $d = 0$ sur $(\mathrm{Lie}\ U^\vee)^1$ et

$$(\mathrm{Lie}\ U^\vee)^1 = \mathrm{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(1)) = F^* \otimes \mathbb{Q}.$$

Puisque $\mathscr{M}_z^{(1)} = [1-z]$, on a

$$\{z\}_1 = (1-z).$$

Proposition 2.3. *Pour $k \geqslant 2$, on a*

$$d\{z\}_k = \{z\}_{k-1} \wedge (z) \quad \text{dans } \bigwedge^2 \mathrm{Lie}\ (U)^\vee .$$

Preuve. On a

$$\omega(\mathscr{M}_z) = \bigoplus_0^N \omega(\mathbb{Q}(k)) = \mathbb{Q}^{[0,N]}.$$

L'action de U sur $\omega(W_{-2}(\mathscr{M}_z)) = \omega(\mathrm{Sym}^{N-1})[z])(11)$ est déduite de son action sur $\omega([z])$. Il en résulte que l'action φ de U sur $\omega(\mathscr{M}_z)$ se factorise par $V \subset \mathrm{GL}_{N+1}$ (1.1).

Les coefficients $c_{\ell,0}$ $(1 \leqslant \ell \leqslant N)$ et $c_{2,1}$ forment une base duale de la base $(\mathrm{ad}\, e_0)^{\ell-1}(e_0)$ et e_0 de Lie V. On a donc dans $\overset{2}{\wedge}$ Lie $(U)^\vee$

$$dc_{\ell,0} = c_{\ell-1,0} \wedge c_{2,1}.$$

On a $\varphi^* c_{\ell,0} = \{z\}_\ell$, $\varphi^* c_{2,1} = (z)$ et 2.3 en résulte.

2.4. Soit $\sum \lambda_\alpha \{x_a\}$ une combinaison linéaire formelle, à coefficients rationnels, d'éléments de $F - \{0, 1\}$. Si

pour $k = 2$: $\quad \sum \lambda_\alpha (1 - z_\alpha) \wedge (z_\alpha) = 0$ dans $\bigwedge^2 F^* \otimes \mathbb{Q}$.

pour $k > 2$: $\quad \sum \lambda_a \{1 - z_\alpha\}_{k-1} \otimes (z_\alpha) = 0$ dans $(\mathrm{Lie}\ U^\vee)^{k-1} \otimes F^*$,

il résulte de 2.3 que $\sum \lambda_\alpha \{z_\alpha\}_k$ est dans le noyau de d, donc correspond à une extension de $\mathbb{Q}(0)$ par $\mathbb{Q}(k)$. Par 1.11 $(\mathrm{B_F})$, cette extension correspond à un élément

$$r\left(\sum \lambda_\alpha \{x_\alpha\}_k\right) \in K_{2k-1}(F)^{(k)}.$$

Soit σ un plongement complexe de F. Nous nous proposons de calculer le régulateur $\mathrm{reg}_{\mathbb{R}}^\sigma$ de $r(\sum \lambda_\alpha \{z_\alpha\}_k)$.

2.5. Soit $\mathscr{H}$ la catégorie tannakienne sur $\mathbb{R}$ des $\mathbb{R}$-structures de Hodge mixtes dont les seuls nombres de Hodge non nuls sont les $h^{p,p}$.

Par 1.2, la donnée de H dans $\mathscr{H}$ équivant à celle d'un espace vectoriel complexe gradué H_k :

$$H_{\mathbb{C}} = \bigoplus H_k ,$$
$$W_{-2k} = \bigoplus_{\ell \geqslant k} H_k ,$$
$$F^{-k} = \bigoplus_{\ell \geqslant k} H_k ,$$

plus une structure réelle $H_{\mathbb{R}}$ sur $\bigoplus H_k$ relativement à laquelle la filtration W soit réelle. La structure réelle $H_{\mathbb{R}}$ induit une structure réelle sur les $\mathrm{Gr}^W_{-2k} H_{\mathbb{C}} = H_k$. On dispose donc de deux structures réelles sur $H = \bigoplus H_k$: la structurale, $H_{\mathbb{R}}$, et sa graduée $\mathrm{Gr}^W H_{\mathbb{R}}$. Pour faciliter la comparaison avec les motifs de Tate mixte, il sera utile d'en considérer une troisième:

$$\omega(H) := \tau(1/2\pi i)\mathrm{Gr}^W H_{\mathbb{R}} = \bigoplus i^k H_{k\mathbb{R}} .$$

Le donnée de $H_{\mathbb{R}}$ équivant à celle la structure réelle $\bigoplus H_{k\mathbb{R}}$ de $\bigoplus H_k$, plus celle d'une donnée (a), (b), (c), ou (d) comme ci-dessous. Soit $X \subset \mathrm{GL}(H_{\mathbb{C}})$ le sous-groupe qui respecte W et induit l'identité sur le gradué. Soit $X(\mathbb{R})$ sa structure réelle rel. la structure $\oplus H_{k\mathbb{R}}$.

Donnée (a): une classe $n \in X/X(\mathbb{R})$: $(H_{\mathbb{R}})$ et n se déterminent mutuellement par

$$H_{\mathbb{R}} = n\left(\bigoplus H_{k\mathbb{R}}\right) .$$

Donnée (b): un élément $b \in X$ tel que $b\bar{b} = 1$: b et n se déterminent mutuellement par

$$b = n\bar{n}^{-1} .$$

En terme de la structure réelle de X définie par $\bigoplus i^k H_{k\mathbb{R}}$, ces formules se récrivent $b = n$ int $\tau(-1)(\bar{n})^{-1}$, et b int $\tau(-1)(\bar{b}) = 1$.

Donnée (c): un élément purement imaginaire N de Lie $X : N = \frac{1}{2}\log b$.

Donnée (d): des éléments purement imaginaire de degré k $N_k \in \mathrm{GL}(\bigoplus H_{\ell\mathbb{C}})$, $(k \geqslant 1)$:

$$N = \sum N_k .$$

En terme de la structure réelle $\omega(H) = \bigoplus i^k H_{k\mathbb{R}}$, N_k est réel pour k impair, purement imaginaire pour k pair.

La construction $H \mapsto (\omega(H)_{\mathbb{R}}, (N_k))$ est une $\otimes$-équivalence de catégories de $\mathscr{H}$ avec la catégorie des espaces vectoriels réels gradués, munis de $N_k \in \mathrm{GL}(\omega(H)_{\mathbb{R}}) \otimes \mathbb{C}(k \geqslant 1)$, avec N_k de degré k, réel pour k impair et purement imaginaire pour k pair. Pour $\otimes$, l'action de N_k est à traiter comme une action de Lie.

2.6. Fixons un plongement complexe σ de F, donc une foncteur "réalisation" de $\mathscr{T}(\operatorname{Spec} F)$ dans les $\mathbb{Q}$-structures de Hodge mixtes. Le foncteur réalisation est automatiquement à valeurs dans les structures de Hodge mixtes dont les seuls nombres de Hodge non nuls sont les $h^{p,p}$. Si on passe à la $\mathbb{R}$-structure de Hodge mixte sous-jacente, on obtient $\text{real}_{\mathbb{R}}$, à valeur dans la catégorie $\mathscr{H}$ de 2.5. Les définitions ont été ajustées pour que

$$\omega(M) \otimes \mathbb{R} = \omega\big(\text{real}_{\mathbb{R}}(M)\big)$$

et chaque N_k fournit un élément de l'algèbre de Lie de U sur $\mathbb{C}$ (réel pour k impair, imaginaire pour k pair).

Si M est une extension de $\mathbb{Q}(0)$ par $\mathbb{Q}(k)$, d'où $\omega(M) = \mathbb{Q}^{\{0,k\}}$, le coefficient $c_{k,0}$ de N_k est par hypothèse le régulateur de $\text{reg}^{\sigma}_{\mathbb{R}}$ l'élément correspondant de $K_{2k-1}(F)^{(k)} \otimes \mathbb{Q}$.

Soit M dans $\mathscr{T}(\operatorname{Spec} F)$. Sa réalisation $\text{real}(M)$ a pour espace vectoriel complexe sous-jacent $\omega(M) \otimes \mathbb{C}$. Si $X \subset \mathrm{GL}\big(\omega(M)\big)$ est le sous-groupe qui respecte W et agit trivialement sur le gradué, le réseau rationnel de $\text{real}(M)$ est de la forme

$$n\tau(2\pi i)\omega(M)$$

avec $n \in X(\mathbb{C})$. Soient $x \in \omega(M)_0$, y une forme linéaire sur $\omega(M)_k$ et $c_{x,y}$ le coefficient correspondant. Par définition, on a

$$\langle N_k\,,\; C_{x,y}\rangle = \text{ coefficient }_{x,y} \text{ de } \tfrac{1}{2}\log\big(n \text{ int } \tau(-1)(\bar{n})^{-1}\big).$$

Pour $M = \mathscr{M}_z$ $(z \in F - \{0\,,\,1\})$, cela donne:

Proposition 2.7. *On a*

$$\langle N_k\,,\,\{z\}_k\rangle = -D_k(z)\,.$$

Proposition 2.8. *Sous les hypothèses de* 1.11, *il existe un monomorphisme de coalgèbre de Lie* $\varphi : \mathscr{L} \hookrightarrow (\text{Lie } U)^{\vee}$, *compatible à* $\{\ \}_k$. *La conjecture de Zagier vaut pour* φ_k *le composé*

$$\operatorname{Ker}\left(\tilde{d}_k : \widetilde{\mathscr{L}}_k \to \bigwedge^2 \bigoplus_1^{k-1} \mathscr{L}^{\ell}\right) \xrightarrow{\varphi} \operatorname{Ker}(d :\ \text{Lie } U^{\vee} \to \bigwedge^2 \text{Lie } U^{\vee})^k$$
$$= \mathit{Ext}^1\big(\mathbb{Q}(0)\,,\ \mathbb{Q}(k)\big) = K_{2k-1}(F)\,.$$

Définissons inductivement

$$\varphi_k : \mathscr{L}^k \hookrightarrow (\text{Lie } U^{\vee})^k$$

compatible à $\{\ \}_k$ et d. Pour $k = 1$, $\mathscr{L}^1 = F^* \otimes \mathbb{Q}$, $(\text{Lie } U^{\vee})^1 = \text{Ext}^1\big(\mathbb{Q}\,,\ \mathbb{Q}(1)\big) = F^* \otimes \mathbb{Q}$ et on prend pour φ_1 l'application identique de $F^* \otimes \mathbb{Q}$.

Supposons $k > 1$ et φ_ℓ défini pour $\ell < k$. Définissons $\widetilde{\varphi}_k : \widetilde{\mathscr{L}}^k \to \operatorname{Lie}(U^\vee)^k$ par

$$\widetilde{\varphi}_k : (\{x\}_k^\sim) = \{z\}_k .$$

D'après 2.3, le diagramme

$$\begin{array}{ccc} \widetilde{\mathscr{L}}^k & \longrightarrow & (\operatorname{Lie} U^\vee)^k \\ \big\downarrow \widetilde{d}_k & & \big\downarrow d \\ \bigwedge^2 \left(\bigoplus_1^{k-1} \mathscr{L}^\ell \right) & \overset{\varphi}{\hookrightarrow} & \bigwedge^2 \bigoplus_1^{k-1} (\operatorname{Lie} U^\vee)^\ell \end{array}$$

est commutatif. L'application φ envoie donc $\operatorname{Ker}(\widetilde{d}_k)$ dans $\operatorname{Ker}(d)^k = \operatorname{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(k)) = K_{2k-1}(F) \otimes \mathbb{Q}$. Pour une extension E de $\mathbb{Q}(0)$ par $\mathbb{Q}(k)$, définissant $x : \mathbb{Q}(0) \to \operatorname{Gr}_0^W(E)$, i.e. $x \in \omega(E)_0$ et $y : \operatorname{Gr}_{-2k}^W(E) \to \mathbb{Q}(k)$, i.e. $y \in \omega(E)_k^\vee$, $\langle N_k, c_{x,y} \rangle$ est le régulateur $\operatorname{reg}_{\mathbb{R}}^\sigma$ de la classe correspondante dans $K_{2k-1}(F) \otimes \mathbb{Q}$, et que la formule proposée pour φ_k est correcte résulte de 2.7. On définit alors φ_k par passage au quotient à partir de $\widetilde{\varphi}_k$.

Remarque 2.9. Pour que le morphisme φ de 2.8 soit un isomorphisme en degrés $k \leqslant N$, il faut et il suffit que la sous-catégorie tannakienne de $\mathscr{T}(F)$ engendrée (par $\oplus$, $\otimes$, dual, sous-quotients) par les $\mathscr{M}_z^{(N)}$ $(z \in F - \{0, 1\})$ contienne tout les M dans $\mathscr{T}(F)$ tels que $M = W_0 M$ et $W_{-2N-2} M = 0$.

Pour $\mathscr{T}(F)$ la catégorie conjecturale de tous les motifs de Tate mixte sur F, ce devrait n'être vrai que pour $N = 1$ ou 2.

3. Extensions et coefficients

Le but de ce paragraphe est d'expliquer plus concrètement le sens de la construction 2.1 d'extensions comme combinaison linéaire de coefficients.

3.1. Comme en 2.1, on supposé donnée une catégorie tannakienne $\mathscr{T}$ et un objet de rang un $\mathbb{Q}(1)$, vérifiant la condition (B1) de 1.10 d'où, comme en 2.1, un foncteur fibre ω de groupe d'automorphismes $\mathbb{G}_m \ltimes U$.

L'algèbre affine $\Gamma(U, \mathscr{O})$ de U est une bigèbre commutative, de coproduit caractérisé par

$$\Delta f = \sum f_i \otimes g_i \Longleftrightarrow f(uv) = \sum f_i(u) g_i(v) .$$

L'action de $\mathbb{G}_m$ par automorphismes intérieurs fournit une graduation de $\Gamma(U, \mathscr{O})$, avec f de degré k si

$$f(\lambda^{-1} u \lambda) = \lambda^k f(u) .$$

Pour M dans $\mathscr{T}$, $x \in \omega(M)_i$, $y \in \omega(M)_j^\vee$, et $k = j - i$, le coefficient $C_{y,x}$ est de degré k. En particulier, une extension E de $\mathbb{Q}(0)$ par $\mathbb{Q}(k)$

fournit un coefficient $C(E)$ de degré k. Cette construction est une bijection

$$\mathrm{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(k)) \tilde{\to} \{e \in \Gamma(U, \mathscr{O})_k | \Delta e = 1 \otimes e + e \otimes 1\}.$$

Comme en 2.1, ceci fournit un moyen pour construire une extension comme combinaison linéaire de coefficients. Prendre garde toutefois qu'il s'agit ici de coefficients au sens des représentations de groupes, plutôt qu'au sens des représentations d'algèbres de Lie. La relation entre les deux constructions sera donnée en 3.3.

3.2. Soient M, x, y comme en 3.1. Soit M_1 le plus petit sous-objet de M tel que $x \in \omega(M_1)_i$. L'espace vectoriel $\omega(M)$ est une représentation de U et $\omega(M_1)$ est la sous-représentation de $\omega(M)$ engendrée par x. Elle est automatiquement stable par $\mathbb{G}_m$ car x est homogène. Dualement, soit M_2 le plus petit quotient de M tel que $y \in \omega(M_2)_j^\vee \subset \omega(M)_j^\vee$: $\omega(M_2)^\vee$ est la sous-représentation de $\omega(M)^\vee$ engendrée par y. Soit $\bar{M}$ l'image de M_1 dans M_2, $\bar{x}$ et $\bar{y}$ les images de x et y dans $\omega(\bar{M})_i$ et $\omega(\bar{M})_j^\vee$. On a

$$C_{y,x} = C_{\bar{y},\bar{x}}.$$

Pour $i \leqslant k \leqslant j$, soient $e(k)_\alpha$ une base de $\omega(\bar{M})_\ell$ et $e(k)^\alpha$ la base duale de $\omega(\bar{M})_k^\vee$. Les $C_{e(k)^\alpha \bar{x}}$ sont linéairement indépendants, et les $C_{\bar{y}, e(k)_\alpha}$ sont linéairement indépendants. On a

$$\Delta C_{\bar{y},\bar{x}} = \sum C_{\bar{y}e(k)_\alpha} \otimes C_{e(k)^\alpha \bar{x}}.$$

Pour que $C_{\bar{y},\bar{x}}$ vérifie $\Delta C_{\bar{y},\bar{x}} = 1 \otimes C_{\bar{y},\bar{x}} + C_{\bar{y},\bar{x}} \otimes 1$, il faut et il suffit donc que soit $C_{\bar{y},\bar{x}} = 0$, auquel cas $M = 0$, soit que $\bar{M}$ soit une extension de $\mathbb{Q}(i)$ par $\mathbb{Q}(j)$.

Etant donnés des triples M_α, x_α, y_α, avec $x_\alpha \in \omega(M_\alpha)_i$, $y_\alpha \in \omega(M_\alpha)_j^\vee$, si $f = \sum \lambda_\alpha C_{y_\alpha x_\alpha}$ vérifie $\Delta f = 1 \otimes f + f \otimes 1$, l'extension correspondante est donc obtenue comme suit.

(a) Si $f = 0$, il existe dans $\bigoplus M_\alpha$ un sous objet M_1 tel que $\sum \lambda_\alpha x_\alpha \in M_1$ et que $\sum y_\alpha$ soit nul sur $\omega(M_1)_j$.

(b) Sinon, l'extension cherchée est le sous-quotient $(\bigoplus M_\alpha)^-$ de $\bigoplus M_\alpha$, muni de $\sum \lambda_\alpha x_\alpha$ et $\sum y_\alpha$; $\mathrm{Gr}_{-2i}((\bigoplus M_\alpha)^-)$ est identifié à $\mathbb{Q}(i)$ par $(\sum \lambda_\alpha x_\alpha)^-$ et $\mathrm{Gr}_{-2j}((\bigoplus M_\alpha)^-)$ est identifié à $\mathbb{Q}(j)$ par $(\sum y_\alpha)^-$.

Cette construction n'utilise guère le produit tensoriel de $\mathscr{T}$. Il joue par contre un rôle essentiel si on veut, comme en 2.1, utiliser plutôt des coefficients de Lie.

3.3. Pour toute fonction f sur un groupe unipotent U, posons $f^\natural(u) = \langle df, \log u\rangle$. Si f est un coefficient, on peut comme suit écrire $f^\natural$ comme un coefficient. Soit (V, φ) une représentation, et $e_1, \dots, e_n$ une base telle que le drapeau correspondant soit fixe par U. Pour le coefficient $C_{ji} =$

$\langle e^j, ue_i\rangle$ $(i < j)$, on a alors

$$C^\natural_{ji}(u) = \langle dC_{ji}, \log u\rangle = \left(\log\varphi(u)\right)_{ji} = \sum_1^\infty \frac{(-1)^{n+1}}{n}\left(\varphi(u)-1\right)^n_{ji} :$$

$$C^\natural_{ji} = \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n} \sum_{i=k_0<\cdots<k_n=j} C_{k_n k_{n-1}} \cdots C_{k_1 k_0}.$$

De plus, un produit de coefficients s'interprète comme un coefficient d'un produit tensoriel de représentations.

Pour que $C_{y,x}$ vérifie $\Delta f = f \otimes 1 + 1 \otimes f$, il faut et il suffit que $C_{y,x} = C^\natural_{y,x}$ et que $c_{y,x} \in \mathrm{Lie}\,(U)^\vee$ vérifie $dc_{y,x} = 0$ dans le complexe $\wedge\ \mathrm{Lie}\ (U)^\vee$. Si des $(M_\alpha, x_\alpha, y_\alpha)$ sont donnés, et que $\sum \lambda_\alpha c_{y_\alpha, x_\alpha}$ définit comme en 2.1 une extension de $\mathbb{Q}(i)$ par $\mathbb{Q}(j)$, alors $\sum \lambda_\alpha C^\natural_{y_\alpha, x_\alpha}$ définit la même extension, d'où finalement une description de l'extension $\sum \lambda_\alpha c_{y_\alpha x_\alpha}$ de 2.1 comme sous-quotient d'une somme de puissances tensorielles des M_α.

3.4. Dans la situation de 1.11, les coefficients associés (cf. fin de 2.1) aux ${}_0\mathcal{M}_z^{(N)}$ sont des fonctions f sur U, de degré N, ayant la propriété suivante.

(3.4.1). Pour u dans le sous-groupe $U^{\geq 2}$ de U d'algèbre de Lie Lie $U^{\geq 2}$, on a

$$f(ug) = f(u) + f(g).$$

Cela résulte de ce que ${}_0\mathcal{M}_z^{(N)}$ est extension de $\mathbb{Q}(0)$ par une représentation de U triviale sur $U^{\geq 2}$. En d'autres termes, f est dans le groupe de Bloch B_N de $\mathcal{T}(F)$ (**[BD]**). Dans **[BD]**, on évite la problème de construire $\mathcal{T}(F)$ en construisant directement, en terme de K-théorie, ce qui devrait être ses groupes de Bloch, et les éléments qui devraient être les coefficients des ${}_0\mathcal{M}_z^{(N)}$.

4. Variante complexe

Au §2, nous avons vu que la conjecture 1.11 implique une expression comme somme de valeurs de D_k pour l'image par

$$\mathrm{reg}^\sigma_{\mathbb{R}} : K_{2k-1}(F) \to \mathbb{C}/(2\pi i)^k\mathbb{R}$$

des éléments de $K_{2k-1}(F)$ dans l'image de φ_k. Ici, nous montrerons que la conjecture 1.10 fournit aussi une formule pour l'image de ces classes par

$$\mathrm{reg}^\sigma : K_{2k-1}(F) \to \mathbb{C}/(2\pi i)^k\mathbb{Q}.$$

4.1. Fixons $z \in \mathbb{C}$. Appelons *détermination généralisée* de $L(z)$ une matrice de la forme $L(z)\tau(2\pi i)v\tau(2\pi i)^{-1}$ avec $v \in V(\mathbb{Q})$, pour $L(z)$ une détermination de la matrice $L(z)$ de 1.1, et *détermination généralisée* de $\left(\log z, \mathrm{Li}_1(z), \ldots, \mathrm{Li}_N(z)\right)$ une suite de nombres $(\ell, L_1, \ldots, L_N)$ telle que la matrice

$$\begin{pmatrix} 1 & & & \\ -L_1 & 1 & & \bigcirc \\ -L_2 & \ell & 1 & \\ \vdots & \frac{\ell^2}{2} & \ell & \ddots \\ & & & \vdots \end{pmatrix} = \exp\left(-\sum L_i(\mathrm{ad}\, e_0)^{i-1}(e_1)\right)\exp(\ell e_0)$$

soit une détermination généralisée de $L(z)$.

Si $(\ell, L_1, \dots, L_N)$ est une détermination généralisée, les autres déterminations généralisées peuvent s'obtenir par une suite d'opérations des types suivants, dans un ordre qu'on peut prescrire:

(a) remplacer ℓ par $\ell + a$, $a \in 2\pi i\mathbb{Q}$ et garder $L_1, \dots, L_N$:

$(\mathrm{b})_k$ $(1 \leqslant k \leqslant N)$ garder $\ell, L_1, \dots, L_{k-1}$ et pour $i \geqslant 0$, remplacer L_{k+i} par $L_{k+i} + a\frac{\ell^i}{i!}$ $(a \in (2\pi i)^k\mathbb{Q})$.

Passons au logarithme. Soit $\Lambda(z) := \log L(z)$ et définissons les $\Lambda_i(z)$ $(1 \leqslant i \leqslant N)$ par

$$\Lambda(z) = \log z e_0 - \sum \Lambda_i(z)(\mathrm{ad}\, e_0)^i(e_1).$$

Une détermination généralisée de $\Lambda(z)$ ou de $(\log z, \Lambda_1(z), \dots, \Lambda_N(z))$ est définie de même, en terme d'une détermination géneralisée de $L(z)$. Utilisant 1.5.4, on vérifie que

$$\Lambda_i(z) = \sum_j b_j \frac{(\log z)^j}{j!} \mathrm{Li}_{i-j}(z).$$

En particulier

$$\Lambda_1(z) = \mathrm{Li}_1(z) = -\log(1-z),$$
$$\Lambda_2(z) = \mathrm{Li}_2(z) + \tfrac{1}{2}\log(z)\log(1-z).$$

A nouveau, une détermination généralisée de $\log z$, $\Lambda_1(z), \dots, \Lambda_{k-1}(z)$ définit celle de $\Lambda_k(z)$ à un multiple rationnel de $(2\pi i)^k$ près.

4.2. Soient F un corps de nombre, et σ un plongement complexe. Dans ce n^0, nous identifierons F à son image dans $\mathbb{C}$ par σ, σ à l'inclusion identique, et nous écrirons reg pour reg^σ. Supposons choisie une détermination généralisée de $\log z$ pour $z \in F^*$, qui soit un homomorphisme de F^* dans $\mathbb{C}$. Par exemple, pour $F \subset \mathbb{R}$, on peut prendre $\log|z|$. Toute autre détermination généralisée qui soit un homomorphisme est de la forme

$$z \to \log z + a(z)$$

pour a un homomorphisme $F^* \to 2\pi i\mathbb{Q}$.

Nous aurons à considérer des familles finies $(z_\alpha)_{\alpha\in A}$ d'éléments de F^*. Une telle famille étant fixée, nous n'aurons à considérer que les $\log(z_\alpha)$ et les $\log(1-z_\alpha)$, mais il sera essentiel les déterminations généralisées choisies proviennent d'un homomorphisme $\log : F^* \to \mathbb{C}$, i.e. que chaque relation multiplicative entre les z_α et $(1-z_\alpha)$ en fournisse une, additive, entre les $\log(z_\alpha)$ et $\log(1-z_\alpha)$.

Une détermination généralisée $\log : F^* \to \mathbb{C}$ comme ci-dessus des $\log(z)$ $(z \in F^*)$ en fournit une des $\mathrm{Li}_1(z) = -\log(1-z)$, d'où une détermination généralisée des $\mathrm{Li}_2(z)$ modulo $(2\pi i)^2\mathbb{Q}$. Fixons $z \in F$, $z \neq 0$, 1 et choisissons une détermination généralisée de $L(z)$ adaptée à celles déjà choisie pour $\log z$ et $\mathrm{Li}_1(z) = -\log(1-z)$. Si on change la détermination choisie de log en log $+a$, et celle de $L(z)$ par

(4.2.1) $$L(z) \mapsto L(z)\exp\left(-a(1-z)e_1 + a(z)e_0\right),$$

$\Lambda(z)$ est remplacé par un polynôme de Lie en $\Lambda(z)$ et $\left(-a(1-z)e_1+a(z)e_0\right)$:

(4.2.2)
$$\Lambda(z) \mapsto \Lambda(z) + \left(-a(1-z)e_1 + a(z)e_0\right) + \tfrac{1}{2}[\Lambda(z), -a(1-z)e_1 + a(z)e_0] + \cdots$$

(loi de Hausdorff). Les termes écrits suffisent au calcul de la nouvelle détermination de Λ_2 :

$$\Lambda_2(z) \mapsto \Lambda_2(z) + \tfrac{1}{2}\left(a(z)\log(1-z) - a(1-z)(\log z)\right).$$

Si une combinaison linéaire formelle $\sum \lambda_\alpha \{z_\alpha\}_2^\sim$ d'éléments z_α de $F - \{0, 1\}$ vérifie dans $\overset{2}{\wedge} F^* \otimes \mathbb{Q}$

(4.2.3) $$\sum \lambda_\alpha (1-z_\alpha) \wedge z_\alpha = 0,$$

la somme $\sum \lambda_\alpha \Lambda_2(z_\alpha)$ est invariante par le changement (4.2.1) de déterminations: elle change par l'addition de l'image de $\frac{1}{2}\sum \lambda_\alpha(1-z_\alpha)\wedge(z_\alpha) = 0$ par l'application

$$\bigwedge^2 (F^*) \otimes \mathbb{Q} \to \mathbb{C} : (z_1)\wedge(z_2) \mapsto a(z_2)\log(1-z_1) - a(z_1)\log(1-z_2).$$

Ceci définit $\sum \lambda_\alpha \Lambda_2(z_\alpha)$ avec une ambiguité dans $(2\pi i)^2\mathbb{Q}$ (changements 4.1 (b_2)).

On verra que la conjecture 1.10 implique que si $\sum \lambda_\alpha (1-z_\alpha)\wedge(z_\alpha) = 0$, alors $\sum \lambda_\alpha \Lambda_2(z_\alpha) \in \mathbb{C}/(2\pi i)^2\mathbb{Q}$ ainsi défini est $\mathrm{reg}\,(\varphi_2 \sum \lambda_\alpha \{z_\alpha\}_2^\sim)$ (notations de 1.8). En particulier elle implique que

(4.2.4)
$$\text{si } \sum \lambda_\alpha \{z_\alpha\}_2 = 0 \text{ dans } \mathscr{L}^2, \text{ i.e. si} \sum \lambda_\alpha (1-z_\alpha)\wedge(z_\alpha) = 0$$
$$\text{et que } \varphi_2 \sum \lambda_\alpha \{z_\alpha\}_2 = 0, \text{ alors } \sum \lambda_\alpha \Lambda_2(z_\alpha) \in (2\pi i)^2\mathbb{Q}.$$

Sans charger la détermination généralisée de log, nous pouvons changer celle de $L(z_\alpha)$ par

(4.2.5) $$L(z_\alpha) \mapsto L(z_\alpha)\exp\left(-a_2(z_\alpha)\ \mathrm{ad}\ e_0(e_1)\right),$$
$$\Lambda_2(z_\alpha) \mapsto \Lambda_2(z_\alpha) + a_2(z_\alpha)$$

pour des $a_2(z_\alpha) \in (2\pi i)^2\mathbb{Q}$. Par (4.2.4), il existe donc une détermination généralisée des $L(z_\alpha)$, adaptée à celle choisie de log, telle que

$$\forall(\lambda_\alpha) \sum \lambda_\alpha\{z_\alpha\}_2 = 0 \Rightarrow \sum \lambda_\alpha \Lambda_2(z_\alpha) = 0. \tag{4.2.6}$$

La condition (4.2.6) est stable par le changement de déterminations (4.2.1), qui change log. Elle est stable par le changement de déterminations (4.2.5) si (et seulement si) les $a_2(z_\alpha)$ vérifient

$$\sum \lambda_\alpha\{z_\alpha\}_2 = 0 \Rightarrow \sum \lambda_\alpha a_2(z_\alpha) = 0. \tag{4.2.7}$$

On verra que si

$$\sum \lambda_\alpha\{z_\alpha\}_2 \otimes (z_\alpha) = 0 \qquad \text{dans } \mathscr{L}^2 \otimes F^*, \tag{4.2.8}$$

et que les déterminations choisies des $L(z_\alpha)$, compatibles à celle de log, vérifient (4.2.6), alors

$$\sum \lambda_\alpha \Lambda_3(z_\alpha) \tag{4.2.9}$$

est invariant par les changements de déterminations (4.2.1) et par les changements (4.2.5) vérifiant (4.2.7). La somme (4.2.9) est ainsi définie avec une ambiguité dans $(2\pi i)^3\mathbb{Q}$. On verra que la conjecture 1.10 implique que cette somme est $\operatorname{reg} \varphi_3 \sum \lambda_\alpha\{z_\alpha\}$, donc que

$$\sum \lambda_\alpha\{z_\alpha\} = 0 \text{ dans } \mathscr{L}^3 \Rightarrow \sum \lambda_\alpha \Lambda_3(z_\alpha) \in (2\pi i)^3\mathbb{Q}.$$

Il existe donc des déterminations vérifiant en outre

$$\sum \lambda_\alpha\{z_\alpha\}_3 = 0 \Rightarrow \sum \lambda_\alpha \Lambda_3(z_\alpha) = 0, \ldots .$$

4.3. Précisons l'argument qui fait passer de (4.2.6) à une ambiguité dans $(2\pi i)^3\mathbb{Q}$ pour (4.2.9), et les ... finaux de 4.2. Soit $k \geqslant 2$. Soit $\sum \lambda_\alpha\{z\}_k^\sim$ une combinaison linéaire formelle d'éléments de $F - \{0, 1\}$. Supposons que quel que soit $v : F^* \to \mathbb{Q}$, on ait

$$\sum \lambda_\alpha v(z_\alpha)^{k-2}((1 - z_\alpha) \wedge (z_\alpha)) \qquad \text{dans } \overset{2}{\wedge} F^* \otimes \mathbb{Q}. \tag*{$(4.3.1)_k$}$$

Forme polarisée: quels que soient $v_1, \ldots, v_{k-2} : F^* \to \mathbb{Q}$,

$$\sum \lambda_\alpha v_1(z_\alpha) \cdots v_{k-2}(z_\alpha)((1 - z_\alpha) \wedge (z_\alpha)) = 0. \tag*{$(4.3.2)_k$}$$

On notera que cette hypothèse implique que quels que soient $2 \leqslant \ell \leqslant k$ et $v_1, \ldots, v_{k-\ell} : F^* \to \mathbb{Q}$,

$$\sum \lambda_\alpha v_1(z_\alpha) \cdots v_{k-\ell}(z_\alpha)\{z_\alpha\}_\ell^\sim \text{ vérifie } (4.3.1)_\ell. \tag{4.3.3}$$

Supposons choisies des déterminationsgénéraliséesdes $L(z_\alpha)$, prolongeant une détermination de log comme en 4.2 et telles que pour $2 \leqslant \ell \leqslant k - 1$ on ait

$(4.3.4)_\ell$ pour toute forme linéaire $v : F^* \to \mathbb{Q}$,

$$\sum (\lambda_\alpha v(z_\alpha)^{k-\ell})\Lambda_\ell(z_\alpha) = 0.$$

Forme polarisée: quels que soient $v_1, \ldots, v_{k-\ell} : F^* \to \mathbb{Q}$

$$\sum \lambda_\alpha v_1(z_\alpha) \cdots v_{k-\ell}(z_\alpha)\Lambda_\ell(z_\alpha) = 0. \tag*{$(4.3.5)_\ell$}$$

Proposition 4.4. *Soient* $\sum \lambda_\alpha \{z_\alpha\}_k^\sim$ *et des détermination généralisées* $L(z_\alpha)$ *prolongeant une de log. Si* $(4.3.2)_k$ *et* $(4.3.4)_\ell$ $(2 \leqslant \ell \leqslant k-1)$ *sont vérifiés, alors les changements de déterminations suivant respectent ces conditions et ne modifient pas la valeur de* $\sum \lambda_\alpha \Lambda_k(z_\alpha)$:

(i) *changement* (4.2.1):

$(\mathrm{ii})_\ell$ $(2 \leqslant \ell \leqslant k-1)$

$$L(z_\alpha) \mapsto L(z_\alpha)\exp\left(a_\ell(z_\alpha)(\operatorname{ad} e_0)^{e-1}(e_1)\right)$$

où $a_\ell(z_\alpha) \in (2\pi i)^\ell \mathbb{Q}$ *et où quels que soient* $v_1, \ldots, v_{k-\ell} : F^* \to \mathbb{Q}$,

$$\sum \lambda_\alpha v_1(z_\alpha) \cdots v_{k-\ell}(z_\alpha) a_\ell(z_\alpha) = 0.$$

Preuve. Prouvons (i). Il s'agit de remplacer $\Lambda(z)$ par un polynome de Lie (4.2.2) en $\Lambda(z)$ et $\left(-a(1-z)e_1 + a(z)e_0\right)$. On a

(4.4.1)

$$[\Lambda(z), -a(1-z)e_1 + a(z)e_0] = -\left(a(z)\log(1-z) - a(1-z)\log z\right) \operatorname{ad} e_0(e_1) + \sum_{\ell=3}^{N} \Lambda_{\ell-1} a(z)(\operatorname{ad} e_0)^{\ell-1}(e_1).$$

Le crochet de (4.4.1) avec $we_0 + \sum w_\ell ade_0^{\ell-1}(e_1)$ ne dépend que de w_0. Il en résulte que $\Lambda_\ell(z_\alpha)$ se modifie par l'addition d'une combinaison linéaire de termes de la forme suivante:

$$\left(a(z_\alpha)\log(1-z_\alpha) - a(1-z_\alpha)\log z_\alpha\right)(\log z_\alpha)^b(\log z_\alpha + a(z_\alpha))^c \quad \text{avec } b+c = \ell-2;$$

$$\Lambda_b(z_\alpha)a(z_\alpha)\log(z_\alpha)^c(\log z_\alpha + a(z_\alpha))^d \quad \text{avec } b+1+c+d = \ell.$$

L'invariance du membre de gauche de $(4.3.5)_\ell$, $2 \leqslant \ell \leqslant k$, résulte alors de $(4.3.1)_k$ et $(4.3.5)_b$, $2 \leqslant b < \ell$, et (i) en résulte.

La preuve de $(\mathrm{ii})_\ell$ est analogue, mais plus simple: $\Lambda_m(z_\alpha)$ est inchangé pour $m < \ell$. Pour $m \geqslant \ell$, il se change par l'addition d'un multiple de

$$a_\ell(z_\alpha)(\log z)^{m-\ell}.$$

L'invariance du membre de gauche de $(4.3.5)_m$, $2 \leqslant m \leqslant k$, résulte alors de l'hypothèse faite sur $a_\ell(z_\alpha)$ et $(\mathrm{ii})_\ell$ en résulte.

4.5. Réciproquement, si $\sum \lambda_\alpha \{z_\alpha\}_k^\sim$ vérifie $(4.3.2)_k$ et que deux systèmes de déterminations des $L(z_\alpha)$ (chacun prolongeant un de log) vérifient $(4.3.5)_\ell$ pour $2 \leqslant \ell \leqslant k-1$, on passe de l'un à l'autre par les opérations 4.4 (i), $(\mathrm{ii})_\ell$ $(2 \leqslant \ell \leqslant k-1)$, suivi d'opérations $(\mathrm{ii})_\ell$ pour $\ell \geqslant k$, sans contrainte sur les a_ℓ. Une opération (i) ramène au cas où les log utilisés sont les mêmes. On retourne ensuite l'argument de 4.4 (ii) pour montrer que les opérations (4.1) (b_k) $(k \geqslant 2)$ à effectuer pour passer d'un système à l'autre sont du type voulu.

Sous les hypothèses de 4.3, la somme

$$\sum \lambda_\alpha \Lambda_k(z_\alpha)$$

est donc définie avec une ambiguité dans $(2\pi i)^k\mathbb{Q}$.

4.6. PROPOSITION. *Admettons la conjecture* 1.10. *Si une combinaison linéaire formelle* $\sum \lambda_\alpha \{z_\alpha\}_k^\sim$ *d'éléments de* $F-\{0, 1\}$ *vérifie* $d_k^\sim(\sum \lambda_\alpha \{z_\alpha\}_k^\sim) = 0$, *alors*

(i) $(4.3.1)_k$ *est vérifié.*

(ii) *Il existe des déterminations* $L(z_\alpha)$ *prolongeant une de log vérifiant* $(4.3.5)_\ell$ *pour* $2 \leqslant \ell \leqslant k-1$. *Pour celles-ci, on a*

(iii) $\sum \lambda_\alpha \Lambda_k(z_\alpha) =$ *reg* $\varphi_k(\sum \lambda_\alpha \{z_\alpha\}_k)$ *dans* $\mathbb{C}/(2\pi i)^k\mathbb{Q}$.

PREUVE. Sur la catégorie tannakienne $\mathscr{T}(S)$, on dispose des deux foncteurs fibres ω et $M \mapsto$ real $(M)_\mathbb{Q}$. Parce que $G = \mathbb{G}_m \ltimes U$, son H^1 est nul et, par la théorie générale des catégories tannakiennes, il existe un isomorphisme de foncteurs fibres $\alpha : \omega \xrightarrow{\sim}$ real $(\ )_\mathbb{Q}$.

On définit comme suit un isomorphisme entre les complexifiés de ces foncteurs fibres:

$$\text{real}\,(M)_\mathbb{C} = \bigoplus \text{real}\,(M)_i = \bigoplus \mathrm{Gr}^W_{-2i}\,\text{real}\,(M)_\mathbb{C} = \bigoplus \text{real}\,(\mathrm{Gr}^W_{-2i}M)_\mathbb{C}$$

$$= \bigoplus \omega(M)_i \otimes \text{real}\,(\mathbb{Q}(i))_\mathbb{C} = \Big(\bigoplus \omega(M)_i\Big) \otimes \mathbb{C}. \tag{4.6.1}$$

Comparant ces isomorphismes, on obtient un automorphisme de $M \mapsto \omega(M)\otimes\mathbb{C}$, i.e., $g \in G(\mathbb{C})$, tel que real(M) (de complexifié identifié à $\omega(M)\otimes\mathbb{C}$ par 4.6.1) soit d'espace vectoriel complexe gradué sous-jacent (cf. 1.2) $\omega(M)\otimes\mathbb{C}$, et de réseau rationnel $g\omega(M)$. Cet élément g est unique à $g \mapsto g\gamma$ près, $\gamma \in G(\mathbb{Q})$. Prenant $M = \mathbb{Q}(1)$, on voit que g se projette dans le quotient $\mathbb{G}_m$ de G sur un multiple rationnel de $2\pi i$. Changeant g par $\gamma \in G(\mathbb{Q})$, on peut supposer que g est de la forme

$$g = u\tau(2\pi i).$$

Le choix de u détermine une détermination généralisée de log z ($z \in F^*$): la matrice de u agissant sur $\omega(E)$, pour E une extension de $\mathbb{Q}(0)$ par $\mathbb{Q}(1)$, est de la forme $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$, avec a additif en E et on prend $\log z = a$ pour $E = [z]$.

De même, le choix de u détermine une détermination généralisée de $L(z)$, pour $z \in F - \{0, 1\}$: la matrice de u agissant sur $\omega(\mathscr{M}_z)$. Elle est compatible à la détermination de log. De plus, chaque relation $\sum \mu_\alpha \{z_\alpha\}_\ell = 0$ fournit une relation analogue entre les coefficients $(\ell, 0)$ de log u agissant sur les $\omega(\mathscr{M}_z)$.

Par hypothèse, on a

$$\sum \lambda_\alpha \{z_\alpha\}_{k-1} \otimes (z_\alpha) = 0 \qquad \text{dans } \mathscr{L}^{k-1} \otimes F^*$$

$$(\text{resp. } \sum \lambda_\alpha (1 - z_\alpha)\Lambda(z_\alpha) = 0 \qquad \text{dans } \overset{2}{\Lambda}\mathscr{L}^1 \text{ si } k = 2),$$

et (cf. 1.7) les hypothèses 4.3 sont donc vérifiées. On a

$$\langle \log u, \sum \lambda_\alpha \{z_\alpha\}_k \rangle = \sum \lambda_\alpha \Lambda_k(z_\alpha) ;$$

$\sum \lambda_\alpha \{z_\alpha\}_k$ est le coefficient de l'extension $\varphi_k(\sum \lambda_\alpha \{z_\alpha\}_k)$ de $\mathbb{Q}(0)$ par $\mathbb{Q}(k)$, et 4.6 en résulte.

BIBLIOGRAPHIE

[BD] A. Beilinson and P. Deligne, *Motivic polylogarithm and Zagier conjecture*, to appear.

[D] P. Deligne, *Le groupe fondamental de la droite projective moins trois points*, in: Galois Groups over R, MSRI publications 16, Springer-Verlag, 1989, pp. 79–297.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA

INSTITUTE FOR ADVANCED STUDY, PRINCETON, NJ

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

The Elliptic Polylogarithm

A. BEILINSON AND A. LEVIN

Introduction

It is known that one may interpret a bunch of monodromy and differential relations for the Euler polylogarithm functions

$$Li_k(z) = \sum \frac{z^n}{n^k}, \qquad k \geq 1,$$

just by saying that the Li_k occur as entries of the period matrix of a single unipotent variation of mixed Hodge structure on $\mathbb{P}^1\setminus\{0, 1; \infty\}$. This variation has motivic origin, which implies a number of remarkable properties for the values $Li_k(z)$ at $z \in \overline{\mathbb{Q}}$ (see **[Z1, BD]**). In this paper we partially extend the classical polylogarithm pattern to the elliptic situation. For any elliptic curve X we define a canonical Hodge sheaf on $X\setminus\{0\}$—the elliptic polylogarithm sheaf. The entries of its period matrix are appropriately q-averaged versions of the Euler polylogarithm functions. The real analytic single-valued functions that describe the variation of $\mathbb{R}$-Hodge structure are certain Eisenstein-Kronecker series. The elliptic polylogarithm sheaf has motivic origin (in the sense of algebraic K-theory). Its fibers over the torsion points of X provide a collection of "Eisenstein" elements in the absolute cohomology groups of symmetric powers of X. Presumably these elements coincide with the Eisenstein symbols from **[B2]** (at least they have the same image under the regulator map). We leave the possible elliptic versions of the Zagier conjecture for a future investigation.

The paper goes as follows. In §1 we construct the elliptic polylogarithm sheaf $\mathscr{P}$ in the framework of any mixed sheaf theory (e.g., as a Hodge sheaf or a mixed system of realizations) and write down its basic properties. In §2 we describe the restriction of $\mathscr{P}$ to torsion points of X and study its p-adic properties. In §§3 and 4 we present the above-mentioned explicit formulas

1991 *Mathematics Subject Classification.* Primary 14H52; Secondary 11G05, 19F27.
This paper is in final form and no version of it will be submitted for publication elsewhere.

for the variation of mixed Hodge structure on $\mathscr{P}$. The motivic version of $\mathscr{P}$ is given in §6, while §5 contains some preliminary information about the absolute motivic cohomology with coefficients in a symmetric power of an elliptic curve. There is a construction of polylogarithm-type sheaves on curves of arbitrary genus, but we do not even know how to compute the corresponding variations of Hodge structures (presumably, they could be expressed in terms of some Green's functions on the universal covering of our curve).

A general remark. Green's functions, such as Eisenstein-Kronecker series, occur naturally in quantum field theory (as expectation values, etc.) and in Hodge theory (as matrix coefficients of a certain operator that describes a variation of mixed $\mathbb{R}$-Hodge structure). In a heuristic dictionary relating Hodge theory and odd-dimensional topology (see **[B1]**) the Green's functions in Hodge theory correspond to link indices in topology; on the other hand, the link indices occur naturally as correlators in some topological field theories. One may wonder whether there exists a "Hodge field theory", parallel to topological ones, that will incorporate the Hodge theoretic Green's functions in a natural way. Perhaps such a field theory would provide the right language for Hodge theory.

We should mention that the elliptic dilogarithm first appeared in Bloch's seminal Irvin lectures **[B1]**. The real analytic single-valued versions of higher elliptic polylogarithm functions were studied by Zagier **[Z2]**; the results of §§3 and 4 give a Hodge-theoretic explanation and proof of Theorem 1 from **[Z2]**. Nori has independently found the results of §§1 and 2. Moreover, he has shown that the corresponding topological construction, being generalized to the case of topological tori of any dimension, describes the values at negative integers of partial ζ-functions of any totally real number field.

The "motivic" papers often hopelessly increase their volume while you write them; they have a tendency to lead the reader astray from the central ideas and simple computations to the burdock thicket of generalities. We apologize as this seems to have happened with this paper.

We would like to thank Sasha Goncharov and Don Zagier for the valuable discussions and interest, the referees for their suggestions that helped to make the text more readable, and Uwe Jannsen for the generous help. We are grateful to M. Bauer for careful typing of the manuscript. The first author would like to thank MSRI, Kyoto, and Mathematisches Institut, Köln, and both authors thank MPI, Bonn, for their hospitality.

This research was partially supported by NSF grant DMS-9008488.

1. The elliptic polylogarithm

In this paper we deal with usual, or topological, sheaves and with mixed sheaves on algebraic varieties. A topological sheaf is either a constructible $\mathbb{Q}_\ell$-sheaf, or an étale sheaf, or, if the base field is $\mathbb{C}$, a constructible sheaf for the classical topology. A mixed sheaf theory is known—at this time—in the following forms:

(i) $\mathbb{Q}_\ell$-theory for schemes of finite type over $\mathbb{F}_p$ or $\mathbb{Q}$ (see **[BBD]**),
(ii) Hodge theory for schemes of finite type over $\mathbb{R}$ or $\mathbb{C}$ (see **[S1]**),
(iii) "mixed system of realizations" theory for schemes of finite type over $\mathbb{Q}$ (see **[S2]**).[1]

When we speak about mixed sheaves, we assume that we are in one of the situations (i)–(iii); we denote by k the base field for our category of schemes. Usually we will deal with lisse sheaves. We normalize the numbering of the cohomological functors in a "classical", not "perverse", way (e.g., a curve has nonzero cohomology groups in degrees $0, 1, 2$).

We denote by F the coefficient ring of our sheaf theory. For topological sheaves F is an arbitrary commutative ring if we deal with the classical topology ($k = \mathbb{C}$), a finite commutative ring of characteristics prime to that of k in the étale situation, and $F = \mathbb{Q}_\ell$, where ℓ is prime to the characteristic of k, in the $\mathbb{Q}_\ell$-case. For mixed sheaves F is $\mathbb{Q}_\ell$ in situation (i), $F = \mathbb{Q}$ or $F = \mathbb{R}$ in situation (ii), and $F = \mathbb{Q}$ in situation (iii). For a scheme B we denote by $\mathscr{S}\ell(B, F)_{\text{top}} = \mathscr{S}\ell(B)_{\text{top}}$ the category of lisse topological F-sheaves on B, and by $\mathscr{S}\ell(B)_{\text{mixed}}$ that of lisse mixed sheaves; one has the "forgetting the mixed structure" functor $o\colon \mathscr{S}\ell(B)_{\text{mixed}} \to \mathscr{S}\ell(B)_{\text{top}}$. We denote by $\mathscr{S}\ell(B)$ either of the above categories. $\mathscr{S}\ell(B)$ is a tensor category; if F is a field, it is a Tannakian F-category assuming that B is connected. We denote by F_B the constant sheaf with fiber F. For $M \in \mathscr{S}\ell(B)$ we put $M' := \underline{\mathrm{Hom}}(M, F_B)$, $M(i) := i$-fold Tate twist of M.

We shall also use prosheaves—projective limits of sheaves. Assume that F is Artinian. Then the F-category $\mathscr{S}\ell(B)^\wedge := \varprojlim \mathscr{S}\ell(B)$ is abelian (since objects of $\mathscr{S}\ell(B)$ have finite length). We extend any functor H between the categories of sheaves to that of prosheaves as $H \varprojlim = \varprojlim H$. If H is an exact functor (from the left or from the right), or H is a cohomological functor, then so is the extension of H to prosheaves.

If F is an arbitrary commutative ring (this happens only if we deal with sheaves on the classical topology), then the prosheaves are projective limits of sheaves with respect to a directed family of epimorphisms. The category $\mathscr{S}\ell(B)^\wedge$ is an exact category. Any right exact functor H between the categories of sheaves extends to prosheaves.

1.1. The unipotent sheaves. Let B be a connected scheme, $p = p_X\colon X \to B$ be a family of elliptic curves over B, and $0 = 0_X \in X(B)$ be the zero section.

1.1.1. Consider the exact tensor functors $\mathscr{S}\ell(B) \underset{0^*}{\overset{p^*}{\rightleftarrows}} \mathscr{S}\ell(X)$. The functor p^* is fully faithful (since the fibers of p are connected) and $0^*p^* = \mathrm{Id}_{\mathscr{S}\ell(B)}$.

[1] At this time the theory (iii) is based on comparison between de Rham and Betti cohomology over $\mathbb{C}$ only; the precise theory that will take into account non-Archimedean places (p-adic Hodge theory) remains to be developed. A definition of lisse system of realizations was first given in **[D, J]**.

The embedding p^* admits left and right adjoints p_* and $p'_*, p'_*(\mathcal{F}) := R^2p_*\mathcal{F}(1)$ (by Poincaré duality), respectively. A sheaf $\mathcal{F} \in \mathcal{S}\ell(X)$ has two canonical filtrations $\mathcal{F} = \mathcal{F}^0 \supset \mathcal{F}^1 \supset \cdots$, $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}$ defined inductively by $\mathcal{F}^{i+1} := \operatorname{Ker}(\mathcal{F}^i \to p^*p'_*\mathcal{F}^i)$, $\mathcal{F}_0 := p^*p_*\mathcal{F}$, $\mathcal{F}_{i+1}/\mathcal{F}_i = (\mathcal{F}/\mathcal{F}_i)_0$. If F is a field then $(\mathcal{F}')_i = (\mathcal{F}/\mathcal{F}^{i+1})'$.

A filtration $A^\cdot$ on $\mathcal{F}$, $\mathcal{F} = A^0\mathcal{F} \supset A^1\mathcal{F} \supset \cdots \supset A^{n+1}\mathcal{F} = 0$, is called *unipotent* (of length n) if $\operatorname{gr}^i_A\mathcal{F} \in p^*\mathcal{S}\ell(B)$. A sheaf that admits such a filtration is called *n-unipotent* (relative to p). Denote by $\mathcal{S}\ell(X)^{\text{un}}_n$ the full subcategory of such sheaves; this subcategory is closed under subquotients and duality. The embedding $\mathcal{S}\ell(X)^{\text{un}}_n \hookrightarrow \mathcal{S}\ell(X)$ has left and right adjoints $\mathcal{F} \mapsto \mathcal{F}/\mathcal{F}^{n+1}$, $\mathcal{F} \mapsto \mathcal{F}_n$ respectively. In particular, $\mathcal{F}$ is n-unipotent iff either $\mathcal{F}^{n+1} = 0$ or $\mathcal{F} = \mathcal{F}_n$. We have the filtration $p^*\mathcal{S}\ell(B) = \mathcal{S}\ell(X)^{\text{un}}_0 \subset \mathcal{S}\ell(X)^{\text{un}}_1 \subset \cdots \subset \mathcal{S}\ell(X)^{\text{un}} := \bigcup \mathcal{S}\ell(X)^{\text{un}}_n \subset \mathcal{S}\ell(X)$. Since $\mathcal{S}\ell(X)^{\text{un}}_n \otimes \mathcal{S}\ell(X)^{\text{un}}_m \subset \mathcal{S}\ell(X)^{\text{un}}_{n+m}$, the category $\mathcal{S}\ell(x)^{\text{un}}$ of unipotent sheaves is a tensor subcategory of $\mathcal{S}\ell(X)$.

Note that a lisse mixed sheaf is unipotent iff the weight filtration is unipotent (since any pure unipotent sheaf lies in $p^*\mathcal{S}\ell(B)$—look at the weights on R^1p_*).

We shall call a projective limit of unipotent sheaves simply a prounipotent sheaf; these form an exact F-category $\mathcal{S}\ell(X)^{un\wedge} := \varprojlim \mathcal{S}\ell(X)^{\text{un}}$ which is abelian if F is an Artinian ring.

1.1.2. Put $\mathcal{H} := R^1p_*F_X(1)$—this is a rank 2 lisse sheaf on B of weight -1 (if we consider mixed sheaves). The trace map $\operatorname{tr}\colon R^2p_*F_X(1) \to F_B$ is an isomorphism; hence, the cup product defines a canonical skew-symmetric intersection pairing $\langle\ ,\ \rangle\colon \mathcal{H}\otimes\mathcal{H} \to F_B(1)$ which provides the isomorphism $\Lambda^2\mathcal{H} \xrightarrow{\sim} F_B(1)$. We shall often consider $\mathcal{H}$ as a commutative Lie algebra of rank 2. Put $S^i := \operatorname{Sym}^i\mathcal{H}$, so that $S^\cdot = \bigoplus S^i$ is the universal enveloping algebra of $\mathcal{H}$. Let $F[[\mathcal{H}]] := \prod_{i\geq 0} S^i$ be the completion of S with respect to powers of the augmentation ideal.

Note that for any sheaf $\mathcal{L} \in \mathcal{S}\ell(B)$ one has canonical isomorphisms $R^ip_*p^*\mathcal{L} = (R^ip_*F_X)\otimes\mathcal{L}$; i.e.,

$$p_*p^*\mathcal{L} = \mathcal{L}, \qquad R^1p_*p^*\mathcal{L} = \mathcal{H}'\otimes\mathcal{L}, \qquad R^2p_*p^*\mathcal{L} = \mathcal{L}(-1).$$

Therefore, for any unipotent sheaf $\mathcal{F}$ and a unipotent filtration $A^\cdot$ on $\mathcal{F}$ the R^ip_*-boundary maps for the short exact sequences $0 \to \operatorname{gr}^{i+1}_A\mathcal{F} \to A^i/A^{i+2}(\mathcal{F}) \to \operatorname{gr}^i_A\mathcal{F} \to 0$ are

$$p_*\operatorname{gr}^i_A\mathcal{F} \xrightarrow{\alpha} \mathcal{H}'\otimes p_*\operatorname{gr}^{i+1}_A\mathcal{F} \xrightarrow{\beta} p_*\operatorname{gr}^{i+2}_A\mathcal{F}(-1).$$

The above morphism α provides a map $\mathcal{H}\otimes p_*\operatorname{gr}^i_A\mathcal{F} \to p_*\operatorname{gr}^{i+1}_A\mathcal{F}$; since $\beta\alpha = 0$, we see that this is an action of $\mathcal{H}$ on $p_*\operatorname{gr}^\cdot_A\mathcal{F}$, i.e., $p_*\operatorname{gr}^\cdot_A\mathcal{F}$ is an $S^\cdot$-module.

1.1.3. An induction by the length of $A^{\cdot}$ shows that $A^{\cdot}$ coincides with the canonical filtration $\mathscr{F}^{\cdot}$ iff $p_* \operatorname{gr}^{\cdot}_A \mathscr{F}$ is generated by $p_* \operatorname{gr}^0_A \mathscr{F}$ as $\mathscr{S}^{\cdot}$-module.

1.2. The logarithm sheaf. In this section we describe a universal unipotent prosheaf on X. For a sheaf $\mathscr{F}$ on X we put $\mathscr{F}_0 := 0^*\mathscr{F}$.

1.2.1. Assume first that $k = \mathbb{C}$ and we deal with the classical topology. Put $\mathscr{H}_{\mathbb{Z}} := R^1 p_* \mathbb{Z}_X(1)$—this is a local system on B with fibers $\mathscr{H}_{\mathbb{Z}b} := H_1(X_b, \mathbb{Z})$. The fiberwise exponential map provides a short exact sequence of sheaves of abelian groups on B

$$0 \to \mathscr{H}_{\mathbb{Z}} \to T \overset{\exp}{\to} X \to 0$$

where $T := 0^*\mathscr{T}_{X/B}$ is the bundle of vertical tangent lines at 0. We get a $p^*\mathscr{H}_{\mathbb{Z}}$-torsor Log on X with fibers $\operatorname{Log}_x := \exp^{-1}(x)$, $x \in X$. In other words Log_x is the set of homotopy paths in $X_{p(x)}$ that connect 0 and x. The torsor Log is trivialized at 0; let "0" be the corresponding section of $\operatorname{Log}_0 = 0^* \operatorname{Log}$. The group structure on T defines, as usual, an isomorphism of $p^*\mathscr{H}_{\mathbb{Z}}$-torsors $\operatorname{pr}_1^* \operatorname{Log} + \operatorname{pr}_2^* \operatorname{Log} \to (+_X)^* \operatorname{Log}$; here $+_X$, pr_1, $\operatorname{pr}_2 \colon X \times_B X \to X$ are the sum operation and projections respectively and $+$ is the sum operation on torsors. This isomorphism satisfies associativity and commutativity constraints, and "0" is the neutral element.

Note that for any locally constant sheaf $\mathscr{L}$ of sets on X the sheaf $\mathscr{L}_0 = 0^*\mathscr{L}$ on B carries a canonical $\mathscr{H}_{\mathbb{Z}}$-action (the monodromy along the fibers of p), and $\mathscr{L} \mapsto \mathscr{L}_0$ gives an equivalence between the categories of locally constant sheaves of sets on X and those on B equipped with an $\mathscr{H}_{\mathbb{Z}}$-action. One may spell this equivalence in a different way. Namely, we have a canonical isomorphism of sheaves

$$p_* \underline{\operatorname{Hom}}(\operatorname{Log}, \mathscr{L}) \overset{\sim}{\to} \mathscr{L}_0, \qquad \varphi \mapsto \varphi(\text{“0”}),$$

which identifies the monodromy action on $\mathscr{L}_0$ with the action provided by the $\mathscr{H}_{\mathbb{Z}}$-action on Log. The inverse equivalence assigns to an $\mathscr{H}_{\mathbb{Z}}$-sheaf $\mathscr{M}$ on B the sheaf $p^*\mathscr{M}_{\operatorname{Log}} := \mathscr{H}_{\mathbb{Z}} \backslash (p^*\mathscr{M} \times \operatorname{Log})$ on X.

1.2.2. One may consider an F-linearized version of the above data. We have the sheaf of rings $F[\mathscr{H}_{\mathbb{Z}}]$ on B (the group algebra of $\mathscr{H}_{\mathbb{Z}}$) and a free $p^*F[\mathscr{H}_{\mathbb{Z}}]$-module of rank one $F[\operatorname{Log}]$ on X. Note that $F[\mathscr{H}_{\mathbb{Z}}]$ is a Hopf algebra and $F[\operatorname{Log}]$ is an F-coalgebra (the comultiplication $\delta \colon F[\operatorname{Log}] \to F[\operatorname{Log}] \otimes F[\operatorname{Log}]$ is $\delta(a) = a \otimes a$ for $a \in \operatorname{Log}$). The zero fiber $F[\operatorname{Log}]_0$ has a canonical base vector $1 :=$ "0". We have the product operation $\operatorname{pr}_1^* F[\operatorname{Log}] \otimes \operatorname{pr}_2^* F[\operatorname{Log}] \to (+)^* F[\operatorname{Log}]$, $+ = +_X$, which is commutative and associative; on the zero fiber this operation coincides with multiplication in $F[\mathscr{H}_{\mathbb{Z}}]$ (i.e., 1 is a unit element for the product).

For any locally constant sheaf $\mathscr{L}$ of F-modules on X we have a canonical isomorphism $p_* \underline{\operatorname{Hom}}(F[\operatorname{Log}], \mathscr{L}) \overset{\sim}{\to} \mathscr{L}_0$, $\varphi \mapsto \varphi(1)$. Therefore, $\mathscr{L}_0$ is a sheaf of $F[\mathscr{H}_{\mathbb{Z}}]$-modules, and $\mathscr{L} \mapsto \mathscr{L}_0$ gives an equivalence between the categories of locally constant sheaves of F-modules on X and those of $F[\mathscr{H}_{\mathbb{Z}}]$-modules on B. The inverse functor is $\mathscr{M} \mapsto p^*\mathscr{M} \otimes_{F[\mathscr{H}_{\mathbb{Z}}]} F[\operatorname{Log}]$.

1.2.3. Denote by $\mathscr{R}$ the completion of $F[\mathscr{H}_{\mathbb{Z}}]$ with respect to the powers of the augmentation ideal. Let $I \subset \mathscr{R}$ be the (completed) image of the augmentation ideal. We have a canonical isomorphism of graded F-algebras $\operatorname{gr}_I^{\cdot}\mathscr{R} \xrightarrow{\sim} S^{\cdot} := \operatorname{Sym}^{\cdot}\mathscr{H}_F$ that sends $(a-1) \bmod I^2 \in I/I^2$, $a \in \mathscr{H}_{\mathbb{Z}}$, to $a \in \mathscr{H}_{\mathbb{Z}} \otimes F = \mathscr{H}_F$. We shall consider $\mathscr{R}$ as a lisse topological prosheaf, $\mathscr{R} = \varprojlim \mathscr{R}/I^n$; it inherits the Hopf algebra structure from $F[\mathscr{H}_{\mathbb{Z}}]$.

Put $G := \mathscr{R} \otimes_{F[\mathscr{H}_{\mathbb{Z}}]} F[\mathrm{Log}]$; equivalently, G is the completion of $F[\mathrm{Log}]$ by powers of the augmentation ideal. We shall call G the *logarithm sheaf*; it is a free $p^*\mathscr{R}$-module of rank one equipped with a comultiplication $\delta: G \to G \otimes G$. We have a canonical isomorphism $G_0 \xrightarrow{\sim} \mathscr{R}$, $1 \mapsto 1$, and a product operation $\mathrm{pr}_1^* G \otimes \mathrm{pr}_2^* G \to (+)^* G$. Note that the monodromy acts on $\operatorname{gr}_I^{\cdot} G$ trivially. Hence, one has a canonical isomorphism of $S^{\cdot}$-modules $\operatorname{gr}_I^{\cdot} G \xrightarrow{\sim} p^* S^{\cdot}$ that coincides with the above identification on the zero fiber. For any lisse unipotent sheaf $\mathscr{L}$ on X one has $p_* \underline{\mathrm{Hom}}(G, \mathscr{L}) = \mathscr{L}_0$, and the functor $\mathscr{L} \mapsto \mathscr{L}_0$ is an equivalence between the category of lisse unipotent sheaves of F-modules on X and that of lisse sheaves of $\mathscr{R}$-modules on B. The inverse functor is $\mathscr{M} \mapsto p^*\mathscr{M} \otimes_{\mathscr{R}} G$. This equivalence is compatible with F-tensor product (use the coproduct δ on G).

Note that if $F \supset \mathbb{Q}$ then we have a canonical isomorphism $\mathscr{R} \xrightarrow{\sim} F[[\mathscr{H}]] := \prod S^i$ that sends $a \in \mathscr{H}_{\mathbb{Z}} \subset \mathscr{R}$ to $\exp(a) \in F[[\mathscr{H}]]$. This is an isomorphism of Hopf algebras where the comultiplication $\delta: F[[\mathscr{H}]] \to F[[\mathscr{H}]] \otimes F[[\mathscr{H}]]$ is $\delta(h) = h \otimes 1 + 1 \otimes h$, $h \in \mathscr{H}$.

1.2.4. Let $\varphi: X \to X'$ be an isogeny of elliptic curves over B. It induces the morphisms $\varphi^{\cdot}: \mathscr{H}_{\mathbb{Z}X} \to \mathscr{H}_{\mathbb{Z}X'}$, $\mathrm{Log}_X \to \varphi^* \mathrm{Log}_{X'}$ and the corresponding morphisms $\varphi_{\cdot\mathscr{H}}: \mathscr{H}_{FX} \to \mathscr{H}_{FX'}$, $\varphi_{\cdot\mathscr{R}}: \mathscr{R}_X \to \mathscr{R}_{X'}$, $\varphi_{\cdot G}: G_X \to G_{X'}$. Note that $\operatorname{gr}_I^i \varphi_{\cdot\mathscr{R}}: \operatorname{gr}_I^i \mathscr{R}_X \to \operatorname{gr}_I^i \mathscr{R}_{X'}$ coincides with $\operatorname{Sym}^i(\varphi_{\cdot\mathscr{H}})$; therefore, if the degree of φ is invertible in F, then $\varphi_{\cdot\mathscr{R}}$, $\varphi_{\cdot G}$ are isomorphisms. Since the finite subgroup $\operatorname{Ker}\varphi$ acts on $\varphi^* G_{X'}$ (by translations), we see that in this situation G_X is also $\operatorname{Ker}\varphi$-equivariant (via φ_G^{-1}). For different φ's these morphisms are compatible; therefore, G_X is equivariant with respect to translations by points of finite order invertible in F; denote this subgroup by $X_{\mathrm{tors}}^{(F)} \subset X$. In particular, for any $x \in X_{\mathrm{tors}}^{(F)}$ we have a canonical isomorphism $G_x \xrightarrow{\sim} G_0 = \mathscr{R}$.

1.2.5. The above constructions have the standard étale versions. One should replace Log by the projective limit of Kummer torsors, $\mathscr{K}_n$, $\mathscr{K}_{nx} = [n]^{-1}(x)$, etc. Note that $\mathscr{R} = \varprojlim_n F[X_n]$, $G = \varprojlim_n \mathscr{K}_n$, n is nilpotent in F (here $X_n = \mathscr{K}_{n0} \subset X$ is the subscheme of points of order n, $[n]: X \to X$ is multiplication by n isogeny). All the properties listed in 1.2.3, 1.2.4 remain valid in this situation.

We may pass to the limit to get the $\mathbb{Q}_\ell$-version of the logarithm sheaf. The sheaf G is mixed, $W_{-i}G = I^i G$. The Hodge and mixed system of realizations versions of it could also be easily constructed "by hands" (see §§3 and 4 for the Hodge versions). Below in 1.2.6–1.2.8 we shall give a general "categorial" construction of G that works in any situation (this construction is a very particular case of a general principle saying that various "homotopy path" sheaves are at hand whenever we have the appropriate cohomology formalism).

Let $G^{(n)}$ be an n-unipotent sheaf, $1^{(n)} \in \mathrm{Hom}(F_B, G_0^{(n)})$ a section of the zero fiber. For a length n unipotent filtration $A^\cdot$ on $G^{(n)}$ denote by $\nu^\cdot : S^\cdot/S^{\geq n+1} \to p_* \mathrm{gr}_A^\cdot G^{(n)}$ the unique morphism of $S^\cdot$-modules that sends 1 to $1^{(0)} := 1^{(n)} \bmod A^1$.

1.2.6. PROPOSITION. *The following conditions on* $(G^{(n)}, 1)$ *are equivalent.*

$(\mathrm{a})_n$ *The morphism* $\nu^\cdot : S^\cdot/S^{\geq n+1} \to p_* \mathrm{gr}^\cdot G^{(n)}$ *for the canonical filtration on* $G^{(n)}$ *is isomorphism.*

$(\mathrm{a})'_n$ *There exists a length* n *unipotent filtration* $A^\cdot$ *such that the corresponding map* $\nu^\cdot : S^\cdot/S^{\geq n+1} \to p_* \mathrm{gr}_A^\cdot G^{(n)}$ *is an isomorphism.*

$(\mathrm{b})_n$ *For any* $\mathcal{F} \in \mathcal{S}\ell(X)_n^{\mathrm{un}}$ *the morphism* $p_*\underline{\mathrm{Hom}}(G^{(n)}, \mathcal{F}) \to \mathcal{F}_0$, $\varphi \mapsto \varphi(1)$, *is an isomorphism.*

Such a pair $(G^{(n)}, 1^{(n)})$ *exists and is unique (up to a unique isomorphism).*

1.2.7. LEMMA (computation of $R^i p_* G^{(n)}$). *For* $(G^{(n)}, 1^{(n)})$ *as in* 1.2.6 *the embedding* $(\nu^n)^{-1} : p^* S^n \hookrightarrow G^{(n)}$ *induces an isomorphism* $S^n \xrightarrow{\sim} p_* G^{(n)}$. *The morphism* $\mathcal{H}(-1) \otimes S^n = R^1 p_* p^* S^n \to R^1 p_* G^{(n)}$ *factors through the isomorphism* $S^{n+1}(-1) \xrightarrow{\sim} R^1 p_* G^{(n)}$. *The image* $Rp_*(1^{(n)})$ *of* $1^{(n)}$ *by the composition*

$$\begin{aligned}\mathrm{Hom}(F_B, G_0^{(n)}) &= \mathrm{Hom}(F_B(-1)[-2], R0^! G^{(n)}) \\ &= \mathrm{Hom}(0_* F_B(-1)[-2], G^{(n)}) \\ &\to \mathrm{Hom}(F_B(-1)[-2], Rp_* G^{(n)})\end{aligned}$$

induces an isomorphism $F_B(-1) \xrightarrow{\sim} R^2 p_* G^{(n)}$.

PROOF OF 1.2.7. Assume $(G^{(n)}, 1^{(n)})$ satisfies $(\mathrm{a})_n$. Let us compute $R^i p_* G^{(n)}$ using the spectral sequence for the filtration $G^{(n)}$. One has $E_1^{p,q} = (R^{p+q} p_* F_X) \otimes S^p$ for $p \leq n$, $E_1^{p,q} = 0$ for $p > n$. The d_1-differentials are the maps α, β from 1.1.2, so E_1 is a truncated Koszul complex. The only nonzero E_2 terms are $E_2^{n,-n}$, $E_2^{n,1-n}$, and $E_2^{0,2}$, which implies 1.2.7. □

PROOF OF 1.2.6. (i) Assume $(G^{(n)}, 1^{(n)})$ satisfies $(\mathrm{a})_n$. Let us check property $(\mathrm{b})_n$. Using induction by n, since

$$(G^{(n-1)}, 1^{(n-1)}) := (G^{(n)}/G^{(n)n}, 1^{(n)} \bmod G^{(n)n})$$

satisfies $(\mathrm{a})_{n-1}$, we may assume that it satisfies $(\mathrm{b})_{n-1}$. Now for $\mathcal{F} \in \mathcal{S}\ell(X)_n^{\mathrm{un}}$ consider the short exact sequence $0 \to \mathcal{F}^n \to \mathcal{F} \to \mathcal{F}/\mathcal{F}^n \to 0$.

Since $\mathscr{F}/\mathscr{F}^n \in \mathscr{S}\ell(X)^{\mathrm{un}}_{n-1}$, $\mathscr{F}^n \in p^*\mathscr{S}\ell(B)$, one has $p_*\underline{\mathrm{Hom}}(G^{(n)}, \mathscr{F}^n) = p_*\underline{\mathrm{Hom}}(G^{(n-1)}, \mathscr{F}/\mathscr{F}^n)$; same for $\mathscr{F}/\mathscr{F}^n$ replaced by $\mathscr{F}^n$. The commutative diagram

$$\begin{array}{ccccccccc}
0\to & \mathscr{F}_0^n & \to & \mathscr{F}_0 & \to & (\mathscr{F}/\mathscr{F}^n)_0 & \to & 0 \\
 & \uparrow\wr & & \uparrow & & \uparrow\wr & & \\
0\to & p_*\underline{\mathrm{Hom}}(G^{(n)},\mathscr{F}^n) & \to & p_*\underline{\mathrm{Hom}}(G^{(n)},\mathscr{F}) & \to & p_*\underline{\mathrm{Hom}}(G^{(n)},\mathscr{F}/\mathscr{F}^n) & \to & R^1p_*\underline{\mathrm{Hom}}(G^{(n)},\mathscr{F}^n) \\
 & \uparrow\wr & & \uparrow & & \uparrow\wr & & \uparrow\mu \\
0\to & p_*\underline{\mathrm{Hom}}(G^{(n-1)},\mathscr{F}^n) & \to & p_*\underline{\mathrm{Hom}}(G^{(n-1)},\mathscr{F}) & \to & p_*\underline{\mathrm{Hom}}(G^{(n-1)},\mathscr{F}/\mathscr{F}^n) & \to & R^1p_*\underline{\mathrm{Hom}}(G^{(n-1)},\mathscr{F}^n)
\end{array}$$

shows that $(\mathrm{b})_n$ follows from the vanishing of the arrow μ.

Put $\mathscr{L} = \mathscr{F}_0^n$; so $\mathscr{F}^n = p^*\mathscr{L} = Rp^!\mathscr{L}(-1)[-2]$. By Poincaré duality, $R^1p_*\underline{\mathrm{Hom}}(G^{(n)}, \mathscr{F}) = R^1\underline{\mathrm{Hom}}(Rp_*G^{(n)}, \mathscr{L}(-1)[-2])$; same for $G^{(n-1)}$.

Consider the sum $\Psi^{(n)}$ of $Rp_*(1^{(n)}) \in \mathrm{Hom}(F_B(-1)[-2], Rp_*G^{(n)})$ (see 1.2.7) and the canonical morphism for τ-truncation

$$\Psi^{(n)}\colon F_B(-1)[-2] \oplus \tau_{\leq 1}Rp_*G^{(n)} \to Rp_*G^{(n)}.$$

By 1.2.7 $\Psi^{(n)}$ and the similar map $\Psi^{(n-1)}$ are quasi-isomorphisms. Therefore, the map $R^1\underline{\mathrm{Hom}}(Rp_*G^{(n)}, \mathscr{L}(-1)[-2]) \to \underline{\mathrm{Hom}}(R^1p_*G^{(n)}, \mathscr{L}(-1)) = \underline{\mathrm{Hom}}(S^{n+1}, \mathscr{L})$ is an isomorphism; same for the $G^{(n-1)}$-case. Since the morphism $R^1p_*G^{(n)} \to R^1p_*G^{(n-1)}$ vanishes by 1.2.7, we see that $\mu = 0$.

(ii) The conditions $(\mathrm{a})_n$, $(\mathrm{a})'_n$ are equivalent by 1.1.3.

(iii) It remains to prove the existence of $(G^{(n)}, 1^{(n)})$ that satisfies $(\mathrm{a})_n$ (since the property $(\mathrm{b})_n$ characterizes our pair uniquely, this will finish the proof of 1.2.6). The induction by n yields us $(G^{(n-1)}, 1^{(n-1)})$ that satisfies $(\mathrm{a})_{n-1}$. The above quasi-isomorphism $\Psi^{(n-1)}$ defines a retraction $Rp_*G^{(n-1)} \to \tau_{\leq 1}Rp_*G^{(n-1)}$ that vanishes on $F_B(-1)[-2]$. Let $\varepsilon \in \mathrm{Hom}(Rp_*G^{(n-1)}, S^{n-1}(-1)[-1])$ be the composition of this retraction with the projection

$$\tau_{\leq 1}Rp_*G^{(n-1)} \to R^1p_*G^{(n-1)} = S^{n-1}(-1)[-1].$$

Again, by the Poincaré duality adjunction, we may consider ε as a morphism $G^{(n-1)} \to Rp^!S^n(-1)[-1] = p^*S^n[1]$; i.e., $\varepsilon \in \mathrm{Ext}^1(G^{(n-1)}, p^*S^n)$.

Let $0 \to p^*S^n \to G^{(n)} \to G^{(n-1)} \to 0$ be the corresponding extension. By the construction ε vanishes on $1^{(n-1)} \in \mathrm{Hom}(0_*F_B(-1)[-2], G^{(n-1)})$; hence, $1^{(n-1)}$ lifts to a section $1^{(n)} \in \mathrm{Hom}(F_B, G_0^{(n)})$. The sheaf $G^{(n)}$ carries an n-unipotent filtration $A^\cdot$ defined as $A^n := p^*S^n$, $A^i/A^n := G^{(n-1)i}$. The isomorphism $\nu^\cdot\colon p_*g^\cdot G^{(n-1)} \xrightarrow{\sim} S^\cdot/S^{\geq n}$ defines an isomorphism $p_*\,\mathrm{gr}^\cdot_A\, G^{(n)} \xrightarrow{\sim} S^\cdot/S^{\geq n+1}$. By construction this is an isomorphism of $S^\cdot$-modules. Hence, $(G^{(n)}, 1^{(n)})$ satisfies $(\mathrm{a})'_n$. $\square$

1.2.8. Consider the array $\{(G^{(n)}, 1^{(n)})\}$. One has the canonical morphisms $G^{(n+1)} \to G^{(n)}$ that send $1^{(n+1)}$ to $1^{(n)}$; they identify $G^{(n)}$ with $G^{(n+1)}/G^{(n+1)n}$. Put $G := \varprojlim G^{(n)}$, $1 = \varprojlim 1^{(n)} \in \mathrm{Hom}(F_B, G_0)$; the prounipotent sheaf G is our *logarithm sheaf*. According to 1.2.6, 1.2.7 it satisfies the following properties.

(a) One has a canonical isomorphism $\nu^{\cdot}: p^*S^{\cdot} \xrightarrow{\sim} \mathrm{gr}^{\cdot} G$.

(b) For any unipotent sheaf $\mathscr{F} \in \mathscr{S}\ell(X)^{\mathrm{un}}$ the morphism $p_*\underline{\mathrm{Hom}}(G, \mathscr{F}) \to \mathscr{F}_0$, $\varphi \mapsto \varphi(1)$, is an isomorphism.

(c) The prosheaves R^ip_*G vanish for $i \neq 2$, and 1 defines an isomorphism $R^2p_*G = F(-1)$.

As follows from 1.2.6(a)$'$, the formation $(G, 1)$ is compatible with base change for our family of curves. Note that the property (b) implies that in the classical or étale topology situation $(G, 1)$ coincides with the logarithm sheaf defined in 1.2.3, 1.2.5.

In the following lemma we assume that F is an Artinian ring.

1.2.9. Lemma. *Let $(G', 1')$ be another pair, $G' \in \mathscr{S}\ell(X)^{un\wedge}$, $1' \in \mathrm{Hom}(F_B, 0^*G')$. Assume that $R^1p_*G' = 0$ and $1'$ defines an isomorphism $F_B(-1) \xrightarrow{\sim} R^2p_*G'$. Then the morphism $\alpha: G \to G'$ that sends 1 to $1'$ is an isomorphism.*

Proof. An object $C \in \mathscr{S}\ell(X)^{un\wedge}$ vanishes iff $R^2p_*C = 0$. Since R^2p_* is right exact, we see that α is surjective (take $C = \mathrm{Coker}\,\alpha$). The long exact sequence of R^ip_* for $0 \to \mathrm{Ker}\,\alpha \to G \to G' \to 0$ shows that $R^2p_*\,\mathrm{Ker}\,\alpha = 0$; i.e., α is injective. □

1.2.10. The sheaf G automatically carries all the structures that were defined in the classical or étale topology situation in 1.2.3–1.2.5. Namely:

(i) One has a canonical coproduct $\delta: G \to G \otimes G$ uniquely defined by the property $\delta(1) = 1 \times 1$ (see 1.2.8(b)); this δ is cocommutative and coassociative.

(ii) One has a canonical product $\mathrm{pr}_1^*G \otimes \mathrm{pr}_1^*G \to (+)^*G$ uniquely defined by the property $1 \cdot 1 = 1$ (since for any unipotent sheaf $\mathscr{F} \in \mathscr{S}\ell(X \times_B X)$ one has $(p \times p)_*\underline{\mathrm{Hom}}(\mathrm{pr}_1^*G \otimes \mathrm{pr}_2^*G, \mathscr{F}) = (p \times p)_*\underline{\mathrm{Hom}}(\mathrm{pr}_1^*G, \underline{\mathrm{Hom}}(\mathrm{pr}_2^*G, \mathscr{L})) = \mathscr{L}_{0\times 0}$ by the base change property for G and 1.2.8(b)). This product is commutative and associative and compatible with δ.

(iii) As follows from the definitions, both product and coproduct are compatible with the filtration $G^{\cdot}$, and the isomorphism $\nu^{\cdot}: p^*S^{\cdot} \xrightarrow{\sim} \mathrm{gr}^{\cdot} G$ is compatible with both product and coproduct ($S^{\cdot}$ carries a standard Hopf algebra structure; see 1.1.2).

(iv) Put $\mathscr{R} := G_0$. The product and coproduct on G define a Hopf algebra structure on $\mathscr{R}$. By (iii) we have an isomorphism $p^*S^{\cdot} \xrightarrow{\sim} \mathrm{gr}^{\cdot}\mathscr{R}$ of Hopf algebras. Hence the filtration $\mathscr{R}^{\cdot}$ on $\mathscr{R}$ coincides with the filtration by the powers of the augmentation ideal $I \subset R$. The product from (ii) defines a

$p^*\mathscr{R}$-module structure on G. One has $G^n = I^n G$ (since this is true on the zero fiber).

(v) For any unipotent sheaf $\mathscr{L}$ the sheaf $\mathscr{L}_0 = p_*\underline{\mathrm{Hom}}(G, \mathscr{L})$ has a canonical $\mathscr{R}$-module structure by (ii). The functor $\mathscr{S}\ell(X)^{\mathrm{un}} \to$ lisse $\mathscr{R}$-modules on B, $\mathscr{L} \mapsto \mathscr{L}_0$, is an equivalence of categories compatible with tensor products. The inverse functor is $\mathscr{M} \mapsto p^*\mathscr{M} \otimes_{p^*\mathscr{R}} G$ (the proof is obvious). Note that $\mathscr{R}$ (being commutative) acts, therefore, canonically on any $\mathscr{L} \in \mathscr{S}\ell(X)^{\mathrm{un}}$.

(vi) If $\varphi: X \to X'$ is an isogeny then one has a canonical morphism $\varphi_{\cdot G}: G_X \to \varphi^* G_{X'}$, uniquely defined by the property $\varphi_{\cdot G}(1) = \varphi^*(1)$ (see 1.2.8(b)). It is compatible with both product and coproduct structures. The corresponding morphism between the zero fibers $\varphi_{\cdot\mathscr{R}}: \mathscr{R}_X \to \mathscr{R}_{X'}$ is a morphism of Hopf algebras, and $\mathrm{gr}^{\cdot}(\varphi_{\cdot\mathscr{R}}): S^{\cdot}_X \to S^{\cdot}_{X'}$ coincides with $\mathrm{Sym}^{\cdot}(\varphi_{\cdot\mathscr{H}})$ where $\varphi_{\cdot\mathscr{H}}: \mathscr{H}_X \to \mathscr{H}_{X'}$ is the trace map for φ. In particular, $\varphi_{\cdot G}$ is an isomorphism if $\deg\varphi$ is invertible in F. As in 1.2.4 this provides an action of the group of translations $X^{(F)}_{\mathrm{tors}}$ on G; hence, we get a canonical isomorphism $G_x = \mathscr{R}$ for $x \in X^{(F)}_{\mathrm{tors}}$.

(vii) In the mixed situation the filtration $G^{\cdot}$ coincides with the weight filtration $G^i = W_{-i}G$ (since $\mathrm{gr}^i G$ is pure of weight $-i$).

(viii) Assume that $F \supset \mathbb{Q}$. Then $\nu^{\cdot}$ lifts uniquely to an isomorphism of Hopf algebras $\nu: F[[\mathscr{H}]] \xrightarrow{\sim} \mathscr{R}$. Also the morphism $G^{(n)} \to \mathrm{Sym}^n G^{(1)}$ that sends $1^{(n)}$ to $1^{(1)n}/n!$ is an isomorphism (since the filtration $A^i :=$ $\mathrm{Sym}^i G^{(1)1} \cdot \mathrm{Sym}^{n-i} G^{(1)}$ on $\mathrm{Sym}^n G^{(1)}$ satisfies 1.2.6(a)$'_n$). Therefore, $G =$ $\varprojlim \mathrm{Sym}^n G^{(1)} := \mathrm{Sym}\, G^{(1)}$.

1.3. The polylogarithm extension. Let $i_D: D \hookrightarrow X$ be a nonempty closed subscheme of X such that the projection $p_D := pi_D: D \to B$ is étale. Put $U = U_D := X \backslash D$; let $j: U \hookrightarrow X$, $p_U := pj: U \to B$ be the open embedding and the projection, respectively. For a sheaf $\mathscr{F}$ on X we will write $\mathscr{F}_U :=$ $j^*\mathscr{F}$.

1.3.1. Put $G[D] := p_{D*} i_D^* G$: this is a locally free sheaf of $\mathscr{R}$-modules on X of rank $\deg D/B$. We have a filtration $G^{\cdot}[D]$ on $G[D]$ such that

$$G[D] = \varprojlim G[D]/G^n[D], \qquad \mathrm{gr}^n G[D] = S^n \underset{F}{\otimes} F[D]$$

where $F[D] := p_* F_D$. Denote by $G[D]^\sharp \subset G[D]$ the kernel of the composition $\sigma: G[D] \twoheadrightarrow \mathrm{gr}^0 G[D] = F[D] \overset{tr}{\twoheadrightarrow} F_B$; so one has $G[D]/G[D]^\sharp = F_B$.

Let us compute $R^i p_{U*} G^{(n)}_U(1)$. Since p_U is affine, these sheaves vanish for $i \neq 0, 1$; one has $R^0 p_{U*} G^{(n)}_U(1) = R^0 p_* G^{(n)}(1) = S^n(1)$. For $i = 1$ we have the residue maps

$$\mathrm{Res}^{(n)}_D: R^1 p_{U*} G^{(n)}_U(1) \to G^{(n)}[D] := G[D]/G^{n+1}[D];$$

these maps are compatible with the projections $G^{(n+1)} \to G^{(n)}$.

1.3.2. LEMMA. *There exists a unique isomorphism of* $\mathcal{R}$*-modules*

$$R^1 p_{U^*} G_U^{(n)}(1) \xrightarrow{\sim} G[D]^\sharp / G^{n+2}[D]$$

such that the composition $R^1 p_{U^*} G_U^{(n+1)}(1) \to R^1 p_{U^*} G_U^{(n)}(1) \to G[D]^\sharp / G^{n+2}[D] \hookrightarrow G^{(n+1)}[D]$ *coincides with* $\mathrm{Res}_D^{(n+1)}$.

PROOF. The exact triangle

$$\cdots \to G^{(n)}(1) \to Rj_* G_U^{(n)}(1) \to i_{D^*} i_D^* G^{(n)}[-1] \to \cdots$$

provides, via Rp_*, the following exact sequence of sheaves on B:

$$0 \to R^1 p_* G^{(n)}(1) \to R^1 p_{U^*} G_U^{(n)}(1) \overset{\mathrm{Res}_D^{(n)}}{\to} G^{(n)}[D] \to R^2 p_* G^{(n)}(1) \to 0.$$

By 1.2.7 the left-most sheaf is S^{n+1} and the right-most sheaf is F_B. The epimorphism on the right coincides with σ; so we get the short exact sequence

$$0 \to S^{n+1} \to R^1 p_{U^*} G_U^{(n)}(1) \overset{\mathrm{Res}_D^{(n)}}{\to} G[D]^\sharp / G^{n+1}[D] \to 0.$$

Note that the morphism $R^1 p_{U^*} G_U^{(n+1)}(1) \to R^1 p_{U^*} G_U^{(n)}(1)$ is surjective (since $R^1 p_{U^*}$ is right exact). It kills the subsheaf $S^{n+2} \subset R^1 p_{U^*} G_U^{(n+1)}(1)$ (since, by 1.2.7, S^{n+2} coincides with the image of $R^1 p_*(p^* S^{n+1}) = R^1 p_*(\mathrm{Ker}(G^{(n+1)} \to G^{(n)}))$). Therefore, it factors through $\mathrm{Res}_D^{(n+1)}$, i.e., we get a canonical surjective morphism $G[D]^\sharp / G[D]^{n+2} \twoheadrightarrow R^1 p_{U^*} G_U^{(n)}$. Since these are sheaves of free F-modules of equal rank, this morphism is isomorphism. The isomorphism from 1.3.2 is the inverse one. □

Passing to the limit we get

1.3.3. COROLLARY. *One has* $R^0 p_{U^*} G_U(1) = 0$, *and the residue at* D *morphism* $\mathrm{Res}_D \colon R^1 p_{U^*} G_U \xrightarrow{\sim} G^\sharp[D]$ *is an isomorphism.*

The adjointness of p_U^*, Rp_{U^*} gives

1.3.4. COROLLARY. *One has* $\mathrm{Hom}(p_U^* \mathcal{L}, G_U(1)) = 0$ *for any prosheaf* $\mathcal{L} \in \mathcal{S}\ell(B)^\wedge$ *and the residue at* D *morphism provides an isomorphism*

$$\mathrm{Ext}^1(p_U^* \mathcal{L}, G_U(1)) \xrightarrow{\sim} \mathrm{Hom}(\mathcal{L}, G[D]^\sharp).$$

If $\mathcal{L}$ *is an* $\mathcal{R}$*-module, then the* Ext^1*-group in the category of* $p_U^* \mathcal{R}$*-modules coincides with* $\mathrm{Hom}_{\mathcal{R}}(\mathcal{L}, G[D]^\sharp)$.

1.3.5. We define the *large elliptic polylogarithm extension*

$$0 \to G_U(1) \to \widetilde{\mathcal{P}}^{(D)} \to P_U^* G[D]^\sharp \to 0$$

as the extension of $p_U^* \mathcal{R}$-modules that corresponds by 1.3.4 to $\mathrm{id}_{G[D]^\sharp}$. By 1.3.4 this extension has no automorphisms and satisfies a universal property. Since the residue morphisms are compatible with base change, $\widetilde{\mathcal{P}}^{(D)}$ is also compatible with the base change.

1.3.6. In 1.3.8 we shall give a convenient explicit construction of $\widetilde{\mathscr{P}}^{(D)}$. We need some preliminaries. For a lisse sheaf $\mathscr{F}$ on X denote by $\mathscr{F}_U^{\sim}$ the sheaf on $U \times_B U$ obtained from $F_U \boxtimes \mathscr{F}_U = \mathrm{pr}_2^* \mathscr{F}_U$ by restriction to the complement of the diagonal $U \times_B U \backslash \Delta(U)$ and the !-extension to $U \times U$. We have a short exact sequence $0 \to \mathscr{F}_U^{\sim} \to \mathrm{pr}_2^*\mathscr{F}_U \to \Delta_* \mathscr{F}_U \to 0$. Put $H_D(\mathscr{F}) := R^1 \mathrm{pr}_{1*} \mathscr{F}_U^{\sim}$. Since $R\mathrm{pr}_{1*}\mathrm{pr}_2^*\mathscr{F}_U = p_U^* Rp_{U*}\mathscr{F}_U$, $R\mathrm{pr}_{1*}\Delta_*\mathscr{F}_U = \mathscr{F}_U$, we get an exact sequence

$$0 \to p_U^* p_* \mathscr{F} \to \mathscr{F}_U \xrightarrow{\alpha} H_D(\mathscr{F}) \to p_U^* R^1 p_{U*} \mathscr{F}_U \to 0$$

of lisse sheaves on U.

1.3.7. LEMMA. (i) *The functor* $H_D: \mathscr{S}\ell(X) \to \mathscr{S}\ell(U)$ *is exact.*

(ii) *For any* $\mathscr{F} \in \mathscr{S}\ell(X)$, $L \in \mathscr{S}\ell(B)$ *one has* $H_D(\mathscr{F} \otimes p_U^* L) = H_D(\mathscr{F}) \otimes p_U^* L$.

(iii) *Let* λ *be a short exact sequence* $0 \to \mathscr{F}_U \to Q_\lambda \to p_U^* M \to 0$, $\mathscr{F} \in \mathscr{S}\ell(X)$, $M \in \mathscr{S}\ell(B)$. *There is a canonical morphism* $\alpha_\lambda: Q_\lambda \to H_D(\mathscr{F})$ *such that* $\alpha_\lambda|_{\mathscr{F}_U} = \alpha$, *and the quotient morphism* $p_U^* M = Q_\lambda / \mathscr{F}_U \to H_D(\mathscr{F})/\alpha(\mathscr{F}_U) = p_U^* R^1 p_{U_*} \mathscr{F}_U$ *is the* p_U*-inverse image of the boundary map* $\partial_\lambda: M = p_{U_*} p_U^* M \to R^1 p_{U_*} \mathscr{F}_U$.

(iv) *Assume that* $p_* \mathscr{F} = 0$. *Then the sequence*

$$0 \to \mathscr{F}_U \to H_D(\mathscr{F}) \to p_U^* R^1 p_{U_*} \mathscr{F}_U \to 0$$

is exact, and the above α_λ *is the unique morphism* $Q_\lambda \to H_D(\mathscr{F})$ *such that* $\alpha_\lambda|_{\mathscr{F}_U} = \alpha$.

PROOF. (i) Clear, since $\mathscr{F}_U^{\sim}$ depends on $\mathscr{F}$ in an exact way and $R^i \mathrm{pr}_{1*} \mathscr{F}_U^{\sim} = 0$ for $i \neq 1$.

(ii) Clear.

(iii) For any $Q \in \mathscr{S}\ell(U)$ one has $\mathrm{Hom}(Q, H_D(\mathscr{F})) = \mathrm{Ext}^1(\mathrm{pr}_1^* Q, \mathscr{F}_U^{\sim})$ by adjointness. We may identify this group with the group of isomorphism classes of pairs (ν, s) where ν is an extension $0 \to \mathrm{pr}_2^* \mathscr{F}_U \to L_\nu \to \mathrm{pr}_1^* Q \to 0$ and s is a trivialization of ν over Δ; i.e., $s \in \mathrm{Hom}(Q, \Delta^* L_\nu)$ is a splitting of the extension $0 \to \mathscr{F}_U \to \Delta^* L_\nu \to Q \to 0$. Now we define the desired morphism α_λ as the pair (ν_λ, s_λ) where ν_λ is the pull-back by the projection $\mathrm{pr}_1^* Q \to \mathrm{pr}_1^* p_U^* M = \mathrm{pr}_2^* p_U^* M$ of the extension $\mathrm{pr}_2^*(\lambda)$ and s is the obvious splitting (note that $\Delta^* \mathrm{pr}_2^* \lambda = \lambda$).

(iv) Clear. □

Passing to limits we get the functor $H_D: \mathscr{S}\ell(X)^\wedge \to \mathscr{S}\ell(U)^\wedge$; the previous lemma remains valid.

1.3.8. LEMMA. *The morphism* $\alpha_\lambda: \widetilde{\mathscr{P}}^{(D)} \to H_D(G(1))$, *which corresponds to the short exact sequence* λ *from* 1.3.5, *is an isomorphism of* $p_U^* \mathscr{R}$*-modules.*

PROOF. Clear, since both $\widetilde{\mathscr{P}}^{(D)}$, $H_D(G(1))$ have the same universal property (see 1.3.4, 1.3.7(iii)). □

Let us describe the $\mathscr{R}$-module structure on $\widetilde{\mathscr{P}}^{(D)} = H_D(G(1))$. Consider the decreasing filtration $\widetilde{\mathscr{P}}^{(D)i} := H_D(G^{i-1}(1))$ on $\widetilde{\mathscr{P}}^{(D)}$; one has $\widetilde{\mathscr{P}}^{(D)} = \varprojlim \widetilde{\mathscr{P}}^{(D)}/\widetilde{\mathscr{P}}^{(D)i}$. Since $I\widetilde{\mathscr{P}}^{(D)i} \subset \widetilde{\mathscr{P}}^{(D)i+1}$, the associated graded sheaf $\operatorname{gr}^{\cdot}\widetilde{\mathscr{P}}^{(D)}$ is a sheaf of $S^{\cdot}$-modules.

1.3.9. LEMMA. (i) *One has a canonical isomorphism of $S^{\cdot}$-modules* $\operatorname{gr}^{\cdot}\widetilde{\mathscr{P}}^{(D)} = p_U^*(S^{\cdot-1} \otimes R^1 p_{U*}F(1))$.

(ii) *Assume that* $\deg D/B = 1$. *Then* $\operatorname{gr}^{\cdot}\widetilde{\mathscr{P}}^{(D)} = p_U^*(S^{\cdot-1} \otimes \mathscr{H})$ *and, if we deal with mixed sheaves, our filtration coincides with weight filtration:* $\widetilde{\mathscr{P}}^{(D)i} = W_{-i}\widetilde{\mathscr{P}}^{(D)}$.

PROOF. (i) follows from 1.3.7(i), (ii) since $H_D(F_X(1)) = p_U^* R^1 p_{U*}F(1)$; (ii) follows from (i). □

Note that 1.3.9 implies that the fibers of $\widetilde{\mathscr{P}}^{(D)}$ are free $\mathscr{R}$-modules of rank $\deg D/B + 1$.

1.3.10. If $D' \supset D$ is another divisor étale over B, $U' := X\backslash D'$, then there is a unique morphism $i_{DD'}: \widetilde{\mathscr{P}}^{(D)}_{U'} \to \widetilde{\mathscr{P}}^{(D')}$ that coincides with the identity map on the subsheaves $G_{U'}(1)$. The quotient morphism $G[D]^\sharp \to G[D']^\sharp$ is an obvious embedding; hence $i_{DD'}$ is injective.

More generally, let $\varphi: X \to X'$ be an isogeny of elliptic curves, $D \subset X$, $D' \subset X'$ be divisors étale over B such that $\varphi(D) \subset D'$. Put $V := X\backslash\varphi^{-1}(D') = \varphi^{-1}(U')$. We have the morphism $\varphi_{\cdot G}: G_X \to \varphi^* G_{X'}$ (see 1.2.4, 1.2.10(vi)); it defines the trace morphism $\varphi_{\cdot D}: G_X[D] \to G_{X'}[D']$ which sends $G_X[D]^\sharp$ to $G_{X'}[D]^\sharp$. By adjointness $\widetilde{\mathscr{P}}^{(D)}$ defines the extension $0 \to \varphi_* G_V(1) \to \widetilde{\mathscr{P}}^{(D)}_\varphi \to p_{U'}^* G_X[D]^\sharp \to 0$ of lisse prosheaves on U'—the pull-back of $\varphi_*\widetilde{\mathscr{P}}^{(D)}_V$ by $p_{U'}^* G_X[D]^\sharp \hookrightarrow \varphi_*\varphi^* p_{U'}^* G_X[D]^\sharp = \varphi_* p_U^* G_X[D]^\sharp$.

1.3.11. LEMMA. *The morphism* $\operatorname{tr}\varphi_{\cdot G}: \varphi_* G_X(1) \to G_{X'}(1)$ *extends in a unique manner to a morphism* $\operatorname{tr}\varphi_{\cdot\widetilde{\mathscr{P}}}: \widetilde{\mathscr{P}}^{(D)}_\varphi \to \widetilde{\mathscr{P}}^{(D')}$. *The quotient morphism* $p_{U'}^* G_X[D]^\sharp \to p_{U'}^* G_{X'}[D]^\sharp$ *is* $p_{U'}^*(\varphi_{\cdot D})$.

PROOF. The first claim follows from the universal property of $\widetilde{\mathscr{P}}^{(D')}$; the second is the compatibility of the trace map with residues. □

1.3.12. Assume now that D consists of points of finite order invertible in F. According to 1.2.10(vi), 1.2.4 we have a canonical isomorphism $G[D] = \mathscr{R}[D] = \mathscr{R} \otimes F[D]$ where $F[D] := p_{D*}F_D$. Therefore, $G[D]^\sharp = \mathscr{R}[D]^\sharp := \ker(\sigma: \mathscr{R}[D] \to F_B)$ where σ is the composition $\mathscr{R}[D] \xrightarrow{\operatorname{tr}} \mathscr{R} \to \mathscr{R}/I = F_B$. Note that $\mathscr{R}[D]^\sharp \supset F[D]^0 := \ker(\operatorname{tr}: F[D] \to F_B)$, so we get an extension

$$0 \to G_U(1) \to \mathscr{P}^{(D)} \to p_U^* F[D]^0 \to 0$$

—the pull-back of $\widetilde{\mathscr{P}}^{(D)}$ by $F[D]^0 \hookrightarrow G[D]^0$.

1.3.13. Assume that $D = \{0\}$ and $F \supset \mathbb{Q}$. Then $G[D] = \mathscr{R}$, $G[D]^\sharp = I \subset \mathscr{R}$. Recall that we have a canonical isomorphism $\mathscr{R} = F[[\mathscr{H}]]$; in particular, we have a canonical embedding $\mathscr{H} \hookrightarrow I$. Let

$$0 \to G_U(1) \to \mathscr{P} \to p_U^* \mathscr{H} \to 0$$

be the pull-back of $\widetilde{\mathscr{P}} := \widetilde{\mathscr{P}}^{(\{0\})}$ by $\mathscr{H} \hookrightarrow I$.

We shall call the above $\mathscr{P}$, $\mathscr{P}^{(D)}$ the *small elliptic polylogarithm extensions*. They are compatible with base change and isogenies. For example, consider the $\mathscr{P}$-extension. Let $\varphi: X \to X'$ be an isogeny. It defines the morphisms $\varphi_{\cdot\mathscr{H}}: \mathscr{H}_X \to \mathscr{H}_{X'}$, $\varphi^{\cdot}_{\mathscr{H}}: \mathscr{H}_{X'} \to \mathscr{H}_X$ such that $\varphi_{\cdot\mathscr{H}}\varphi^{\cdot}_{\mathscr{H}} = \deg\varphi \cdot \mathrm{id}_{\mathscr{H}_{X'}}$. Let

$$0 \to G_{U'}(1) \to \mathscr{P}_\varphi \to p_{U'}^* \mathscr{H}_{X'} \to 0$$

be the $\operatorname{tr}\varphi_{\cdot}$-push-forward and $\varphi^{\cdot}_{\mathscr{H}}$-pull-back of the extension $\varphi_* \mathscr{P}_X$. We see that this extension is (canonically) isomorphic to the $\deg\varphi$-multiple of $\mathscr{P}_{X'}$.

Let us write down the universal property of $\mathscr{P}$. Consider the morphisms $\operatorname{End}\mathscr{H} \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \operatorname{Ext}^1(p_U^*\mathscr{H}, G_U(1))$, where α sends $\varphi \in \operatorname{End}\mathscr{H}$ to the pull-back of $\mathscr{P}$ by φ, and β is the composition $\operatorname{Ext}^1(p_U^*\mathscr{H}, G_U(1)) \to \operatorname{Ext}^1(p_U^*\mathscr{H}, F_U(1)) = \operatorname{Hom}(\mathscr{H}, Rp_{U*}F_U(1)[1]) \to \operatorname{Hom}(\mathscr{H}, R^1p_{U*}F(1)) = \operatorname{End}\mathscr{H}$ (the first arrow comes from the projection $G \to G^{(0)} = F_X$). Clearly $\beta\alpha = \mathrm{id}_{\operatorname{End}\mathscr{H}}$.

1.3.14. LEMMA. *If either we deal with mixed sheaves or our family $p: X \to B$ is not isotrivial (i.e., not étale locally constant) then the above α, β are mutually inverse isomorphisms.*

PROOF. By 1.3.4 $\operatorname{Ext}^1(p_U^*\mathscr{H}, G_U(1)) = \operatorname{Hom}(\mathscr{H}, I)$, and 1.3.14 follows since our conditions imply that $\operatorname{Hom}(\mathscr{H}, S^i) = 0$ for $i = 1$. □

1.3.15. LEMMA. *The sheaf $\mathscr{P}$ carries a unique F-Lie algebra structure such that $G_U(1) \subset \mathscr{P}$ is an abelian ideal, and the adjoint action of $\mathscr{P}/G_U(1) = p_U^*\mathscr{H}$ on it coincides with the $\mathscr{H}$-action defined by the canonical $F[[\mathscr{H}]] = \mathscr{R}$-module structure on $G_U(1)$.*

PROOF. Unicity. Since $\mathscr{H}$ acts on $G_U(1)$ faithfully, we see that $\mathscr{P}/G_U(1)$ must be an abelian quotient. Therefore, the two possible brackets on $\mathscr{P}$ differ by a morphism $\Lambda^2\mathscr{P} \to \mathscr{P}$ that factors through a morphism $p_U^*\Lambda^2\mathscr{H} \to G_U(1)$. By 1.2 one has $\operatorname{Hom}(p_U^*\Lambda^2\mathscr{H}, G_U(1)) = 0$.

Existence. Denote by $\mathscr{K} \subset \Lambda^2\mathscr{P}$ the kernel of the projection $\Lambda^2\mathscr{P} \twoheadrightarrow \Lambda^2 p_U^*\mathscr{H} = F_U(1)$ so that we have a short exact sequence

$$0 \to \Lambda^2(G_U(1)) \to \mathscr{K} \to p_U^*\mathscr{H} \otimes G_U(1) \to 0.$$

Our conditions define the restriction to $\mathscr{K}$ of the desired bracket $[\ ,\ ]: \Lambda^2\mathscr{P} \to G_U(1) \subset \mathscr{P}$—it coincides with the composition $\theta: \mathscr{K} \to p_U^*\mathscr{H} \otimes G_U(1) \subset p_U^*\mathscr{R} \otimes G_U(1) \to G_U(1)$. The definition of $[\ ,\ ]$ is equivalent, therefore, to a splitting of the short exact sequence $0 \to G_U(1) \to \mathscr{F} \to F_U(1) \to 0$ which is the θ-push-forward of $0 \to \mathscr{H} \to \Lambda^2\mathscr{P} \to F_U(1) \to 0$. We may assume that our family X/B is not isotrivial (since $\mathscr{P}$ is compatible with base change). Then $\mathrm{Ext}^1(F_U(1), G_U(1)) = \mathrm{Hom}(F_B, \prod_{i\geq 1} S^i(-1)) = 0$ (see 1.3.4); therefore, the splitting exists. It is unique since $\mathrm{Hom}(F_U(1), G_U(1)) = 0$. We get the bracket; the verification of the Jacobi identity is left to the reader. □

1.4. A quotient of the fundamental Lie algebra. In this section we shall give another construction of $\mathscr{P}$ which provides, in particular, an explanation of 1.3.15; the results of 1.4 will not be used in other sections. We assume that $F \supset \mathbb{Q}$ is a field.

1.4.1. Consider the category $\mathscr{S}\ell(U)^{\mathrm{un}}$ of unipotent (with respect to p_U) lisse sheaves on U. By definition, this is the Serre subcategory of $\mathscr{S}\ell(U)$ generated by $p_U^*(\mathscr{S}\ell(B))$; its objects are successive extensions of sheaves from $p_U^*(\mathscr{S}\ell(B))$. This is a Tannakian subcategory of $\mathscr{S}\ell(U)$. As usual, we put $\mathscr{S}\ell(U)^{\mathrm{un}\wedge} := \varprojlim \mathscr{S}\ell(U)^{\mathrm{un}}$—this is an abelian category with tensor product.

Let L be the p_U-relative fundamental Lie algebra for $\mathscr{S}\ell(U)^{\mathrm{un}}$. This is a pronilpotent Lie algebra object of $\mathscr{S}\ell(U)^{\mathrm{un}\wedge}$ that acts canonically on any $\mathscr{F} \in \mathscr{S}\ell(U)^{\mathrm{un}}$; this action is compatible with tensor product and duality. The Lie algebra L satisfies the following properties:

(i) Let $L = L^1 \supset L^2 \supset \cdots$ be the lower central series filtration. Then $\mathrm{gr}^{\cdot} L$ is the free Lie algebra generated by $\mathrm{gr}^1 L = p_U^*\mathscr{H}$. In particular, $\mathrm{gr}^2 L = \Lambda^2 p_U^* \mathscr{H} = F_U(1)$ (see 1.1.2).

(ii) Let $\mathscr{F} \in \mathscr{S}\ell(U)^{\mathrm{un}}$ be a sheaf. Then $\mathscr{F} \in p_U^*\mathscr{S}\ell(B)$ iff L acts on $\mathscr{F}$ trivially. Also $\mathscr{F}$ extends (as a lisse sheaf) to X iff $L^2 \subset L$ acts on $\mathscr{F}$ trivially; in this case the $p_U^*\mathscr{H} = L/L^2$-action on $\mathscr{F}$ coincides with those defined in 1.2.10(v), (viii).

Let $T := 0^*\mathscr{T}_{X/B}$ be the vertical tangent line at 0, and let $\dot{T} \subset T$ be the complement to the zero section (so $\dot{T}$ is an $\mathscr{O}^*$-torsor over B). We have the specialization at 0 functor $\mathrm{Sp}_0 \colon \mathscr{S}\ell(U)^{\mathrm{un}} \to \mathscr{S}\ell(\dot{T})$. This is an exact faithful tensor functor; for $\mathscr{K} \in \mathscr{S}\ell(X)$ one has $\mathrm{Sp}_0 \mathscr{K}_U = p_{\dot{T}}^* \mathscr{K}_0$.

(iii) We have a canonical morphism $\alpha \colon F_{\dot{T}}(1) \to \mathrm{Sp}_0 L^2$ such that the composition $F_{\dot{T}}(1) \xrightarrow{\alpha} \mathrm{Sp}_0 L^2 \to \mathrm{Sp}_0 L^2/L^3 \underset{(\mathrm{i})}{=} F_{\dot{T}}(1)$ is the identity.

(iv) The action of L on itself, as on an object of $\mathscr{S}\ell(U)^{\mathrm{un}\wedge}$, coincides with the adjoint action.

1.4.2. Remarks. (i) If we are over $\mathbb{C}$ and deal with the classical topology, then L is the Lie algebra of the Mal′cev completion of the sheaf formed by the fundamental groups of the fibers. Precisely, on U we have a local system $\sqcap$ of groups with fibers $\sqcap_x = \pi_1(U_{p(x)}, x)$. Let $F[\sqcap]^\wedge$ be the completion of the group algebra $F[\sqcap]$ with respect to powers of the augmentation ideal. Then $F[\sqcap]^\wedge \in \mathscr{S}\ell(U)^{\mathrm{un}\wedge}$ is a Hopf algebra, and $L \subset F[\sqcap]^\wedge$ is the subsheaf of Lie algebra type elements for the coproduct. The canonical action of L

on a sheaf $\mathcal{F} \in \mathcal{S}\ell(U)^{\mathrm{un}}$ comes from the monodromy action of $\sqcap$, or $F[\sqcap]$, on $\mathcal{F}$. The map α from 1.4.1(iii) assigns to a generator of $F(1)$ the logarithm of a small loop around 0. A similar construction is possible in the étale situation. A general construction parallel to 1.2.6–1.2.8 is fairly straightforward; we skip the details.

(ii) In the mixed situation the filtration $L^{\cdot}$ coincides with the weight filtration: one has $L^i = W_{-i}L$ (since, by 1.4.1(i), $\mathrm{gr}^i L$ is pure of weight $-i$).

(iii) The bracket $[\ ,\ ]\colon \Lambda^2 L \to L$ is strictly compatible with the filtration $L^{\cdot}$ (this follows from 1.4.1(i); in the mixed situation this is simply a usual property of weight filtrations).

Consider the quotient algebra $\overline{L} := L/[L^2, L^2]$; let $\overline{L}^i = L^i/L^i \cap [L^2, L^2]$ be its lower central series filtration. The adjoint action of $\overline{L}$ on $\overline{L}^2$ factors through the $p_U^*\mathcal{H}$-action. By 1.4.1(ii), (iv) $\overline{L}^2$ extends as a lisse sheaf to X; denote this sheaf on X by $\mathcal{J}$. The morphism α from 1.4.1(iii) defines a section $\overline{\alpha} \in \mathrm{Hom}(F_B(1), \mathcal{J}_0)$; by 1.2.8(b) we get a morphism $\varphi\colon G(1) \to \mathcal{J}$ that sends 1 to $\overline{\alpha}$.

1.4.3. Proposition. (i) *This φ is an isomorphism.*

(ii) *It extends canonically to an isomorphism of extensions*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & G_U(1) & \longrightarrow & \mathcal{P} & \longrightarrow & p_U^*\mathcal{H} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\wr} & & \downarrow{\scriptstyle\mathrm{id}_{\mathcal{H}}} & & \\ 0 & \longrightarrow & \mathcal{J}_U & \longrightarrow & \overline{L} & \longrightarrow & p_U^*\mathcal{H} & \longrightarrow & 0 \end{array}$$

compatible with the Lie brackets (*see* 1.3.15).

Proof. (i) By 1.4.1(ii), (iv) the canonical action of $p_X^*\mathcal{H}$ on $\mathcal{J}_U$ comes from the adjoint action of $\overline{L}$ on $\overline{L}^2$. By 1.2.1(a)$'$ our statement will follow if we show that $\mathrm{gr}^{\geq 2}\overline{L}$ is a free $S^{\cdot}$-module generated by $\mathrm{gr}^2\overline{L} = F(1)$.

Denote by ϕ the free Lie algebra generated by $\mathcal{H}$. It carries an obvious grading with $\phi^1 = \mathcal{H}$. By 1.4.1(i) we need to show that $\phi^{\geq 2}/[\phi^{\geq 2}, \phi^{\geq 2}]$ is a free $S^{\cdot}$-module generated by $\phi^2 = F(1)$. Clearly $F(1)$ generates our $S^{\cdot}$-module. To show that there are no relations, it suffices to construct a morphism of graded $S^{\cdot}$-modules $\gamma\colon \phi^{\geq 2}/[\phi^{\geq 2}, \phi^{\geq 2}] \to S^{\cdot}(1)$ which is $\mathrm{id}_{F(1)}$ on the degree 2 component.

The sheaf $M := \mathcal{H} \oplus S^{\cdot}(1)$ carries a bracket $[\ ,\ ]\colon \Lambda^2 M \to M$ that is zero on $S^{\cdot}(1)$, coincides with the intersection pairing $\Lambda^2\mathcal{H} \to F(1) = S^{\circ}(1)$ on $\mathcal{H}$, and coincides with the multiplication map $\mathcal{H} \otimes S^{\cdot}(1) \to S^{\cdot+1}(1)$ on $\mathcal{H} \otimes S^{\cdot}(1) \subset \Lambda^2 M$. One immediately checks the Jacobi identity for $[\ ,\ ]$; so we have the morphism of Lie algebras $\phi \to M$ that induces the identity on $\mathcal{H}$. It kills $[\phi^{\geq 2}, \phi^{\geq 2}]$; hence, we get the desired map $\gamma\colon \phi^{\geq 2}/[\phi^{\geq 2}, \phi^{\geq 2}] \to S^{\cdot}(1)$. $\square$

(ii) According to 1.3.14, 1.4.1(iv) it suffices to show that the bracket on the quotient L/L^3 reduces to the intersection pairing $\Lambda^2\mathscr{H} = \Lambda^2 L/L^2 \to L^2/L^3 = F(1)$, which is obvious. The compatibility with the brackets follows from the unicity of the bracket on $\mathscr{P}$ (see 1.3.15). □

1.5. A connection with the classical polylogarithm. We shall show that the elliptic polylogarithm sheaf $\mathscr{P}$ tends to the classical one when our elliptic curve degenerates. In this section we deal with mixed sheaves.

1.5.1. Recall a definition of the classical polylogarithm sheaf (see, e.g., **[BD]**). Put $X_\infty := \mathbb{P}^1_k\backslash\{0, \infty\}$, $U_\infty := \mathbb{P}^1_k\backslash\{0, 1, \infty\}$ (below we shall consider X_∞ as a degenerate fiber of a family of elliptic curves; this explains the notation). Consider a mixed Tate prosheaf G_{X_∞} of weight ≤ 0 such that $\dim \mathrm{gr}^W_{-2i} G_{X_\infty} = 1$ for $i \geq 0$, the fiber $G_{X_\infty, 1}$ splits, and each extension

$$0 \to \mathrm{gr}^W_{-2i-2} G_{X_\infty} \to W_{-2i}/W_{-2i-4} G_{X_\infty} \to \mathrm{gr}^W_{-2i} G_{X_\infty} \to 0$$

for $i \geq 0$ is nontrivial. Such G_{X_∞} exists and is unique up to isomorphism; one has $\mathrm{End}\, G_{X_\infty} = F$. We can make the choice of G_{X_∞} canonical by fixing an isomorphism $\mathrm{gr}^W_0 G_{X_\infty} = G_{X_\infty}/W_{-2} G_{X_\infty} \xrightarrow{\sim} F_{X_\infty}$. Note that $W_{-2} G_{X_\infty}$ is isomorphic to $G_{X_\infty}(1)$.

Consider the morphism

$$(\mathrm{Res}_0, \mathrm{Res}_1)\colon \mathrm{Ext}^1(F_{U_\infty}, G_{U_\infty}(1)) \to F \oplus F$$

that assigns to an extension $0 \to G_{U_\infty}(1) \to \mathscr{F} \to F_{U_\infty} \to 0$ the images of the class of $0 \to F_{U_\infty}(1) \to \mathscr{F}/W_{-4}\mathscr{F} \to F_{U_\infty} \to 0$ by the residue maps $\mathrm{Res}_0, \mathrm{Res}_1\colon H^1(U_\infty, F(1)) \to F$. According to **[BD]** the map $(\mathrm{Res}_0, \mathrm{Res}_1)$ is an isomorphism (this is an analog of the statement 1.3.14). For $(\lambda_0, \lambda_1) \in F \oplus F$ denote by $\mathscr{P}^\infty_{\lambda_0, \lambda_1}$ the extension that corresponds to (λ_0, λ_1). Then $\mathscr{P}^\infty_{0,1}$ is the classical polylogarithm sheaf, and $\mathscr{P}^\infty_{1,0} = G_{U_\infty}$. Let

$$0 \to G_{U_\infty}(1) \to \mathscr{P}^\infty \to F \oplus F \to 0$$

be the corresponding universal extension (so $\mathscr{P}^\infty_{\lambda_0, \lambda_1}$ is the pull-back of $\mathscr{P}^\infty$ by $(\lambda_0, \lambda_1)\colon F \to F \oplus F$).

1.5.2. Assume that our base B is a smooth curve. Let $\overline{B} \supset B$ be a smooth compactification of B and $\infty \in (\overline{B}\backslash B)(k)$ be a point such that our family X has multiplicative reduction at ∞. Let X_∞ be the connected component of the fiber at ∞ of the Neron model $p_{\overline{X}}\colon \overline{X} \to \overline{B}$. We may identify X_∞ with $\mathbb{G}_m = \mathbb{P}^1\backslash\{0, \infty\}$ as algebraic groups. The neutral point of X_∞ is 1; therefore, $U_\infty = X_\infty\backslash\{1\}$. Let $X_\infty \overset{i}{\hookrightarrow} \overline{X} \overset{k}{\hookleftarrow} X$ be the closed and open embeddings.

1.5.3. Lemma. *One has an isomorphism*

$$\mathscr{P}^\infty \xrightarrow{\sim} i^* k_* \mathscr{P}.$$

PROOF. Denote by τ the tangent line to $\overline{B}$ at ∞; put $\dot{t} := \tau\backslash\{0\}$. The differential of $p_{\overline{X}}$ identifies the normal bundle of X_∞ with the constant bundle with fiber τ. Consider the "specialization to X_∞" functors

$$\mathrm{Sp}_\infty : \mathscr{S}\ell(X)^{\mathrm{un}} \to \mathscr{S}\ell(X_\infty \times \dot{t}), \ \mathscr{S}\ell(U)^{\mathrm{un}} \to \mathscr{S}\ell(U_\infty \times \dot{t}).$$

The specialization functors for X and U are compatible with the restriction of sheaves from X to U. They are exact faithful tensor functors compatible with p-pull-back: for $\mathscr{L} \in \mathscr{S}\ell(B)$ one has $\mathrm{Sp}_\infty p^*\mathscr{L} = p_\infty^* \mathrm{Sp}_\infty \mathscr{L}$ where $p_\infty : X_\infty \times \dot{t} \to \dot{t}$ is the projection and $\mathrm{Sp}_\infty \mathscr{L} \in \mathscr{S}\ell(\dot{t})$ is the specialization of $\mathscr{L}$. For any $\mathscr{F} \in \mathscr{S}\ell(U)^{\mathrm{un}}$ we have a canonical operator $N_{\mathscr{F}} : \mathrm{Sp}_\infty \mathscr{F} \to \mathrm{Sp}_\infty \mathscr{F}(-1)$ (the logarithm of the unipotent part of the monodromy along $\dot{t}$). We have a canonical embedding $\mathrm{pr}_1^* i^* k_* \mathscr{F} \hookrightarrow \ker N_{\mathscr{F}}$; if the monodromy action on $\mathrm{Sp}_\infty \mathscr{F}$ has all eigenvalues equal to one, then this embedding is an isomorphism.

Let us return to our situation. We have a short exact sequence

$$0 \to F(1)_{\dot{t}} \to \mathrm{Sp}_\infty \mathscr{H} \to F_{\dot{t}} \to 0.$$

The operator $N_{\mathscr{H}}$ is the composition $\mathrm{Sp}_\infty \mathscr{H} \twoheadrightarrow F_{\dot{t}} \xrightarrow{\sim} F_{\dot{t}} \hookrightarrow \mathrm{Sp}_\infty \mathscr{H}(-1)$. Therefore, for any $i \geq 0$, $\mathrm{Sp}_\infty S^i$ is a mixed Tate sheaf on $\dot{t}$ with $\mathrm{gr}^W_{-2j} \mathrm{Sp}_\infty S^i$ equal to $F(j)$ for $j = 0, \ldots, i$ and to zero otherwise; N_{S^i} acts as a Jordan block: it induces isomorphisms $\mathrm{gr}^W_{-2j} \mathrm{Sp}_\infty S^i \xrightarrow{\sim} (\mathrm{gr}^W_{-2j-2} \mathrm{Sp}_\infty S^i)(-1)$ for $j = 0, \ldots, i-1$.

Put $\mathscr{M} := \mathrm{Sp}_\infty \mathscr{P}$, $\mathscr{M}^i := \mathrm{Sp}_\infty W_{-i}\mathscr{P}$. We have $\mathscr{M} = \mathscr{M}^1 \supset \mathscr{M}^2 \supset \cdots$, $\mathscr{M} = \varprojlim \mathscr{M}/\mathscr{M}^i$, $\mathscr{M}/\mathscr{M}^2 = p_\infty^* \mathrm{Sp}_\infty \mathscr{H}$, $\mathscr{M}^i/\mathscr{M}^{i+1} = p_\infty^* \mathrm{Sp}_\infty S^{i-2}(1)$ for $i \geq 2$. We see that $\mathscr{M}$ is a mixed Tate prosheaf. The eigenvalues of the monodromy action on $\mathscr{M}$ are trivial, so $\mathrm{pr}_1^* i^* k_* \mathscr{P} \xrightarrow{\sim} \mathscr{M}^{N_{\mathscr{P}}} := \ker N_{\mathscr{P}}$.

Let us show that the obvious embedding $\mathrm{gr}^\cdot(\mathscr{M}^{N_{\mathscr{P}}}) \hookrightarrow (\mathrm{gr}^\cdot \mathscr{M})^{\mathrm{gr}^\cdot N_{\mathscr{P}}}$ is an isomorphism. Note that the projection $\mathscr{M}/W_{-2}\mathscr{M} \to \mathrm{Sp}_\infty \mathscr{H}/W_{-2}\mathrm{Sp}_\infty \mathscr{H} = F$ is an isomorphism; hence, $\mathscr{M}^{N_{\mathscr{P}}} = (W_{-2}\mathscr{M})^{N_{\mathscr{P}}}$. One has $N_{\mathscr{P}}(W_{-2}\mathscr{M}) \subset W_{-2}(\mathscr{M}(-1)) = W_{-2}(\mathscr{M}^3(-1))$ by weight reasons, and the above discussion shows that $N_{S^i(1)}(\mathrm{Sp}_\infty S^i(1)) = W_{-2}\mathrm{Sp}_\infty S^i$. Therefore, $N_{\mathscr{P}}(W_{-2}\mathscr{M}) = W_{-2}(\mathscr{M}(-1))$. The same holds on any quotient $W_{-2}\mathscr{M}/\mathscr{M}^i$; hence, $\dim(\mathscr{M}/\mathscr{M}^i)^{N_{\mathscr{P}}} = \dim(W_{-2}\mathscr{M}/\mathscr{M}^i)^{N_{\mathscr{P}}} = \dim(W_{-2}\mathscr{M}/N_{\mathscr{P}}(W_{-2}\mathscr{M}) + \mathscr{M}^i) = \dim \mathrm{gr}^W_{-2}\mathscr{M}/\mathscr{M}^i$. This proves our claim $\mathrm{gr}^\cdot \mathscr{M}^{N_{\mathscr{P}}} = (\mathrm{gr}^\cdot \mathscr{M})^{\mathrm{gr}^\cdot N_{\mathscr{P}}}$.

Look at the sheaf $i^* k_* G_X$. We already know that $\mathrm{pr}_1^* i^* k_* G_X = \ker N_{G_X}$, and our results about $\mathscr{P}$ show that $i^* k_* G_X$ is a mixed Tate prosheaf on X_∞ of weights ≤ 0 and $\dim \mathrm{gr}^W_{-2i} i^* k_* G_X = 1$ for $i \geq 0$. It splits over 1 since the restriction of $\mathrm{Sp}_\infty G_X$ to $\{1\} \times \dot{t}$ coincides with $\mathrm{Sp}_\infty G_0 = \mathrm{Sp}_\infty \prod S^i$. A simple computation of the monodromy shows that it acts on $i^* k_*(G_X/W_{-2}G_X)$ nontrivially. By 1.2.10(viii) the sheaf $i^* k_* G_X$ satisfies the conditions from 1.5.1; hence, we may identify it with G_{X_∞}. The above considerations show that we have a short exact sequence

$$0 \to G_{U_\infty} \to i^* k_* \mathscr{P} \to F_{U_\infty} \to 0.$$

A computation of local monodromy around 1 shows (by 1.3.5, 1.3.13) that the class of this extension has nontrivial residue at 1, which implies 1.5.2. □

1.6. Some remarks on the case of a modular family. In this section we deal with some specific properties that hold when our family $p_X \colon X \to B$ is modular. So let us assume that the morphism $B \to$ (moduli of elliptic curves) is a finite étale covering which is a quotient of a moduli space of curves with some level structure.

We shall deal with mixed sheaves; for a mixed sheaf $\mathscr{F}$ on a scheme Y we put $H^i(Y, \mathscr{F}) := R^i \pi_{Y*} \mathscr{F} \in \mathscr{S}\ell(k)_{\text{mixed}}$, where $\pi_Y \colon Y \to \operatorname{Spec} k$ is the structure morphism, and $H^i_{\text{abs}}(Y, \mathscr{F}) := \operatorname{Ext}^i(F_Y(0), \mathscr{F})$. For a mixed sheaf $\mathscr{K}$ over $\operatorname{Spec} k$ we shall simply write $H^i_{\text{abs}}(\mathscr{K}) := H^i_{\text{abs}}(\operatorname{Spec} k, \mathscr{K})$. We have the Leray spectral sequence $E_2^{p,q} = H^p_{\text{abs}}(H^q(Y, \mathscr{F}))$ converging to $H^{\cdot}_{\text{abs}}(Y, \mathscr{F})$.

The following lemma is well known.

1.6.1. Lemma. *The mixed sheaf $H^1(B, S^i(1))$ on* $\operatorname{Spec} k$ *splits into a direct sum of pure components of weights $-1-i$ and 0. The weight $-1-i$ part (the parabolic component) does not contain any subobject isomorphic to a Tate twist of an Artin sheaf. The weight 0 part $\mathscr{E}is^{i+2}(B)$ (the Eisenstein component) is an Artin sheaf.*

Here "Artin sheaf on $\operatorname{Spec} k$" = "a sheaf isomorphic to $F(0)^n$ on some étale covering of $\operatorname{Spec} k$" = "a representation of the Galois group of k that factors through a finite quotient". For such a sheaf $\mathscr{E}$, $H^0_{\text{abs}}(\mathscr{E})$ is the subspace of Galois invariants in the corresponding representation.

1.6.2. Corollary. *For $i > 0$ one has $H^1_{\text{abs}}(B, S^i(1)) = H^0_{\text{abs}}(\mathscr{E}is^{i+2}(B))$ and $H^1_{\text{abs}}(B, S^i) = 0$.*

Proof. One has $H^j(B, S^i(1)) = 0$ for $j \neq 1$ (recall that B is affine). Hence 1.6.2 follows from 1.6.1. □

1.6.3. Corollary. (i) *Let $G' = G'^0 \supset G'^1 \supset \cdots$, $G' = \varprojlim G'/G'^i$, be a filtered prosheaf on X equipped with isomorphisms $G'^i/G'^{i+1} = p_X^* S^i$ such that the differentials $S^i = R^0 p_{X*}(G'^i/G'^{i+1}) \to R^1 p_{X*}(G'^{i+1}/G'^{i+2}) = (R^1 p_{X*} F_X) \otimes S^{i+1} = \mathscr{H}' \otimes S^{i+1}$ are the standard Koszul ones. Then there exists a unique isomorphism $G = G'$ that coincides with $\mathrm{id}_{F(0)_X}$ on gr^0.*

(ii) *The projection $G \to F_X$ induces an isomorphism*

$$\operatorname{Ext}^1_U(p_U^* \mathscr{H}, G(1)_U) \xrightarrow{\sim} \operatorname{Ext}^1_U(p_U^* \mathscr{H}, F(1)_U) = F.$$

Proof. (i) The section $1 \in F_B = 0^*(G'/G'^1)$ lifts to a section of $0^*G'$ by 1.6.2. Therefore (i) is a consequence of 1.2.6.

(ii) By 1.3.14 it suffices to show that $\mathrm{Ext}^1_U(p_U^*\mathscr{H}, F(1)_U)$ is an F-vector space of dimension one. One has $\mathrm{Ext}^1_U(p_U^*\mathscr{H}, F(1)_U) = H^1_{\mathrm{abs}}(U, p_U^*\mathscr{H}) = H^1_{\mathrm{abs}}(B, \mathscr{H} \otimes Rp_{U*}F_U) = H^0_{\mathrm{abs}}(B, \mathscr{H} \otimes \mathscr{H}') = F$. Here the third equality holds since by 1.6.2 one has $H^1_{\mathrm{abs}}(S, \mathscr{H}) = 0$ and the fourth one holds since $\mathscr{H} \otimes \mathscr{H}' = F_B \oplus S^2(-1)$ and $H^0(B, S^2) = 0$. □

2. Residues at infinity and values at torsion points

In this section we describe the restriction of the elliptic polylogarithm to the torsion points of our elliptic curve $p\colon X \to B$.

2.1. The Eisenstein classes. In §§2.1–2.2 we assume that $F \supset \mathbb{Q}$, so $\mathscr{R} = \mathscr{F}[[\mathscr{H}]] = \prod_{i \geq 0} S^i$. Let $D \subset X$ be a subscheme of points of finite order; on $U := X \setminus D$ we have the polylogarithm extension (see 1.3.12)

$$0 \to G_U(1) \to \widetilde{\mathscr{P}}^{(D)} \to p_U^*\mathscr{R}[D]^\sharp \to 0.$$

2.1.1. Let $x \in X(B)$ be a point of finite order; so there is a canonical isomorphism $G_x = \mathscr{R}$ of sheaves on B. Therefore, if $x \notin D$, we have the extension of $\mathscr{R}$-modules

$$0 \to \mathscr{R}(1) \to \widetilde{\mathscr{P}}^{(D)}_x \to \mathscr{R}[D]^\sharp \to 0.$$

If $x \in D$, we should consider the specialization to x of our sheaves; we get the extension

$$0 \to p_{\dot{T}}^*\mathscr{R}(1) \to \mathrm{Sp}_x \widetilde{\mathscr{P}}^{(D)} \to p_{\dot{T}}\mathscr{R}[D]^\sharp \to 0$$

of sheaves on the punctured tangent line to x (identified with the punctured tangent line to zero $\dot{T}$ by x-translation). Our aim is to compute these extensions.

2.1.2. Our extensions do not split fiberwise as extensions of $\mathscr{R}$-modules. The obstruction to the splitting lies in $H^0(B, \underline{\mathrm{Ext}}^1_{\mathscr{R}}(\mathscr{R}[D]^\sharp, \mathscr{R}(1))) = F$; for our extensions this obstruction equals to 1. To kill this obstruction, we may split off a standard extension $\mathscr{K}^{(D)}$ with the same invariant. To construct it, first consider the Koszul extension

$$0 \to \mathscr{R}(1) \xrightarrow{\alpha} \mathscr{H} \otimes \mathscr{R} \xrightarrow{\beta} I \to 0.$$

Here β is the multiplication map $\mathscr{H} \otimes F[[\mathscr{H}]] \to \mathscr{H}F[[\mathscr{H}]] = I$, and for $h_1 = h_2 \in \mathscr{H}$ one has $\alpha(r\langle h_1, h_2\rangle) = h_1 \otimes h_2 r - h_2 \otimes h_1 r$. Now define $\mathscr{K}^{(D)}$ as the pull-back of the Koszul extension by the "sum of coordinates" projection $\mathscr{R}[D]^\sharp \to I$; if D is a single point, it coincides with the Koszul extension.

2.1.3. The Baer difference $\widetilde{\mathscr{P}}^{(D)}_x - \mathscr{K}^{(D)}$ splits fiberwise as an extension of $\mathscr{R}$-modules. Such extensions are classified by the group

$$H^1_{\mathrm{abs}}(B, \underline{\mathrm{Hom}}_{\mathscr{R}}(\mathscr{R}[D]^\sharp, \mathscr{R}(1)))$$

(or $H^1_{\text{abs}}(\dot{T}, \underline{\text{Hom}}_{\mathscr{R}}(\mathscr{R}[D]^\sharp, \mathscr{R}(1)))$ if $x \in D$). Note that the embeddings of $\mathscr{R}$-modules $I[D] \subset \mathscr{R}[D]^\sharp \subset \mathscr{R}[D]$ induce isomorphisms between the $\underline{\text{Hom}}_{\mathscr{R}}(\cdot, \mathscr{R}(1))$'s, and $\underline{\text{Hom}}_{\mathscr{R}}(\mathscr{R}[D], \mathscr{R}(1)) = \underline{\text{Hom}}_F(F[D], \mathscr{R}(1))$. Therefore, the above group of classes of extensions coincides with

$$H^1_{\text{abs}}(B, \underline{\text{Hom}}_F(F[D], \mathscr{R}(1))) = H^1_{\text{abs}}(D, \mathscr{R}(1))$$

(or $H^1_{\text{abs}}(D \times_B \dot{T}, \mathscr{R}(1))$ respectively). Denote by $\mathscr{E}_x^{(D)}$ the class of $\widetilde{\mathscr{P}}_x^{(D)} - \mathscr{K}^{(D)}$.

2.1.4. The functorial properties of $\widetilde{\mathscr{P}}$ imply that $\mathscr{E}_x^{(D)} = \mathscr{E}_x^{(d)}$ where $\mathscr{E}_x^{(d)}$ is the class for the D-family of curves $X_D = X \times_B D$, d is the divisor that reduces to a single (canonical) point $d \in X_D(D)$. Therefore, we may assume from the very beginning that D consists of a single point $d \in X(B)_{\text{tors}}$. The translation invariance of $\widetilde{\mathscr{P}}$ implies that (replacing x by $x - d$) we may assume that $d = 0$. So our aim is to compute $\mathscr{E}_x := \mathscr{E}_x^{(0)} \in H^1_{\text{abs}}(B, \mathscr{R}(1))$ for $x \neq 0$, $\mathscr{E}_0 \in H^1_{\text{abs}}(\dot{T}, \mathscr{R}(1))$.

2.1.5. One has $\mathscr{R}(1) = \prod S^i(1)$; so $\mathscr{E}_x = (\mathscr{E}_x^{i+2})$ where $\mathscr{E}_x^{i+2} \in H^1(B, S^i(1))$ for $x \neq 0$, $\mathscr{E}_0^{i+2} \in H^1(B \times \dot{T}, S^i(1))$ are the *Eisenstein cohomology classes of weight* $i+2$. Here is a simpler equivalent construction of the class $\mathscr{E}_x^{i+2}$. One has a canonical decomposition $\mathscr{H}^* \otimes S^{i+1}(1) \xrightarrow{\sim} S^i(1) \oplus S^{i+2}$, the first projection is $h' \otimes h^{i+1} \mapsto \frac{1}{2}\langle h', h\rangle h^{i+1}$, the second one is the multiplication map $\mathscr{H}' \otimes S^{i+1}(1) = \mathscr{H} \otimes S^{i+1} \to S^{i+2}$. Let $\widetilde{\mathscr{E}}_x^{i+2} \in H^1_{\text{abs}}(B, \mathscr{H}' \otimes S^{i+1}(1))$ be the image of the class of the extension $0 \to \mathscr{R}(1) \to \mathscr{P}_x \to \mathscr{H} \to 0$ by the $i+1$st projection

$$\text{Ext}^1(\mathscr{H}, \mathscr{R}(1)) = H^1_{\text{abs}}\left(B, \prod \mathscr{H}' \otimes S^j(1)\right) \to H^1_{\text{abs}}(B, \mathscr{H}' \otimes S^{i+1}(1)).$$

It is easy to see that $\mathscr{E}_x^{i+2}$ coincides with the first component of $\widetilde{\mathscr{E}}_x^{i+2}$ (the second component vanishes).

2.1.6. Consider the case $x = 0$. We have a short exact sequence

$$0 \to H^1_{\text{abs}}(B, S^i(1)) \to H^1_{\text{abs}}(\dot{T}, S^i(1)) \xrightarrow{\text{Res}_0} H^0_{\text{abs}}(B, S^i) \to 0.$$

If $i > 0$ then $H^0_{\text{abs}}(B, S^i) = 0$ if either we are in a mixed situation or our family is not isotrivial. Anyway we see that $\mathscr{E}_0^{i+2} \in H^1_{\text{abs}}(B, S^i(1))$.

In case $i = 0$ our sequence splits canonically: we have an isomorphism

$$H^1_{\text{abs}}(\dot{T}, F(1)) = H^1_{\text{abs}}(B, F(1)) \oplus H^0_{\text{abs}}(B, F) = H^1_{\text{abs}}(B, F(1)) \oplus F,$$

such that $1 \in F$ corresponds to $\frac{1}{12}$ (class of Δ), where $\Delta \in \mathscr{O}^*(\dot{T})$ is the discriminant (the parabolic form of weight 12 considered as a function on T).

As follows from the very definition of $\mathscr{P}$, one has $\text{Res}_0 \mathscr{E}_0^2 = 1$. We shall see in a moment that actually $\mathscr{E}_0 = 1 \in F \subset H^1_{\text{abs}}(\dot{T}, F(1))$.

2.2. The residues of $\mathcal{E}_x$ at infinity. To compute $\mathcal{E}_x$, we may assume, by functoriality, that B is a modular curve. Then, by 1.6.2, for $i \geq 1$ one has $H^1_{\text{abs}}(B, S^i(1)) = H^0_{\text{abs}}(\mathcal{E}is^{i+2}(B))$.

For $i = 0$ we have a short exact sequence

$$0 \to H^1_{\text{abs}}(H^0(B, F)(1)) \to H^1_{\text{abs}}(B, F(1)) \to H^0_{\text{abs}}(\mathcal{E}is^2(B)) \to 0.$$

It splits uniquely in a way compatible with the action of Hecke operators; by 1.3.13, $\mathcal{E}^2_x$ lives in the $H^0_{\text{abs}}(\mathcal{E}is^2(B))$-subspace. If $x = 0$ then we may assume that B is a modular "curve" of zero level, hence $\mathcal{E}is^2(B) = 0$, and the same Hecke argument shows that the $H^1_{\text{abs}}(B, F(1))$-component of $\mathcal{E}^2_0$ vanishes.

2.2.1. Let $B \subset \overline{B}$ be a smooth compactification of B; so $B^\infty := \overline{B} \backslash B$ is the (finite) scheme of parabolic points. Let $P = H^0(B^\infty, F)$ be the Artin object of F-valued functions on B^∞. There is a canonical "residue at ∞" morphism $\text{Res}_\infty : H^1(B, S^i(1)) \to P$, $\text{Res}_\infty = (\text{Res}_{b_\infty})$, $b_\infty \in B^\infty$, with kernel equal to the parabolic part of $H^1(B, S^i(1))$. It induces an isomorphism $\mathcal{E}is^{i+2}(B) \xrightarrow{\sim} P$ for $i > 0$; for $i = 0$ we have a short exact sequence $0 \to \mathcal{E}is^2(B) \to P \xrightarrow{\sigma} F \to 0$, where σ is the trace ("sum of the values") map. Therefore, the $\mathcal{E}^i_x$'s are completely determined by their residues $\text{Res}_\infty(\mathcal{E}^i_x) \in H^0_{\text{abs}}(B^\infty, F)$, which we are going to compute.

2.2.2. We may assume that our base field is $\mathbb{C}$; we shall work in classical topology. Recall a definition of Res_{b_∞}. For $b_\infty \in B^\infty$ choose a point $b \in B$ close to b_∞; let γ be a small loop around b_∞ oriented in the usual "counterclockwise" way. Put $\Gamma_b := H_1(X_b, \mathbb{Z})$; we have a short exact sequence $0 \to \mathbb{Z} \xrightarrow{\delta} \Gamma_b \xrightarrow{\pi} \mathbb{Z} \to 0$ such that the monodromy along γ acts on Γ_b as $l \mapsto l + n\delta\pi(l)$, $n > 0$. Consider the corresponding projections $\pi_F : \mathcal{H}_b = \Gamma_b \otimes F \to F$, $\pi^i_F : S^i_b \to F$. For $\alpha \in H^1(B, S^i(1))$ put

$$\text{Res}_{b_\infty}(\alpha) := \frac{i!}{n} \text{Res}_{b_\infty} \pi^i_F(\alpha) = \frac{i!}{2\pi\sqrt{-1}n} \int_\gamma \pi^i_F \alpha.$$

The exponential map identifies $X_{b\ \text{tors}}$ with $\Gamma_b \otimes \mathbb{Q}/\mathbb{Z}$; so π defines the projection $\pi_{\text{tors}} : X_{b\ \text{tors}} \to \mathbb{Q}/\mathbb{Z}$.

2.2.3. PROPOSITION. *One has* $\text{Res}_{b_\infty}(\mathcal{E}^{i+2}_x) = -\frac{1}{i+2} B_{i+2}(\langle \pi_{\text{tors}}(x) \rangle)$, *where* B_{i+2} *is the Bernoulli polynomial and for* $\mu \in \mathbb{Q}/\mathbb{Z}$, $\langle \mu \rangle \in \mathbb{Q}$ *denotes the "fractional part" of* μ *(so* $0 \leq \langle \mu \rangle < 1$, $\langle \mu \rangle = \mu \bmod \mathbb{Z}$*)*.

We shall prove 2.2.3 in 2.4.3–2.4.9.

2.3. A digression on Bernoulli polynomials. Below we describe a "topological" setting for Bernoulli polynomials. Essentially it reduces to a remark that the generating function $te^{yt}/(e^t - 1)$ is the unique solution of the differential equation $\partial_y f(y, t) = tf(y, t)$ such that $f(1, t) - f(0, t) = t$. As we shall see in 2.5 this approach also gives a simple explanation of the p-adic patterns.

Below F is our coefficient ring; we do not assume that $F \supset \mathbb{Q}$.

2.3.1. Consider the ring $R = R_F := \varprojlim F[q]/(q-1)^n$; so R is completion of the group algebra $F[\mathbb{Z}] = F[q, q^{-1}]$ with respect to powers of the augmentation ideal. Put $R^\sim := R[(q-1)^{-1}]$, $I := (q-1)R$; one has $\mathrm{gr}_I^{\cdot} R = F[\bar{t}]$ where $\bar{t} := q - 1 \bmod I^2$. Our R is a Hopf algebra: the comultiplication $\delta\colon R \to R \otimes_F R$ is $\delta(q) = q \otimes q$. For $n \in \mathbb{Z}$ we have the corresponding "nth power" endomorphism $[n].\colon R \to R$, $[n].(q) = q^n$; so $[n].(\bar{t}) = n\bar{t}$. If n is invertible in F then $[n].$ is an automorphism; it extends to an automorphism of $R^\sim$. If $F \supset \mathbb{Q}$ then we have canonical isomorphisms $R = F[[t]]$, $R^\sim = F((t))$, $q = e^t$; one has $\delta(t) = t \otimes 1 + 1 \otimes t$.

2.3.2. Consider the circle $Y = \mathbb{R}/\mathbb{Z}$; this is a topological group. Denote by G_Y the local system of R-modules on Y with fiber at zero $G_{Y,0} = R$ and the monodromy along the loop $[0, 1]$ equal to multiplication by q. Put $G_Y^\sim = G_Y[(q-1)^{-1}] = G_Y \otimes_R R^\sim$. We have a canonical map $\delta\colon G_Y \to G_Y \otimes_F G_Y$ with fiber over 0 equal to the map δ from 2.3.1. We also have a canonical isomorphism $(+)^* G_Y^{(\sim)} \xrightarrow{\sim} p_1^* G_Y^{(\sim)} \otimes p_2^* G_Y^{(\sim)}$, $G_{Y,x+y}^{(\sim)} \xrightarrow{\sim} G_{Y,x}^{(\sim)} \otimes_{R^{(\sim)}} G_{Y,y}^{(\sim)}$, of local systems of $R^{(\sim)}$-modules on $Y \times Y$ that sends $1 \in R^{(\sim)} = G_{Y,0}^{(\sim)}$ to $1 \otimes 1 \in G_{Y,0}^{(\sim)} \otimes G_{Y,0}^{(\sim)}$. It satisfies the usual associativity and commutativity constraints. Therefore, for $n \in \mathbb{Z}$ we have a canonical morphism $[n].\colon G_Y \to [n]^* G_Y$ (where $[n]\colon Y \to Y$ is multiplication by n) with fiber over zero equal to the $[n].$ map from 2.3.1.

2.3.3. Assume that n is invertible in F. Then $[n].$ is an isomorphism; it extends to $[n].\colon G_Y^\sim \xrightarrow{\sim} [n]^* G_Y^\sim$. The group $n^{-1}\mathbb{Z}/\mathbb{Z} = \ker[n]$ of translations of Y acts canonically on $[n]^*(?)$; therefore, it also acts on G_Y, $G_Y^\sim$. For different n's these actions are compatible; so G_Y, $G_Y^\sim$ are $\mathbb{Q}/\mathbb{Z}^{(F)}$-equivariant local systems, where $\mathbb{Q}/\mathbb{Z}^{(F)} \subset \mathbb{Q}/\mathbb{Z}$ is the group of translations of Y of finite order invertible in F. In particular, for $x \in \mathbb{Q}/\mathbb{Z}^{(F)} \subset Y$ we have a canonical isomorphism $G_{Y,x}^{(\sim)} \xrightarrow{\sim} G_{Y,0}^{(\sim)} = R^{(\sim)}$.

2.3.4. Note that $H^0(Y, G_Y) = H^0(Y, G_Y^\sim) = H^1(Y, G_Y^\sim) = 0$ and that the morphism $H^1(Y, G_Y) \to H^1(Y, F) = F$, induced by the projection $G_Y \to G_Y/IG_Y = F_Y$, is isomorphism (since the map $\partial\colon R \to R$, $\partial r = (q-1)r$, is injective with image I).

For a finite subset $D \subset Y$ put

$$G_Y^{(\sim)}[D] := \bigoplus_{y \in D} G_{Y,y}^{(\sim)}, \qquad G_Y[D]^\sharp := \ker(\sigma : G_U[D] \to F)$$

where $\sigma(\nu_y) = \sum_{y\in D} \bar{\nu}_y$, $\bar{\nu}_y = \nu_y \bmod IG_{Y,y} \in F$. Consider the maps $\mathrm{Res}_y\colon H^0(Y \backslash D, G_Y^{(\sim)}) \to G_Y^{(\sim)}(y)$, $\mathrm{Res}_y(\gamma) := \gamma(y+\varepsilon) - \gamma(y-\varepsilon)$ where $\varepsilon > 0$ is (very) small; put $\mathrm{Res}_D := \bigoplus_{y\in D} \mathrm{Res}_y\colon H^0(Y\backslash D, G_Y^{(\sim)}) \to G_Y^{(\sim)}[D]$. The above remark about the cohomology of $G_Y^{(\sim)}$ implies that the residue maps Res_D define isomorphisms

$$H^0(Y\backslash D, G_Y^{\sim}) \xrightarrow{\sim} G_Y^{\sim}[D], \qquad H^0(Y\backslash D, G_Y) \xrightarrow{\sim} G_Y[D]^{\sharp}.$$

2.3.5. For $\alpha = (\alpha_y) \in G_Y^{\sim}[D]$ put $\gamma_\alpha := \mathrm{Res}_D^{-1}(\alpha) \in H^0(Y\backslash D, G_Y^{\sim})$. If $\alpha \in G_Y[D]$ then $\gamma_\alpha \in H^0(Y\backslash D, (q-1)^{-1}G_Y)$; if in addition $\sigma(\alpha) = 0$ then $\gamma_\alpha \in H^0(Y\backslash D, G_Y)$. Note that for $y \in D$ we may correctly define the elements $\gamma_\alpha(y)^{\pm} := \gamma_\alpha(y \pm \varepsilon) \in G_{Y,y}$ such that $\gamma_\alpha(y)^+ - \gamma_\alpha(y)^- = \alpha_y$.

Put $\gamma := \gamma_{1(0)} \in H^0(Y\backslash\{0\}, G_Y^{\sim})$. Then one has $\gamma_\alpha(x) = \sum_{y\in D} \gamma(x-y)\alpha_y$ where $\gamma(x-y)\alpha_y$ is the image of $\gamma(x-y) \otimes \alpha_y$ under the isomorphism $G_{Y,x-y} \otimes G_{Y,y} \xrightarrow{\sim} G_{Y,x}$ from 2.3.2.

All the above constructions are compatible with changes of the coefficient ring F.

2.3.6. Assume that $F = \mathbb{R}$; so $R = \mathbb{R}[[t]]$, $R^{\sim} = \mathbb{R}((t))$. Let $\mathscr{F}$ be a trivialized (pro) C^∞-bundle on Y with fiber $\mathbb{R}[[t]]$, $\mathscr{F}^{\sim} := \mathscr{F}[t^{-1}] = \mathscr{F} \otimes_R R^{\sim}$. We have a "trivial" action of the group of translations $\mathbb{R}/\mathbb{Z}$ of Y on $\mathscr{F}^{(\sim)}$, the "constant" maps $\delta: \mathscr{F} \to \mathscr{F} \otimes_F \mathscr{F}$, $[n]_\cdot: \mathscr{F} \to [n]^*\mathscr{F}$ defined by the maps $\delta, [n]_\cdot$ from 2.3.1, and the "$\mathrm{id}_{R^{(\sim)}}$" map $\mathscr{F}_{x+y}^{(\sim)} \xrightarrow{\sim} \mathscr{F}_x^{(\sim)} \otimes_{R^{(\sim)}} \mathscr{F}_y^{(\sim)}$. Consider the connection ∇ on $\mathscr{F}^{(\sim)}$ that is defined by $\nabla_{\partial_y}(\varphi) = \partial_y\varphi - t\varphi$. Its monodromy along $[0, 1]$ is multiplication by $q = e^t$; hence, we have the canonical identifications $G_Y = \mathscr{F}^{\nabla}$, $G_Y^{\sim} = \mathscr{F}^{\sim\nabla}$ (which coincide with $\mathrm{id}_{R^{(\sim)}}$ in the zero fibers). The above operations on $\mathscr{F}^{(\sim)}$ are ∇-horizontal. Hence they define corresponding operations on the sheaves of horizontal sections; they coincide with those from 2.3.2. The restriction to $\mathbb{Q}/\mathbb{Z}$ of the $\mathbb{R}/\mathbb{Z}$-action coincides with the $\mathbb{Q}/\mathbb{Z}$-action from 2.1.3. For $y \in Y\backslash\{0\}$ put $e^{yt} := e^{\{y\}t} \in \mathbb{R}[[t]]$, where $\{y\}$ is the fractional part of y. Then e^{yt} is a ∇-horizontal section of $\mathscr{F}$ on $Y\backslash\{0\}$, and $\mathrm{Res}_0 e^{yt} = 1 - e^t$. We see that $\gamma = e^{yt}/(1-e^t) = -\sum_{k\geq 0} B_k(\{y\})t^{k-1}/k!$, where the B_k are the Bernoulli polynomials.

2.4. Computation of the residue at infinity. Let us compare the patterns from 1.3 and 2.1. Let X be an elliptic curve over $\mathbb{C}$, let T be the tangent line to zero, and let $\Gamma = H_1(X, \mathbb{Z})$. The exponential map defines an isomorphism $T/\Gamma \xrightarrow{\sim} X$, so, as a compact torus, X coincides with $\Gamma \otimes \mathbb{R}/\mathbb{Z}$.

2.4.1. Let $\pi \in H^1(X, \mathbb{Z}) = \mathrm{Hom}(\Gamma, \mathbb{Z})$ be a fixed nonzero cohomology class. It defines the projection

$$\pi_X := \pi \otimes \mathbb{R}/\mathbb{Z}: X = \Gamma \otimes \mathbb{R}/\mathbb{Z} \to Y = \mathbb{R}/\mathbb{Z}$$

of compact tori, and the morphism $\pi_R: \mathscr{R} \to R$ between the completed F-group algebras of Γ and $\mathbb{Z}$. Consider the local systems G_X, G_Y on X, Y, respectively, defined in 1.2 and 2.3; we have a canonical morphism $\pi_G: G_X \to \pi_X^* G_Y$ that coincides with π_R on the zero fibers.

We put on X the standard orientation defined by the complex structure, and, similarly, identify $\mathbb{Z}(1)$ with $\mathbb{Z}$. Since Y is oriented, we get orientations of the fibers of π_X, hence the corresponding trace map $\mathrm{tr}\colon R^1\pi_{X*}\pi_X^*\mathscr{L} \to \mathscr{L}$, where $\mathscr{L}$ is any sheaf on Y.

Let $D \subset X$ be a finite set, $\overline{D} := \pi_X(D) \subset Y$. Denote by

$$\mathrm{tr}\,\pi_X\colon H^1(X\backslash D, G_X(1)) \to H^0(Y\backslash\overline{D}, G_Y)$$

the composition

$$H^1(X\backslash D, G_X(1)) \to H^1(X\backslash\pi_X^{-1}(\overline{D}), G_X(1)) \overset{\pi_G}{\to} H^1(X\backslash\pi_X^{-1}(\overline{D}), \pi_X^*(G_Y)) \to H^0(Y\backslash D, R^1\pi_*\pi_X^*G_Y) \overset{\mathrm{tr}}{\to} H^0(Y\backslash D, G_Y)$$

and by $\mathrm{tr}\,\pi_G\colon G_X[D] \to G_Y[\overline{D}]$ the map $(\mathrm{tr}\,\pi_G(\alpha))_y = \sum_{\pi_X(x)=y}\alpha_x$. Clearly $\mathrm{tr}\,\pi_G$ sends $G_X[D]^\sharp$ to $G_Y[\overline{D}]^\sharp$.

2.4.2. LEMMA. *The diagram*

$$\begin{array}{ccc} H^1(X\backslash D, G_X(1)) & \xrightarrow[\sim]{\mathrm{Res}_D} & G_X[D]^\sharp \\ {\scriptstyle \mathrm{tr}\,\pi_X}\downarrow & & \downarrow{\scriptstyle \mathrm{tr}\,\pi_G} \\ H^0(Y\backslash\overline{D}, G_Y) & \xrightarrow[\sim]{\mathrm{Res}_{\overline{D}}} & G_Y[\overline{D}]^\sharp \end{array}$$

commutes.

PROOF. Clear. □

2.4.3. Let us start the proof of 2.2.3. Assume we are in the situation 2.2.2. Our calculation is purely topological. We may replace the base B by a small circle $Z \subset B$ around b_∞, oriented in a standard way, so $p_X\colon X \to Z$ is a fibration by two-dimensional oriented tori over the circle. The homology groups $\Gamma_z = H_1(X_z, \mathbb{Z})$, $z \in Z$, form a local system Γ on Z; we have a short exact sequence $0 \to \mathbb{Z}_Z \overset{\delta}{\to} \Gamma \overset{\pi}{\to} \mathbb{Z}_Z \to 0$ such that the monodromy along Z acts on Γ as $e \mapsto e + n\delta\pi(e)$, $n > 0$. The identification $X_z = \Gamma_z \otimes \mathbb{R}/\mathbb{Z}$ defines on our family of tori a standard flat connection. As in 2.4.1 π defines the projection $\pi_X : X \to Y = \mathbb{R}/\mathbb{Z}$, and δ defines a horizontal embedding $Q_Z \hookrightarrow X$, where Q is another copy of $\mathbb{R}/\mathbb{Z}$, $Q_Z = Q \times Z$. The group Q acts on X by translations, and the projection $(\pi_X, p_X)\colon X \to Y \times Z$ induces an isomorphism $Q\backslash X \overset{\sim}{\to} Y \times Z$.

2.4.4. It is clear that X is an Eilenberg-Mac Lane space with the "Heisenberg" fundamental group. Precisely, let us fix a point $z \in Z$. Consider the loops l_Q, l_Y in $X_z = p_X^{-1}(z)$, where $l_Q := Q_z$ and l_Y is the image of Y by a section of $\pi_{X_z} : X_z \to Y$. Let l_Z be the image of Z by the zero section 0_X. Then l_Q, l_Y, l_Z generate $\pi_1(X, 0_z)$ with the only relations $l_Q l_Z = l_Z l_Q$, $l_Q l_Y = l_Y l_Q$, $l_Z l_Y = l_Y l_Z l_Q^n$.

2.4.5. Consider the lisse prosheaf $\overline{G}_X := G_X/\delta_F(1)G_X$; here $\delta_F(1) \in \Gamma_F = \mathscr{H} \subset F[[\mathscr{H}]] = \mathscr{R}$. The projection π_G from 2.4.1 induces an isomorphism

$\theta\colon \overline{G}_X \xrightarrow{\sim} \pi_X^* G_Y$. Put $\overline{\mathscr{P}} := \mathscr{P}/\delta_F(1)G_X$; so on U we have the extension of lisse (pro)sheaves

$$0 \to \overline{G}_U \to \overline{\mathscr{P}} \to p_U^*\mathscr{H} \to 0$$

(as in 2.4.1 we identified F and $F(1)$). Let $\mathscr{L}_U \subset \overline{\mathscr{P}}$ be the preimage of $F \overset{\delta_F}{\hookrightarrow} \mathscr{H}$; we have the filtration $\overline{G}_U \subset \mathscr{L}_U \subset \overline{\mathscr{P}}$, $\mathscr{L}_U/\overline{G}_U = F_U$, $\overline{\mathscr{P}}/\mathscr{L}_U = \mathscr{H}/\delta_F(F) = F_U$. The local monodromy around 0 acts on $\mathscr{L}_U$ trivially; therefore, $\mathscr{L}_U$ extends as a lisse sheaf to X. So we have an extension

$$0 \to \overline{G}_X \to \mathscr{L}_X \to F_X \to 0$$

of lisse sheaves on X.

The Leray spectral sequence for the projection $p_X\colon X \to Z$ shows that $H^0(X, \overline{G}_X) = 0$, $H^1(X, \overline{G}_X) = H^0(Z, R^1p_{X*}\overline{G}_X) = F$. Therefore, $\pi_X^*\colon H^{\cdot}(Y, G_Y) \to H^{\cdot}(X, \overline{G}_X)$ is an isomorphism. Let

$$0 \to G_Y \to \mathscr{L}_Y \to F_Y \to 0$$

be an extension with invariant $1 \in F = H^1(Y, G_Y)$. The previous remark shows that we have a unique isomorphism $\theta\colon \mathscr{L}_X \xrightarrow{\sim} \pi_X^*\mathscr{L}_Y$ of extensions that coincides with the above θ on $\overline{G}_X$.

We may (and shall) assume that $F = \mathbb{R}$. Then, in terms of 2.3.6, one has $\mathscr{L}_Y = (t^{-1}\mathscr{F}_Y)^\nabla \subset \mathscr{F}^{\sim\nabla}$ (the identification $F_Y = \mathscr{L}_Y/G_Y$ is $1 \mapsto t^{-1}$ mod G_Y).

2.4.6. Consider the extension $0 \to \mathscr{L}_U \to \overline{\mathscr{P}} \to F_U \to 0$; let $c(\overline{\mathscr{P}}) \in H^1(U, \mathscr{L}_U)$ be its class. Take $x \in X_{\text{tors}}(Z)$, $x \neq 0$. Note that the loop $x(Z)$ lies in a single fiber $\pi_X^{-1}(y_x)$; therefore, we have $\int_{x(Z)} c(\overline{\mathscr{P}}) \in \mathscr{L}_{y_x} = t^{-1}\mathbb{R}[[t]]$. As follows from the definitions 2.1.5, 2.2.2 one has

$$\int_{x(Z)} c(\overline{\mathscr{P}}) = nt^{-1} + nt\sum_{i\geq 0} \operatorname{Res}_{b_\infty}(\mathscr{E}_x^{i+2})\frac{t^i}{i!}.$$

2.4.7. To compute this integral, consider the local system E over Y with fibers $E_y := H_1(\pi_X^{-1}(y), \mathbb{R})$. It has a global section e_0, $e_0(y) :=$ class of the circle $\pi_X^{-1}(y) \cap X_z$. The zero fiber has a distinguished element $e_1(0) :=$ class of the zero section $0_X\colon Z \hookrightarrow X$. Then $e_0(0), e_1(0)$ is a basis of E_0, and the monodromy along Y fixes $e_0(0)$ and sends $e_1(0)$ to $e_1(0) - ne_0(0)$ (see 2.4.4). Denote by $E^- \subset E$ the subsheaf with fibers $E_y^- = E_y$ for $y \neq 0$, $E_0^- = \mathbb{R}e_1(0) = H_1(\pi_X^{-1}(0) \cap U, \mathbb{R})$. We have a canonical morphism of sheaves $A\colon E^- \to \mathscr{L}$, $A(e) := \int_e c(\overline{\mathscr{P}})$. By 2.4.2, 2.3.6 one has

$$A(e_0)(y) = \frac{te^{yt}}{1-e^t} = -\sum B_i(\{y\})\frac{t^i}{i!}.$$

2.4.8. Consider the C^∞-section φ of E, defined by $\varphi(y) := e_1(\{y\}) + n\{y\}e_0(y)$; here e_1 is a horizontal section of E on the interval $0 \leq y < 1$

that equals $e_1(0)$ at $y=0$. One has $\nabla\varphi = ne_0$ and $\varphi(0)=e_1(0)\in E_0^-$. Therefore, $A(\varphi)$ is a continuous section of $\mathscr{L}_Y$ (i.e., that of $t^{-1}\mathscr{F}_Y$) such that $\nabla A(\varphi) = nA(e_0)$. These properties determine $A(\varphi)$ uniquely (since $H^0(Y, \mathscr{L}_Y)=0$). This implies that $A(\varphi)=n(\partial_t - t^{-1})A(e_0)$; i.e.,

$$A(\varphi)(y) = nt^{-1} - nt\sum_{i\geq 0}\frac{B_{i+2}(\{y\})}{i+2}\cdot\frac{t^i}{i!}.$$

2.4.9. In the situation 2.4.6 the class of $x(Z)$ in $E_{y_x}^-$ coincides with $\varphi(y_x)$. (To see this consider, for any y, $0\leq y<1$, a loop γ_y in $\pi_X^{-1}(y)$ that is the composition of a horizontal section γ_y' of p_X with a jump at z, and a path γ_y'' in $\pi_X^{-1}(y)\cap X_z$ that connects the ends of γ_y'. Assume that the γ_y form a continuous family and that for $y=0$ the path y_y'' is trivial. Then γ_y represents $e_1(y)$, γ_{y_x}' is a Q-translation of $x(Z)$, and γ_{y_x}'' represents $-nye_0(y)$. Now 2.2.3 follows from the formulas 2.4.6, 2.4.8.

2.5. p-adic measures. We shall explain how the p-adic properties of the Eisenstein classes fit into the above formalism. As a warm-up let us start with the Bernoulli polynomials situation.

2.5.1. Assume we are in the situation 2.3; let $m>0$ be an integer. The morphism $[m].$ from 2.3.2 defines, by adjointness, the trace morphism $\mathrm{tr}[m].: [m]_* G_Y \to G_Y$. For finite sets $D, D'\subset Y$, with $[m](D)\subset D'$, we have the corresponding map $\mathrm{tr}[m].: H^0(Y\backslash D, G_Y)\to H^0(Y\backslash D', G_Y)$. The trace map is compatible with residues. Namely, let $[m].: G_Y[D]\to G_Y[D']$ be the sum of the $[m]._y$, $y\in D$; one has $[m].\mathrm{Res}_D = \mathrm{Res}_{D'}\mathrm{tr}[m]$. Therefore, for $\alpha\in G_Y[D]^\sharp$ one has

$$\mathrm{tr}[m].(\gamma_\alpha) = \gamma_{[m].\alpha}.$$

If m is invertible in F we may replace G_Y by $G_Y^\sim$ (and $G_Y[D]^\sharp$ by $G_Y^\sim[D]$) in the above statement.

2.5.2. Take $D=D'=n^{-1}\mathbb{Z}/\mathbb{Z}$, with n invertible in F. The translations $G_y \xrightarrow{\sim} G_0 = R$ from 2.3.3 define a canonical isomorphism $G_Y[D]=R[D]$. They commute with $[m].$; hence the above $[m].\in \mathrm{End}\, G_Y[D]$ identifies with the endomorphism $[m].\in\mathrm{End}\, R[D]$, $([m].(r))_x = \sum_{y\in[m]^{-1}(x)}[m].(r_y)$. This operator preserves the subspace $F[D]\subset R[D]$; for $a=(a_y)\in F[D]$ one has $([m].(a))_x = \sum_{y\in[m]^{-1}(x)} a_y$. Note that $[m].\in \mathrm{End}\, F[D]$ is invertible iff $(m,n)=1$; if $m=1 \bmod n$ then $[m].=\mathrm{id}_{F[D]}$.

2.5.3. We shall consider the sections $\gamma_\alpha\in H^0(Y\backslash D, G_Y)$ for $\alpha\in F[D]^0 := \ker(\sigma : F[D]\to F) = F[D]\cap G_Y[D]^\sharp$ (or, if m is invertible in F, the sections $\gamma_\alpha\in H^0(Y\backslash D, (q-1)^{-1}G_Y)$ for $\alpha\in F[D]$). As follows from 2.5.1 if $\alpha\in F[D]$ has property $[m].\alpha=\alpha$ then

$$\mathrm{tr}[m].(\gamma_\alpha)=\gamma_\alpha.$$

For $\gamma_\alpha = \gamma$ this is the distribution property for Bernoulli polynomials (see 2.3.6). Now let P be a set of primes. Denote by $\{P\}$ the monoid of positive integers that are a product of primes in P and by $\mathbb{Q}/\mathbb{Z}^{(F,P)} \subset \mathbb{Q}/\mathbb{Z}$ the subgroup of elements of order prime to P and invertible in F. Put $\mathbb{Z}_P^\wedge := \prod_{p\in P}\mathbb{Z}_p = \varprojlim_{m\in\{P\}} \mathbb{Z}/m$. Let $\mathrm{Meas}(\mathbb{Z}_P^\wedge, R)$ be the set of R-valued measures on $\mathbb{Z}_P^\wedge$.

2.5.4. Lemma-Definition. *For $\alpha \in F[\mathbb{Q}/\mathbb{Z}^{(F,P)}]^0$ there exists a unique measure $\mu_\alpha^+ \in \mathrm{Meas}(\mathbb{Z}_P^\wedge, R)$ such that for any $m \in \{P\}$ such that $[m].\alpha = \alpha$ and $f \in \mathbb{Z}_P^\wedge/m\mathbb{Z}_P^\wedge = \mathbb{Z}/m$ one has*

$$\mu_\alpha^+(f + m\mathbb{Z}_P^\wedge) = [m].\gamma_\alpha(f/m) \quad \textit{for } f \neq 0, \quad \mu_\alpha^+(m\mathbb{Z}_P^\wedge) = [m].\gamma_\alpha(0)^+.$$

Proof. The subgroups $m\mathbb{Z}_P^\wedge$ with $[m].\alpha = \alpha$ are arbitrarily small (just consider such m that $m \equiv 1 \bmod n$, $\alpha \in n^{-1}\mathbb{Z}/\mathbb{Z}$, see 2.5.2). Therefore our μ_α^+ is uniquely determined. The fact that our μ_α^+ is a measure follows immediately from 2.5.3. □

2.5.5. Remarks. (i) One also has a measure μ_α^- such that $\mu_\alpha^-(f+m\mathbb{Z}_P^\wedge) = \mu_\alpha^+(f + m\mathbb{Z}_P^\wedge)$ for $f \neq 0$, $\mu_\alpha^-(m\mathbb{Z}_P^\wedge) = [m].\gamma_\alpha^-(0)$. The difference $\mu_\alpha^+ - \mu_\alpha^-$ equals $\alpha_0\delta_0$, where δ_0 is the δ-measure at $0 \in \mathbb{Z}_P^\wedge$.

(ii) For any integer $l > 0$ invertible in F and prime to P one has $[l]_*^{-1}\mu_\alpha^+(l \cdot U) = \mu_{[l]^*\alpha}^+(U)$. Here $U \subset \mathbb{Z}_P^\wedge$ is an open subset and $([l]^*\alpha)_y := \alpha_{ly}$.u

(iii) The ring R carries a standard differentiation ∂_t, $\partial_t(q) = q$. Therefore, μ_α^+ defines a sequence of F-valued measures $\mu_\alpha^{(n)}$, $\mu_\alpha^{(n)}(U) := \partial_t^n\mu_\alpha^+(U) \bmod I$.

(iv) If all the primes in P are invertible in F, then the Definition 2.5.4 works for any $\alpha \in F[\mathbb{Q}/\mathbb{Z}^{(F,P)}]$. The measure μ_α^+ takes values in $(q-1)^{-1}R \subset R^\sim$, and its residue—the image of μ_α^+ by the projection $(q-1)^{-1}R \to (q-1)^{-1}R/R \overset{t}{\underset{\sim}{\to}} F$—is the Haar measure of total volume $\sigma(\alpha)$.

(v) Assume that $F \supset \mathbb{Q}$. Then $R = F[[t]]$ and $\mu_\alpha^+ = \sigma(\alpha)\mu_{\mathrm{Haar}}t^{-1} + \sum_{n\geq 0}\mu_\alpha^{(n)}t^n/n!$ where μ_{Haar} is the normalized Haar measure; if $\sigma(\alpha) = 0$ then the $\mu_\alpha^{(n)}$ are the measures from (iii) above. For example, take $\alpha = 1_0$. Then the $\mu^{(n)} := \mu_{1_0}^{(n)}$ are the standard ζ-measures. Namely, for an open subset $U \subset \mathbb{Z}_P^\wedge$ let $\zeta(U, s)$ be the corresponding partial ζ-function,

$$\zeta(U, s) = \sum_{n\in\mathbb{Z}_{>0}\cap U} n^{-s} \quad \text{for } \mathrm{Re}\, s > 1.$$

One has $\mu^{(n)}(U) = \zeta(U, -n)$.

Note that for any $l > 0$ prime to P the element α_l: $[l]^*1_0 - l1_0$ belongs to $F[\mathbb{Q}/\mathbb{Z}^{(F,P)}]^0$ with $F = \mathbb{Z}[l^{-1}]$. One has $\mu_{\alpha_l}^{(n)}(U) = l^{-n}\mu^{(n)}(lU) - l\mu^{(n)}(U)$;

i.e., $\mu_{\alpha_l}^{(n)}$ is the standard l-regularization of $\mu^{(n)}$. The fact that $\mu_{\alpha_l}^{(n)}$ takes values in $\mathbf{Z}[l^{-1}]$ is a basic integrality property of the ζ-measure.

(vi) For any $P' \subset P$ and $\alpha \in F[\mathbb{Q}/\mathbb{Z}^{(F,P)}] \subset F[\mathbb{Q}/\mathbb{Z}^{(F,P')}]$ the measure μ_α on $\mathbb{Z}_{P'}^\wedge$ coincides with the push-forward of the measure μ_α on $\mathbb{Z}_P^\wedge$ by the projection $\mathbb{Z}_P^\wedge \twoheadrightarrow \mathbb{Z}_{P'}^\wedge$.

2.5.6. Now assume that $P = \{p\}$ consists of a single prime and p is nilpotent in F. The convolution gives the space of measures $\mathrm{Meas}(\mathbb{Z}_p, F)$ the structure of an F-algebra. We have a canonical Iwasawa isomorphism of F-algebras $R = \varprojlim F[q, q^{-1}]/(q^{p^n} - 1) \xrightarrow[\sim]{\partial} \mathrm{Meas}(\mathbb{Z}_p, F)$ that sends q to the δ-measure δ_1 at $1 \in \mathbb{Z}_p$. It defines the isomorphism $\mathrm{Meas}(\mathbb{Z}_p, R) \xrightarrow{\sim} \mathrm{Meas}(\mathbb{Z}_p \times \mathbb{Z}_p, F)$ which identifies an R-valued measure μ on $\mathbb{Z}_p$ with an F-valued one on $\mathbb{Z}_p \times \mathbb{Z}_p$ with $\mu(U_1 \times U_2) := (\partial\mu(U_1))(U_2)$.

2.5.7. Lemma. *For any $\alpha \in F[\mathbb{Q}/\mathbb{Z}^{(p)}]^0$ the measure μ_α on $\mathbb{Z}_p \times \mathbb{Z}_p$ is supported on the diagonal $\mathbb{Z}_p \subset \mathbb{Z}_p \times \mathbb{Z}_p$.*

Proof. We need to show that for any n the push-forward of our measure to $\mathbb{Z}/p^n \times \mathbb{Z}/p^n$ is supported on the diagonal. This means that $\mu_\alpha(f + p^n\mathbb{Z}_p)$ mod $(q^{p^n} - 1)R$ belongs to $Fq^f \subset R/(q^{p^n} - 1) = F[q, q^{-1}]/(q^{p^n} - 1)$, which is clear from the definition of μ_α since $Fq^f + (q^{p^n} - 1)R = [p^n].(G_{Y, f/p^n})$. □

We see immediately that μ_α coincides with the push-forward of $\mu_\alpha^{(0)}$ by the diagonal embedding $\mathbb{Z}_p \hookrightarrow \mathbb{Z}_p \times \mathbb{Z}_p$. Hence $\mu_\alpha^{(n)} = x^n \mu_\alpha^{(0)}$, where $x\colon \mathbb{Z}_p \to F$ is the structure map (recall that F is a $\mathbb{Z}_p$-algebra).

2.5.8. Now let us pass to the Eisenstein measures. Let n be a positive integer invertible in F and in $\mathscr{O}_B$. The subscheme $X_n \subset X$ of points of order n is étale over B; put $U = U_n := X \setminus X_n$. Consider the small polylogarithm extension

$$0 \to G_U(1) \to \mathscr{P}^{(X_n)} \to p_U^* F[X_n]^0 \to 0$$

from 1.3.12. For a section α of $F[X_n]^0$ we get the α-pull-back extension

$$0 \to G_U(1) \to \mathscr{P}^{(\alpha)} \to F_U \to 0.$$

Denote by $\gamma_\alpha \in H^1_{\mathrm{abs}}(U_n, G_{U_n}(1))$ the corresponding absolute cohomology class. If $\varphi\colon X \to X'$ is any isogeny then the morphism $\mathrm{tr}\,\varphi._G\colon \varphi_* G_X \to G_{X'}$ defines the trace map $\mathrm{tr}\,\varphi.\colon H^1_{\mathrm{abs}}(U_n, G_{U_n}(1)) \to H^1_{\mathrm{abs}}(U'_n, G_{U'_n}(1))$. As follows from 1.3.11 one has $\mathrm{tr}\,\varphi.(\gamma_\alpha) = \gamma_{\varphi\cdot(\alpha)}$. In particular, for a multiplication by m isogeny $[m]\colon X \to X$, $m = 1 \bmod n$, one has $\mathrm{tr}[m].(\gamma_\alpha) = \gamma_\alpha$.

2.5.9. Let P be a set of primes invertible in $\mathscr{O}_B$ and prime to n. Let $\{P\}$ be the monoid of products of primes in P, $T_P(X) := \varprojlim X_m$, $m \in \{P\}$. For a sheaf $\mathscr{F}$ on B we define the space $\mathrm{Meas}(T_P(X), H^1_{\mathrm{abs}}(B, , \mathscr{F}))$ of $H^1_{\mathrm{abs}}(B, \mathscr{F})$-valued measures on $T_P(X)$ as $\varprojlim H^1_{\mathrm{abs}}(X_m, \mathscr{F}_{X_m})$, $m \in \{P\}$, where the projective limit is taken with respect to the trace maps for the projections $X_{m_1 m_2} \overset{[m_2]}{\to} X_{m_1}$. Assume for simplicity that $\alpha \in F[X_n \backslash \{0\}]^0$. Then for any $m \in \{P\}$ we have $\gamma_\alpha(X_m) \in H^1_{\text{abs}}(X_m, G_{X_m}(1))$.

2.5.10. Lemma-Definition. *There exists a unique measure* $\mu_\alpha = (\mu_\alpha^{(m)}) \in \text{Meas}(T_P(X), \mathscr{R}(1))$ *such that for any* $m \in \{P\}$, $m \equiv 1 \bmod n$, *the component* $\mu_\alpha^{(m)} \in H^1_{\text{abs}}(X_m, \mathscr{R}(1)_{X_m})$ *is the image of* $i^*_{X_m}(\gamma_\alpha) \in H^1_{\text{abs}}(X_m, G_{X_m}(1))$ *by* $[m]_{G^\cdot}\colon G_{X_m}(1) \to [m]^* G_0(1) = \mathscr{R}(1)_{X_m}$.

Proof. Clear, since $\text{tr}[m].\gamma_\alpha = \gamma_\alpha$; see 2.5.8. □

Assume that $P = \{p\}$ and p is nilpotent in F. As in 2.5.6 we may identify $\text{Meas}(T_p(X), H^1_{\text{abs}}(B, \mathscr{R}(1))$ with $\text{Meas}(T_p(X) \times T_p(X), H^1_{\text{abs}}(B, F(1)))$.

2.5.11. Lemma. *The Eisenstein measure* $\mu_\alpha \in \text{Meas}(T_p(X) \times T_p(X), H^1_{\text{abs}}(B, F(1)))$ *is supported on the diagonal* $T_p(X) \subset T_p(X) \times T_p(X)$.

Proof. Same as 2.5.7. □

2.5.12. It would be interesting to compare the above picture with Nick Katz's theory of p-adic Eisenstein series.

3. Kronecker double series

In this section we shall give an explicit construction of an $\mathbb{R}$-Hodge version of the elliptic polylogarithm sheaf. We consider here, as well as in sections 4 and 6, only the case $D = \{0\}$ (see 1.3.5, 1.3.13). The constructions for arbitrary D are similar; we leave the details to the reader.

3.1. $\mathbb{R}$-Hodge structures. For details see, for example, [**D**]. Let V be a finite-dimensional $\mathbb{R}$-vector space. There are two ways to define a Hodge structure on V. Namely, a Hodge structure on V is either of the following linear algebra data:

(i) An increasing finite filtration W on V and a decreasing finite filtration $F^\cdot$ on $V_\mathbb{C} = V \otimes \mathbb{C}$ such that $F, \overline{F}$ induce on $\text{gr}^W_j V_\mathbb{C}$ the j-complementary filtrations $F^\cdot_j := F^\cdot \cap W_j V_\mathbb{C} / F^\cdot \cap W_{j-1} V_\mathbb{C}$, $\overline{F}^\cdot_j := \overline{F}^\cdot \cap W_j V_\mathbb{C} / \overline{F}^\cdot \cap W_{j-1} V_\mathbb{C}$ (i.e., one has $\text{gr}^W_j V_\mathbb{C} = \bigoplus_{p+g=j} F^p_j \cap \overline{F}^q_j$).

(ii) A bigrading $V_\mathbb{C} = \bigoplus_{p,q} V^{p,q}$ and a linear operator $N \in \text{End}\, V_\mathbb{C}$ such that $\overline{V^{p,q}} = V^{p,q}$, $\overline{N} = -N$, and $N(V^{p,q}) \subset \bigoplus_{\substack{a<p \\ b<q}} V^{a,b}$.

The data (i) and (ii) are in a canonical one-to-one correspondence. It assigns to (ii)-data $(V^{p,q}, N)$ the following (i)-data: $W_j = \bigoplus_{p+q \le j} V^{p,q}$, $F^i = (\exp N)(\bigoplus_{p \ge i} V^{p,q})$.

We shall call N the *Green's operator* of a Hodge structure. The name comes from the following

3.1.1. Example. Let X be a compact Riemannian surface, $A, B \subset X$ be finite sets of points such that $A \cap B = \varnothing$. Put $V = H^1(X \setminus A, B; \mathbb{R})$.

According to Deligne, V carries a canonical mixed Hodge structure. One has $V = V^{1,1} \oplus V^{0,1} \oplus V^{1,0} \oplus V^{0,0}$ so that the only nonzero component of N is $N\colon V^{1,1} \to V^{0,0}$. One has $V^{0,0} = W_0V = \mathrm{Im}(H^0(B, \mathbb{R}) \to H^1(X\backslash A, B; \mathbb{R})) = H^0(B, \mathbb{R})/H^0(X, \mathbb{R}) = \mathbb{R}[B]/\mathbb{R}$ and $V^{1,1} = V/W_1V = H^1(X\backslash A, B; \mathbb{R})/H^1(X, B; \mathbb{R}) = \mathrm{Im}(\mathrm{Res}\colon H^1(X\backslash A, \mathbb{R}) \to H^0(A, \mathbb{R})(-1)) = \{\sum_i n_ia_i \in \mathbb{R}[A](-1) : \sum n_i = 0\}$. It is easy to see that $N(\sum_i n_ia_i) = \sum_{i,j} n_iG(a_i, b_j)b_j$ where $G(x, y)$ is the Green's function of the Laplacian of X.

3.2. Quasi-Hodge sheaves. Let S be a complex analytic manifold. A *quasi-*$\mathbb{R}$*-Hodge sheaf* on S is a C^∞-class $\mathbb{R}$-vector bundle V on S with a C^∞-connection ∇, equipped with $\mathbb{R}$-Hodge structures on the fibers V_s, $s \in S$, that depend continuously on s. The following axioms (i)–(iii) should hold.

(i) The curvature $\mathrm{curv}(\nabla)$ lies in $\Omega_S^{1,1} \otimes W_{-2}\,\mathrm{End}\,V \subset \Omega_S^2 \otimes \mathrm{End}\,V$ (here $W_{-2}\,\mathrm{End}\,V = \{f \in \mathrm{End}\,V : f(W_iV) \subset W_{i-2}V\}$).

(ii) The filtration $W_\cdot$ is ∇-horizontal.

(iii) One has $\nabla(F^i) \subset \Omega_S^{0,1} \otimes F^i + \Omega_S^{1,0} \otimes F^{i-1}$.

The axioms imply that $\bar\partial = \nabla^{0,1}$ is integrable; i.e., it defines a holomorphic structure on $V_{\mathbb{C}}$ and the F^i are holomorphic subbundles. The curvature $\mathrm{curv}(\nabla)$ maps W_j to W_{j-2} and F^i to F^{i-1}; hence, one has

(iv) $\mathrm{curv}(\nabla) \in \Omega_S^{1,1} \otimes \mathrm{Hom}_{\mathscr{H}}(V, V(-1))$, where $\mathrm{Hom}_{\mathscr{H}}$ means morphisms of vector bundles compatible with the Hodge structures on the fibers.

The quasi-Hodge sheaves on S form in an obvious manner a Tannakian $\mathbb{R}$-category $\mathscr{QH}(S)$; the flat sheaves (i.e., the ones with curvature 0) form a full Tannakian subcategory $\mathscr{QH}^0(S)$. If S is an algebraic manifold, then the category $\mathscr{H}(S)$ of $\mathbb{R}$-Hodge sheaves is a full Tannakian subcategory of $\mathscr{QH}^0(S)$.

We shall need the following example. Let $\mathscr{I}$ be a quasi-Hodge sheaf with zero curvature such that $\mathscr{I} = W_{-2}\mathscr{I}$, $F^0\mathscr{I} = 0$, and the Green's operator $N_{\mathscr{I}} = 0$.

3.2.1. Lemma. *The vector space* $\mathrm{Ext}^1_{\mathscr{QH}(S)}(\mathbb{R}(0)_S, \mathscr{I})$ *is canonically isomorphic to the space of* C^∞*-sections* g *of* $\mathscr{I}_{\mathbb{C}}$ *that satisfy* $\bar g = -g$, $\partial g \in \Omega_S^{1,0} \otimes F^{-1}$.

Proof. Let $0 \to \mathscr{I} \to P \to \mathbb{R}(0)_S \to 0$ be an extension. The Green's operator $N_P(1)$ maps P to $\mathscr{I}$ and $\mathscr{I}$ to zero, hence is a morphism from $\mathbb{R}(0)$ to $\mathscr{I}$; put $g_P = N_P(1)$. This is the desired section of $\mathscr{I}$ that corresponds to P; by 3.1(ii) $\bar g_P = g_P$. According to 3.1(ii) we have a canonical direct sum decomposition of F as C^∞-vector bundle $P = W_{-2}P \oplus P^{0,0}$. In terms of this direct sum decomposition the Hodge filtration on P is given by $F^1P = 0$, $F^0P = \mathbb{C}(g+1)$, $F^iP = F^i\mathscr{I} + F^0P$ for $i < 0$; here 1 is the generator of $\mathbb{R}_S \subset \mathcal{I} \oplus \mathbb{R}_S$. Since $F^0 P$ projects isomorphically to $F^0\mathbb{R}(0)_S$ as holomorphic bundle, we see that $\overline{\partial}(g+1) = \overline{\partial}g + \overline{\partial}(1) = 0$; hence $\overline{\partial(1)} = -\overline{\partial}g$ and $\partial(1) = \overline{\overline{\partial}(1)} = \partial g$. We have $\partial(g+1) = 2\partial g \in \Omega_S^{1,0} \otimes F^{-1}\mathcal{I}$ by 3.2(iii). This implies the conditions of the lemma. Conversely, given g as in 3.2.1, we may construct the extension P_g as $\mathcal{I} \oplus \mathbb{R}_S$ equipped with the connection ∇_g such that $\nabla_g(1) = \partial g - \overline{\partial} g$ and the Hodge structure defined by the above formulas. □

3.2.2. REMARK. One has $\operatorname{curv}\nabla_g = 2\overline{\partial}\partial g \in \Omega_S^{1,1} \otimes \mathcal{I}^{1,1} = \Omega_S^{1,1} \otimes \operatorname{Hom}(\mathbb{R}(0), \mathcal{I}^{1,1}) \subset \Omega_S^{1,1} \otimes \operatorname{End} P_g$.

3.3. The Eisenstein series. Let $p_X : X \to B$ be our family of elliptic curves where B is a smooth $\mathbb{C}$-scheme.

3.3.1. Consider the $\mathbb{R}$-Hodge sheaf $\mathcal{H} = \mathcal{H}_{\mathbb{R}}$ on B. For $b \in B$ one has $\mathcal{H}_b = H_1(X_b, \mathbb{R}) = \Gamma_b \otimes \mathbb{R}$ where $\Gamma_b = H_1(X_b, \mathbb{Z})$. The exponential map identifies X_b, considered as a topological space, with the compact torus $\mathcal{H}_b/\Gamma_b$. The canonical connection ∇ on $\mathcal{H}$ defines a flat C^∞-connection on this family of tori. The C^∞-tangent bundle to $X_b = \mathcal{H}_b/\Gamma_b$ is a trivialized bundle with fiber $\mathcal{H}_b$. Therefore, this connection defines a closed C^∞-class $p_X^*\mathcal{H}$-valued 1-form ν on X: for a tangent vector λ one has $\nu(\lambda) = \nabla$-vertical component of λ. Any local section $\gamma \in \Gamma$ defines a C^∞-function $\chi_\gamma : X \to S^1 \subset \mathbb{C}^*$; for $h \in X_b = \Gamma_b \otimes \mathbb{R}/\mathbb{Z}$ one has $\chi_\gamma(h) := \exp\langle\gamma, h\rangle$, where $\langle\ ,\ \rangle : \Gamma \times \Gamma \to 2\pi i\mathbb{Z}$ is the intersection pairing. One has $d\log\chi_\gamma = \langle\gamma, \nu\rangle$.

Let $\mathcal{H} \otimes \mathbb{C} = \mathcal{H}^{0,-1} \oplus \mathcal{H}^{-1,0}$ be the Hodge decomposition. For $\gamma \in \mathcal{H}$, we denote by $\gamma^{0,-1}$, $\gamma^{-1,0}$ the components of γ. For $a, b \in \mathbb{Z}$ consider the series

$$g_{a,b} = \sum{}' \chi_\gamma \frac{(\gamma^{0,-1})^{a-1}(\gamma^{-1,0})^{b-1}}{\langle\gamma^{-1,0}, \gamma^{0,-1}\rangle^{a+b-1}}$$

of $(\mathcal{H}^{0,-1})^{\otimes a-1} \otimes (\mathcal{H}^{-1,0})^{\otimes b-1}$-valued functions on X; here, as usual, $\sum'$ means summation over $\Gamma\setminus\{0\}$. For $a+b > 2$ this series converges absolutely; for arbitrary a, b it converges in the sense of distributions to a generalized section which is of C^∞-class on $U = X\setminus 0(B)$.

For example $g_{1,1}$ is a normalized Green's function—this is (the only) L^1-class function on X smooth on U such that the restriction $g_{1,1b}$ to the torus X_b satisfies the equation $\partial\overline{\partial}g_{1,1b} = \mu_b - \delta_0$, where μ_b is invariant volume form with $\int_{X_b}\mu_b = 2\pi i$ and δ_0 is the δ-measure at 0, and such that one has $\int_{X_b} g_{1,1}\mu_b = 0$.

3.3.2. Here are explicit formulas for the above functions. We may consider the (locally) universal standard family of curves over the upper half-plane H. Let τ, ξ be the standard parameters on H, $\mathbb{C}$, respectively. For $\tau \in H$ one has $\Gamma_\tau = \mathbb{Z} + \mathbb{Z}\tau \overset{\alpha}{\hookrightarrow} \mathbb{C}$, $X_\tau = \mathbb{C}/\Gamma_\tau$. Denote by $l \in \mathcal{H}_\tau^{-1,0}$, $\overline{l} \in \mathcal{H}_\tau^{0,-1}$ the base vectors dual to $d\xi$, $\overline{d\xi}$ respectively; then $\gamma^{-1,0} = \alpha(\gamma)l$, $\gamma^{0,-1} = \overline{\alpha(\gamma)l}$

for $\gamma \in \Gamma_\tau$. One has the following formulas:

$$\langle l , \bar{l} \rangle = \frac{2\pi i}{\tau - \bar{\tau}},$$

$$\nu = \nu^{1,0} + \nu^{0,1} = \left(d\xi - \frac{\xi - \bar{\xi}}{\tau - \bar{\tau}} d\tau \right) l + \left(\overline{d\xi} - \frac{\xi - \bar{\xi}}{\tau - \bar{\tau}} \overline{d\tau} \right) \bar{l},$$

$$\chi_\gamma(\tau , \xi) = \exp 2\pi i \left(\frac{\alpha(\gamma)\bar{\xi} - \overline{\alpha(\gamma)}\xi}{\tau - \bar{\tau}} \right),$$

$$\nabla(l) = -\frac{l d\tau + \overline{l d\tau}}{\tau - \bar{\tau}}, \qquad \nabla(\bar{l}) = \frac{l d\tau + \overline{l d\tau}}{\tau - \bar{\tau}},$$

$$g_{a,b}(\tau , \xi) = \sum{}' \chi_\gamma(\tau , \xi) \left(\frac{\tau - \bar{\tau}}{2\pi i} \right)^{a+b-1} \frac{l^{b-1}\bar{l}^{a-1}}{\alpha(\gamma)^a \overline{\alpha(\gamma)}^b}.$$

3.3.3. Consider the connection ∇ on $\mathscr{H}$. One has $\nabla = \partial + \bar{\partial} = \nabla^0 + \lambda$ where $\partial , \bar{\partial}$ are the $(1,0)$- and $(0,1)$-components of ∇, ∇^0 is the component of ∇ that preserves the Hodge decomposition, and

$$\lambda = \lambda^{1,0} + \lambda^{0,1} \in \mathrm{Hom}(\mathscr{H}^{0,-1}, \mathscr{H}^{-1,0}) \otimes \Omega_B^{1,0} \oplus \mathrm{Hom}(\mathscr{H}^{-1,0}, \mathscr{H}^{0,-1}) \otimes \Omega_B^{0,1}.$$

In the setting 3.3.2 one has

$$\lambda^{1,0} = \frac{1}{\tau - \bar{\tau}} \frac{l}{\bar{l}} d\tau, \qquad \lambda^{0,1} = \overline{\lambda^{1,0}} = \frac{-1}{\tau - \bar{\tau}} \frac{\bar{l}}{l} \overline{d\tau}.$$

For $a, b \geq 1$ one has $(\mathscr{H}^{0,-1})^{\otimes a-1} \otimes (\mathscr{H}^{-1,0})^{\otimes b-1} = [S^{a+b-2}]^{-b+1,-a+1} = [S^{n+b-2}(1)]^{-b,-a}$. Therefore, we may consider $g_{a,b}$ as a section of $p_X^* S^{a+b-2}$ and differentiate it.

LEMMA. (i) *One has*

$$\partial g_{a,b} = -\nu^{1,0} g_{a,b-1} + (a g_{a+1,b-1} + (a-1) g_{a,b}) \lambda^{1,0},$$

$$\bar{\partial} g_{a,b} = \nu^{0,1} g_{a-1,b} + (b g_{a-1,b+1} + (b-1) g_{a,b}) \lambda^{0,1}.$$

(ii) $\bar{\partial}\partial g_{1,1} = -\frac{1}{2}\langle \nu , \nu \rangle + \delta_0$.

PROOF. This follows by direct calculation using the above formulas; note that

$$\partial \langle \gamma^{-1,0}, \gamma^{0,-1} \rangle^{-1} = \frac{d\tau}{2\pi i \alpha(\gamma)^2},$$

$$\partial \gamma^{-1,0} = -\frac{\overline{\alpha(\gamma)}}{\tau - \bar{\tau}} l d\tau,$$

$$\partial \gamma^{0,-1} = \frac{\overline{\alpha(\gamma)}}{\tau - \bar{\tau}} l d\tau,$$

$$\partial \chi_\gamma = \frac{2\pi i \overline{\alpha(\gamma)}}{\tau - \bar{\tau}} \left(-d\xi + \frac{\xi - \bar{\xi}}{\tau - \bar{\tau}} d\tau \right) \chi_\gamma.$$

For $n \geq 2$ put

$$g_n := \sum_{a,b\geq 1, a+b=n} (-1)^a g_{a,b}.$$

This is an L^1-class section of $p_X^* S^{n-2}(1)$ smooth on U (and smooth everywhere for $n \geq 3$).

3.3.5. COROLLARY. *One has* $\overline{g}_n = -g_n$, $dg_n + \nu g_{n-1} = (-1)^n g_{n-1,0}\nu^{1,0} + (-1)^{n-1}(n-1)g_{n,0}\lambda^{1,0} - g_{0,n-1}\nu^{0,1} - (n-1)g_{0,n}\lambda^{0,1}$.

PROOF. The first statement is clear since $\overline{\chi}_\gamma = \chi_{-\gamma}$ (note that the (1)-twist changes the sign of complex conjugation). The second one follows from 3.3.4. □

3.4. The logarithm sheaf. Let us construct the sheaf G. As C^∞-bundle of $\mathbb{R}$-Hodge structures, G coincides with $\prod_{n\geq 0} p_X^* S^n$. The connection ∇_G is equal to $\nabla_S + \nu$, where ∇_S is the direct sum of the usual connections ∇_{S^i} on the $p_X^* S^i$ and ν is the "multiplication by ν endomorphism" which is a ∇_S-closed C^∞-class 1-form with values in $\prod_i \operatorname{Hom}(p_X^* S^i, p_X^* S^{i+1}) \subset \operatorname{End} G$. It is clear that G is a flat quasi $\mathbb{R}$-Hodge sheaf on X. It is a $p_X^* S^\cdot$-module in an obvious manner. Note that $G_0 = 0^* G$ coincides with $\prod_{n\geq 0} S^n$ as an $\mathbb{R}$-Hodge sheaf. In particular, we have a canonical section $1 \in \mathbb{R} = S^0 \subset G_0$.

3.4.1. LEMMA. *G is an $\mathbb{R}$-Hodge sheaf on X. The pair $(G, 1)$ is canonically isomorphic to the $\mathbb{R}$-Hodge version of the pair $(G, 1)$ from* 1.2.8.

PROOF. It suffices to check that our G is a Hodge sheaf since G (equipped with the weight filtration) obviously satisfies the conditions 1.2.6(a)$'$. One has an obvious identification $G/W_{-n-1}G = \operatorname{Sym}^n(G/W_{-2}G)$. Therefore, it suffices to check that $G/W_{-2}G$ is a Hodge sheaf. For $x \in X_b$ consider the relative homology group $H_1(X_b, \{0, x\}; \mathbb{R})$; we have a canonical short exact sequence of $\mathbb{R}$-Hodge structures $0 \to \mathscr{H}_b \to H_1(X_b, \{0, x\}; \mathbb{R}) \xrightarrow{\partial} \mathbb{R}(0) \to 0$ where the projection ∂ assigns to a relative 1-cycle its boundary at zero. This sequence splits (since $\mathscr{H}_b$ has type $(0, -1), (-1, 0)$); therefore, we have a canonical isomorphism of mixed Hodge structures $\varphi_{x,b}: (G/W_{-2}G)_x \xrightarrow{\sim} H_1(X_b\{0, x\}; \mathbb{R})$ that is identity on $\mathscr{H}_b$ and $\mathbb{R}(0)$. When (x, b) varies the groups $H_1(X_b, \{0, x]\}; \mathbb{R})$ form a lisse Hodge sheaf $\widetilde{\mathscr{H}}$ on X; so our assertion will follow if we show that our isomorphism $\varphi: G/W_{-2}G \to \widetilde{\mathscr{H}}$ is horizontal. It suffices to show that the image of the section 1 of $\mathbb{R} \subset G/W_{-2}G$ in $\widetilde{\mathscr{H}}$ satisfies the condition $\nabla\varphi(1) = \nu$. The element $\varphi(1)_x \in H_1(X_b, \{0, x\}\mathbb{R})$ is the only element of $F^0 \cap H^1(X_b, \{0, x\}; \mathbb{R})$ such that $\partial\varphi(1)_x = 1$. Therefore, in the setting of 3.3.2, one has $\varphi(1)_x = \gamma_\xi + \xi l + \overline{\xi l}$ where $x = \xi \bmod \Gamma$ and γ_ξ is the straight path from ξ to 0. One has $\nabla\varphi(1) = \nabla(\xi l + \overline{\xi l}) = \nu$ by 3.3.2. □

3.5. The auxiliary quasi-Hodge extensions. Consider the section $g = \sum_{n\geq 2} g_n$ of $G(1)_U$. It satisfies, according to 3.3.5, the conditions of 3.2.1. Let $\mathcal{F}^{\text{ell}}$ be the corresponding quasi-Hodge extension of $\mathbb{R}(0)_U$ by $G(1)_U$. We have also the corresponding $S^{\cdot}$-modules extension

$$0 \to G(1)_U \to \widetilde{\mathcal{F}}^{\text{ell}} \to \prod_{i\geq 0} p_U^* S^i \to 0.$$

Here $\widetilde{\mathcal{F}}^{\text{ell}}$ is the push-out of

$$0 \to \prod p_U^* S^i \otimes G(1)_U \to \prod_{i\geq 0} p_U^* S^i \otimes \mathcal{F}^{\text{ell}} \to \prod_{i\geq 0} p_U^* S^i \to 0$$

along the multiplication map $\prod p_U^* S^i \otimes G(1)_U \to G(1)_U$. Let

$$0 \to G(1)_U \to \widetilde{\mathcal{P}}' \to p_U^* I \to 0 \tag{3.5.1}$$

be the pull-back of $\widetilde{\mathcal{F}}^{\text{ell}}$ by $I = \prod_{i\geq 1} S^i \hookrightarrow \prod_{i\geq 0} S^i$. This is also an extension of $S^{\cdot}$-modules.

Define another quasi-Hodge extension $\widetilde{\mathcal{P}}''$ of $p_X^* I$ by $G(1)$ as follows. As C^∞-bundle of Hodge structures, $\widetilde{\mathcal{P}}''$ coincides with $\prod_{i\geq 0} p_X^* S^i \otimes \mathcal{H}$. The connection $\nabla_{\widetilde{\mathcal{P}}''}$ equals $\nabla_{S\otimes\mathcal{H}} + \nu_{\widetilde{\mathcal{P}}''}$ where $\nabla_{S\otimes\mathcal{H}}$ is the direct sum of the standard connections on the $S^i \otimes \mathcal{H}$ and ν is a 1-form with values in $\prod \operatorname{Hom}(S^i \otimes \mathcal{H}, S^{i+1} \otimes \mathcal{H}) \subset \operatorname{End} \widetilde{\mathcal{P}}''$ given by the formula $\nu_{\widetilde{\mathcal{P}}''}(\delta)(f \otimes h) = f\nu(\delta) \otimes h - fh \otimes \nu(\delta)$. Here δ is a tangent vector at the point $(x, b) \in X$, $h \in \mathcal{H}_b$, $f \in S_b^i$, $\nu(\delta) \in \mathcal{H}_b$.

3.5.2. LEMMA. *$\widetilde{\mathcal{P}}''$ is a quasi-Hodge sheaf on X with curvature given by the formula* $\operatorname{curv}_{\widetilde{\mathcal{P}}''}(\delta_1 \wedge \delta_2)(f \otimes h) = fh\nu(\delta_1) \otimes \nu(\delta_2) - fh\nu(\delta_2) \otimes \nu(\delta_2)$.

PROOF. Direct check. □

Consider the Koszul short exact sequence of C^∞-bundles of Hodge structures with S-action

$$0 \to G(1)_X \xrightarrow{\alpha} \widetilde{\mathcal{P}}'' \xrightarrow{\beta} p_X^* I \to 0. \tag{3.5.3}$$

Here $\alpha(f_1\langle h_1, h_2\rangle) = f_1 h_1 \otimes h_2 - f_1 h_2 \otimes h_1$, and $\beta(f_2 \otimes h_3) = f_2 h_3$, where $h_i \in \mathcal{H}_{\mathbb{R}}$, $\langle h_1, h_2\rangle \in 2\pi i\mathbb{R}$ is the intersection pairing, f_1 is a section of G identified (as a bundle of Hodge structures) with $\prod S^i$ (see 3.4), and f_2 is a section of $\prod S^i$. As follows from 3.5.2 and the definition of G, the morphisms in 3.5.3 are horizontal; i.e., 3.5.3 is an extension of quasi Hodge sheaves.

3.6. The polylogarithm sheaf. Let

$$0 \to G(1)_U \to \widetilde{\mathcal{P}} \to p_U^* I \to 0$$

be the Baer sum of the extensions (3.5.1) and (3.5.3) (restricted to U).

3.6.1. Proposition. *$\widetilde{\mathscr{P}}$ is a Hodge sheaf on U, and the extension* 3.6 *is canonically isomorphic to the polylogarithm extension from* 1.3.5 *(for $D = (0)$).*

Proof. By 3.2.2, 3.3.4(ii) the curvature of the sheaf $\mathscr{F}^{\mathrm{ell}}$ from 3.5 is the composition $\mathscr{F}^{\mathrm{ell}} \to \mathbb{R} \to \Omega_U^{1,1} \otimes G(1) \subset \Omega_U^{1,1} \otimes \mathscr{F}^{\mathrm{ell}}$ where the first arrow is the projection of our extension and the second one maps $1 \in \mathbb{R}$ to $-\langle \nu, \nu \rangle \in \Omega_U^{1,1} \otimes \mathbb{R}(1) \subset \Omega_U^{1,1} \otimes G(1)$. Therefore, the curvature of $\widetilde{\mathscr{P}}'$ is the composition $\widetilde{\mathscr{P}}' \to p_U^* I \to \Omega_U^{1,1} \otimes G(1) \subset \Omega_U^{1,1} \otimes \widetilde{\mathscr{P}}'$ where the first arrow is the projection and the second is multiplication by $-\langle \nu, \nu \rangle$.

By 3.5.2 this implies that our $\widetilde{\mathscr{P}}$ has zero curvature. To prove that $\widetilde{\mathscr{P}}$ is a Hodge sheaf, one needs to check the asymptotic conditions at infinity (see, e.g., [**K**]). Note that $\widetilde{\mathscr{P}}$ is generated, as $S^{\cdot}$-module, by the quasi-Hodge subsheaf $\mathscr{P}$ defined as the preimage of $p_X^* \mathscr{H} \subset p_X^* I$ in $\widetilde{\mathscr{P}}$ (cf. 3.2). Therefore, it suffices to check that $\mathscr{P}$ is a Hodge sheaf.

Assume that the base B is a single point. Then X is compact. The local monodromy around $0 = X \backslash U$ acts trivially on $\mathrm{gr}^W_{\cdot} \mathscr{P}$; therefore, W coincides with the relative weight filtration. Let $\mathscr{P}^{\vee}_{\mathscr{O}_X}$ be the Deligne extension of $\mathscr{P}_{\mathscr{O}_X}$ to X. One has a short exact sequence of $\mathscr{O}_X$-modules $0 \to G(1)_{\mathscr{O}_X} \to \mathscr{P}^{\vee}_{\mathscr{O}_X} \to p_X^* \mathscr{H}_{\mathscr{O}_X} \to 0$. The condition to check is that near 0 there exists a holomorphic section $\gamma^{\vee}: p_X^* \mathscr{H}_{\mathscr{O}_X} \to \mathscr{P}^{\vee}_{\mathscr{O}_X}$ compatible with the Hodge filtrations on U. Equivalently, we need to construct a holomorphic section $\gamma: p_U^* \mathscr{H}_{\mathscr{O}_U} \to \mathscr{P}_{\mathscr{O}_U}$ on a punctured neighbourhood of 0 compatible with the Hodge filtrations and such that $\xi \partial_\xi \gamma: p_U^* \mathscr{H}_{\mathscr{O}_U} \to G(1)_{\mathscr{O}_U}$ is holomorphic at 0; here ξ is a local parameter at 0.

Consider the quasi Hodge sheaf $\mathscr{F}^{\mathrm{ell}}$ on U from 3.5. We have the holomorphic section $g+1$ of $F^0 \mathscr{F}^{\mathrm{ell}}$ (see the proof of 3.2.1) such that $\partial(g+1) = 2\partial g$. By 3.3.4(ii), 3.3.5 the section $\xi \partial_\xi (g+1)$ of $G(1)$ is continuous at 0. Denote by $\gamma': p_U^* \mathscr{H} \to \mathscr{P}'$ the section $\gamma'(h) = h(g+1)$ (here $\mathscr{P}' \subset \widetilde{\mathscr{P}}'$ is the preimage of $p_U^* \mathscr{H} \subset p_U^* I$). Then γ' is compatible with the Hodge filtrations and $\xi \partial_\xi \gamma'$ is continuous at 0. Let $\gamma'': p_X^* \mathscr{H} \to \mathscr{P}'' \subset \widetilde{\mathscr{P}}''$ be any holomorphic section defined on a neighbourhood of 0 and compatible with Hodge filtrations (recall that $\mathscr{P}''$ is defined at 0). Put $\gamma = \gamma' + \gamma''$. This is a holomorphic section of $\mathscr{P}$ compatible with the Hodge filtrations and such that $\xi \partial_\xi \gamma$ is continuous and, hence, holomorphic at 0.

Denote by $\widetilde{\mathscr{P}}^{\mathrm{pol}}$ the polylogarithm extension from 1.3.5. If B is a point then we have shown that $\widetilde{\mathscr{P}}$ is a Hodge sheaf. By 1.3.4 and the construction of $\widetilde{\mathscr{P}}''$ we have a canonical isomorphism of extensions $\varphi: \widetilde{\mathscr{P}}^{\mathrm{pol}} \xrightarrow{\sim} \widetilde{\mathscr{P}}$. If B is arbitrary then we have the canonical isomorphisms $\varphi_b: \widetilde{\mathscr{P}}^{\mathrm{pol}}_{X_b} \to \widetilde{\mathscr{P}}_{X_b}$ for each fiber X_b. Together they form a C^∞-isomorphism φ of the corresponding extensions of bundles of Hodge structures horizontal along the fibers. The derivative $\nabla\varphi$ of φ is a closed 1-form with values in the local system $\mathrm{Hom}(p_X^*\mathscr{H}, G(1))$. The restriction of $\nabla\varphi$ to the fibers vanishes, i.e., $\nabla\varphi$ belongs to $p^*\Omega_B^1 \subset \Omega_X^1$. Hence, for each tangent vector ∂ at $b \in B$ the C^∞-class section $\nabla\varphi(\partial)$ of $\mathrm{Hom}(\mathscr{H}_b, G(1)_{X_b})$ is horizontal. But this local system has no horizontal sections; hence, $\nabla\varphi = 0$ and φ is horizontal. Therefore, φ is an isomorphism of flat quasi-Hodge sheaves; in particular, $\widetilde{\mathscr{P}}$ is a Hodge sheaf. □

3.7. Remarks. (i). We have used the quasi-Hodge sheaves with curvature to split the "polylogarithm gerb". Compare the above $\mathbb{R}$-Hodge construction with 2.1. By the way, we may compute the residues at ∞ from §2 using the formulas from the present section (since the $\mathbb{R}$-Hodge (p, q)-splitting of G being restricted to the torsion points coincides with the splitting from 2.2, the formulas reduce the computation of residues at ∞ for $\mathscr{E}_x^i$ to one for holomorphic Eisenstein series). Conversely, the results of §2 determine the restriction of $\widetilde{\mathscr{P}}$ to torsion points, hence the Green's operator of $\widetilde{\mathscr{P}}$ (since the torsion points are dense). This would give an alternate proof of 2.2.3.

4. Elliptic polylogarithm functions

In this section we describe explicitly the $\mathbb{Q}$-Hodge realization of the elliptic polylogarithm. Our construction follows the pattern of §3. Using Schottky uniformization we enlarge the category of Hodge sheaves on our elliptic curve to the one of quasi-Hodge sheaves to trivialize the polylogarithm gerb. This helps to represent the elliptic polylogarithm extension as a Baer sum of two quasi-Hodge ones: a ζ-regularized q-averaging of a classical polylogarithm extension and a Koszul type correction.

4.1. Quasi-Hodge sheaves. Let S be a connected complex analytic manifold. A quasi $\mathbb{Q}$-Hodge sheaf V on S is the same as a variation of mixed $\mathbb{Q}$-Hodge structures: such V is a $\mathbb{Q}$-local system on S together with a $\mathbb{Q}$-Hodge structure on each fiber V_s, $s \in S$, such that $W_\cdot$ is a filtration by $\mathbb{Q}$-local subsystems, $F^\cdot$ is a holomorphic filtration, and Griffiths' transversality condition $\partial F^i \subset F^{i-1}$ holds. Denote by $\mathscr{QH}(S)_\mathbb{Q}$ the category of quasi $\mathbb{Q}$-Hodge sheaves. This is a Tannakian $\mathbb{Q}$-category. If S is an algebraic variety then the category $\mathscr{H}(S)_\mathbb{Q}$ of lisse $\mathbb{Q}$-Hodge sheaves on S is a full Tannakian subcategory of $\mathscr{QH}(S)_\mathbb{Q}$ that consists of graded polarizable objects that satisfy certain asymptotic conditions at infinity (see, e.g., **[K]**).

Now let $\pi: \widetilde{S} \to S$ be a local homeomorphism such that $\widetilde{S}$ is nonempty and connected. A quasi $\mathbb{Q}$-Hodge sheaf $\mathscr{F}$ on S relative to $\widetilde{S}$ is a collection $(\mathscr{F}_{\widetilde{S}}, \mathscr{F}_{S^i}, \varphi_i)$ where $\mathscr{F}_{\widetilde{S}} \in \mathscr{QH}(\widetilde{S})_\mathbb{Q}$, $\mathscr{F}_{S^i} \in \mathscr{QH}(S)_\mathbb{Q}$ $(i \in \mathbb{Z})$, and $\varphi_i: \pi^*\mathscr{F}_{S^i} \xrightarrow{\sim} \mathrm{gr}_i^W \mathscr{F}_{\widetilde{S}}$ are isomorphisms. Clearly $\mathscr{F}_{S^i}$ is a variation of pure Hodge structures of weight i on S. Such $\mathscr{F}$'s form a category $\mathscr{QH}_{\widetilde{S}}(S)$ which is a Tannakian category in an obvious manner. One has $\mathscr{QH}_S(S) = \mathscr{QH}(S)_\mathbb{Q}$. If $\pi': \widetilde{S}' \to S$ is another such local homeomorphism and $f\colon \widetilde{S}' \to \widetilde{S}$ is an S-morphism then we have an exact tensor functor $f^*\colon \mathcal{QH}_{\widetilde{S}}(S) \to \mathcal{QH}_{\widetilde{S}'}(S)$.

4.1.1. LEMMA. *f^* is a fully faithful embedding. In particular, $\pi^*\colon \mathcal{QH}(S)_{\mathbb{Q}} \to \mathcal{QH}_{\widetilde{S}}(S)$ is a fully faithful embedding.*

PROOF. Clearly f^* is faithful, and our lemma reduces easily to the following statement. Let $U \subset S$ be an open connected subset, $\mathcal{F}_1, \mathcal{F}_2 \in \mathcal{QH}(S)_{\mathbb{Q}}$, and $\psi_U\colon \mathcal{F}_{1U} \to \mathcal{F}_{2U}$ be a morphism over U such that $\mathrm{gr}^W \psi_U\colon \mathrm{gr}^W \mathcal{F}_{1U} \to \mathrm{gr}^W \mathcal{F}_{2U}$ extends to a morphism $\mathrm{gr}^W \psi$ over S. Then ψ_U extends to $\psi\colon \mathcal{F}_1 \to \mathcal{F}_2$. To prove this note that since

$$\mathrm{Hom}(\mathcal{F}_1, \mathcal{F}_2) = \mathrm{Hom}(\mathbb{Q}(0), \underline{\mathrm{Hom}}(\mathcal{F}_1, \mathcal{F}_2)),$$

we may assume that $\mathcal{F}_1 = \mathbb{Q}(0)_S$ (replacing $\mathcal{F}_2$ by $\mathcal{F} = \underline{\mathrm{Hom}}(\mathcal{F}_1, \mathcal{F}_2)$). Then $\mathrm{gr}^W \psi = \mathrm{gr}_0^W \psi \in \mathrm{Hom}(\mathbb{Q}(0)_S; W_0\mathcal{F}/W_{-1}\mathcal{F})$ and $\psi_U\colon \mathbb{Q}(0)_U \to W_0\mathcal{F}_U$ is a lifting of $\mathrm{gr}_0^W \psi$ on U. Clearly ψ_U is the unique section of the $\mathbb{Q}$-local system $W_0\mathcal{F}_{\mathbb{Q}}$ on U that projects to $\mathrm{gr}_0^W \psi$ and belongs to F^0. The extension ψ of ψ_U to S as a multi-valued section of $W_0\mathcal{F}_{\mathbb{Q}}$ also projects to $\mathrm{gr}_0^W \psi$ and belongs to F^0 (since F^0 is holomorphic). By uniqueness of such a lifting of $\mathrm{gr}_0^W \psi$ we see that it is actually single valued; i.e., $\psi \in \mathrm{Hom}(\mathbb{Q}(0)_S, \mathcal{F})$. □

So we may consider $\mathcal{QH}_{\widetilde{S}}(S)$ as an enlargement of the category $\mathcal{QH}(S)_{\mathbb{Q}}$ (and of $\mathcal{H}(S)_{\mathbb{Q}}$ if S is algebraic).

4.1.2. LEMMA. *Let $\mathcal{J} \in \mathcal{QH}_{\widetilde{S}}(S)$ be a quasi-Hodge sheaf such that $\mathcal{J} = W_{-2}\mathcal{J}$ and $F^0\mathcal{J} = 0$. Then the $\mathbb{Q}$-vector space $\mathrm{Ext}^1_{\mathcal{QH}_{\widetilde{S}}(S)}(\mathbb{Q}(0)_S, \mathcal{J})$ is canonically isomorphic to the space of sections g of $\mathcal{J}_{\mathcal{O}_{\widetilde{S}}}/\mathcal{J}_{\widetilde{S}\mathbb{Q}}$ such that $\partial g \in F^{-1}\mathcal{J}_{\widetilde{S}} \otimes \Omega^1_{\widetilde{S}}$.*

PROOF. Note that ∂ kills $\mathcal{J}_{\widetilde{S}\mathbb{Q}}$, so ∂g is well defined. Let g be such a section. For any (local) lifting $\tilde{g} \in \mathcal{J}_{\mathcal{O}_{\widetilde{S}}} = \mathcal{J}_{\widetilde{S}} \otimes \mathcal{O}_{\widetilde{S}}$ of g we have the extension $0 \to \mathcal{J}_{\widetilde{S}} \to P_{\tilde{g}} \to \mathbb{Q}(0)_{\widetilde{S}} \to 0$ where $P_{\tilde{g}\mathbb{Q}} - \mathcal{J}_{\widetilde{S}\mathbb{Q}} \oplus \mathbb{Q}_{\widetilde{S}}$, $W_0 P_{\tilde{g}} = P_{\tilde{g}}$, $W_{-i}P_{\tilde{g}} = W_{-i}\mathcal{J}_{\widetilde{S}}$ for $i \geq 1$, and $F^1 P_{\tilde{g}} = 0$, $F^0 P_{\tilde{g}} = \mathcal{O}_{\widetilde{S}}(\tilde{g} + 1)$, $F^{-i}P_{\tilde{g}} = F^0 P_{\tilde{g}} + F^{-i}\mathcal{J}_{\widetilde{S}}$ for $i \geq 1$. If $\tilde{g}'$ is another lifting of g, i.e., $\tilde{g}' = h + \tilde{g}$, $h \in \mathcal{J}_{\widetilde{S}\mathbb{Q}}$, we have a canonical isomorphism of extensions $P_{\tilde{g}} \xrightarrow{\sim} P_{\tilde{g}'}$, $1 \mapsto h+1$. This shows that we actually have a canonical extension $0 \to \mathcal{J}_{\widetilde{S}} \to P_g \to \mathbb{Q}(0)_{\widetilde{S}} \to 0$ that depends on g only. This is the desired extension of quasi-Hodge sheaves.

Conversely, for an extension $0 \to \mathcal{J}_{\widetilde{S}} \to P \to \mathbb{Q}(0)_{\widetilde{S}} \to 0$ the projection $F^0 P \to F^0\mathbb{Q}(0)_{\widetilde{S}} = \mathcal{O}_{\widetilde{S}}$ is an isomorphism. Denote by $\tilde{1}_F \in F^0 P$ the section that projects to $1 \in \mathcal{O}_{\widetilde{S}}$, and by $\tilde{1}_{\mathbb{Q}} \in P_{\mathbb{Q}}$ any local section that projects to $1 \in \mathbb{Q}$. Then $\tilde{g}_P = \tilde{1}_F - \tilde{1}_{\mathbb{Q}} \in \mathcal{I}_{\mathcal{O}_{\tilde{S}}}$, and $g_P = \tilde{g}_P \bmod \mathcal{I}_{\mathbb{Q}} \in \mathcal{I}_{\tilde{S}}/I_{\mathbb{Q}}$ is independent of the choice of $\tilde{1}_{\mathbb{Q}}$ and satisfies the conditions of 4.1.2. Clearly $P \mapsto g_P$, $g \mapsto P_g$ are mutually inverse one-to-one correspondences. □

4.2. The Schottky uniformization. Let us return to elliptic curves. Let $D = \{q \in \mathbb{C}^* : 0 < |q| < 1\} \subset \mathbb{C}^*$ be the coordinate punctured disc, $\tilde{X}_D := \mathbb{C}^* \times D$. For $n \in \mathbb{Z}$ denote by $(q)^n$ the automorphism $(z, q) \mapsto (q^n z, q)$ of $\tilde{X}_D$; this defines a $\mathbb{Z}$-action on $\tilde{X}_D$. Put $X_D := (q)^{\mathbb{Z}} \backslash \tilde{X}_D$. Then $p_{X_D} : X_D \to D$ is the standard family of elliptic curves over D. Note that for $q \in D$ the curve X_q is equipped with a canonical embedding $2\pi i \mathbb{Z} = H_1(\mathbb{C}^*, \mathbb{Z}) \hookrightarrow H_1(X_q, \mathbb{Z})$ such that $H_1(X_q, \mathbb{Z})/2\pi i \mathbb{Z}$ has no torsion and X_D/D is the universal family of elliptic curves equipped with such data.

Let $\mathcal{M}$ be the modular stack of elliptic curves, $p_{X_{\mathcal{M}}} : X_{\mathcal{M}} \to \mathcal{M}$ be the universal family of elliptic curves. The projection $\pi_D : D \to \mathcal{M}$ defined by X_D is an (infinite) unramified covering, hence so is the composition $\tilde{\pi}$ of the projections $\tilde{X}_D \to X_D \to X_{\mathcal{M}}$. Put $\tilde{U}_D := \tilde{\pi}^{-1}(U_{\mathcal{M}}) = \mathbb{C}^* \times D \backslash \{(q^n, q), n \in \mathbb{Z}\}$; the projection $\tilde{\pi}_U := \tilde{\pi}|_{\tilde{U}_D} \to U_{\mathcal{M}}$ is also an unramified covering.

We shall construct our elliptic polylogarithm sheaves as objects of the category $\mathcal{QH}_{\tilde{U}_D}(U)$.

4.3. The logarithm sheaf. First let us describe the sheaf $\mathcal{H} = \mathcal{H}_D = \pi_D^* \mathcal{H}_{\mathcal{M}}$ on D. By 4.2 we have a canonical embedding $2\pi i \mathbb{Q} \hookrightarrow \mathcal{H}_{\mathbb{Q}}$; denote by $e \in \mathcal{H}_{\mathbb{C}}$ the image of 1. Let f be the section of $F^0\mathcal{H}$ defined by the property $\langle e, f \rangle = 1$. The sections e, f form a basis of $\mathcal{H}_{\mathcal{O}_D}$; the connection is given by the formulas $\nabla e = 0$, $\nabla f = -e\,dq/q$. The $\mathbb{Q}$-lattice $\mathcal{H}_{\mathbb{Q}} \subset \mathcal{H}_{\mathcal{O}_D}$ is generated by $2\pi i e$ and the multi-valued horizontal section $f + (\log q)e$.

4.3.1. To describe the sheaf G_{X_D} on X_D, consider first the mixed sheaf $G^0_{\mathbb{C}^*}$ on $\mathbb{C}^*$ denoted as G_{X^∞} in 1.5.1. One has $\mathrm{gr}^W G^0_{\mathbb{C}} = \bigoplus_{k \geq 0} \mathbb{Q}(k)_{\mathbb{C}^*}$. An explicit construction of $G^0_{\mathbb{C}^*}$ goes as follows. One has $G^0_{\mathcal{O}_{\mathbb{C}^*}} = \prod_{k \geq 0} \mathcal{O}_{\mathbb{C}^*} e^k$, $F^j G^0_{\mathcal{O}_{\mathbb{C}^*}} = \bigoplus_{-j \geq k \geq 0} \mathcal{O}_{\mathbb{C}^*} e^k$, $W_{2n} G^0_{\mathcal{O}_{\mathbb{C}^*}} = W_{2n+1} G_{\mathcal{O}_{\mathbb{C}^*}} = \prod_{k \geq -n} \mathcal{O}_{\mathbb{C}^*} e^k$, $\nabla e^k = -e^{k+1} dz/z$ (where z is the parameter on $\mathbb{C}^*$). The $\mathbb{Q}$-structure is generated by the multi-valued sections $\tilde{e}^k := (2\pi i)^k (e^k + \log z e^{k+1} + \frac{(\log z)^2}{2} e^{k+2} + \cdots$ (i.e., $\tilde{e}^k$ is the horizontal section characterized by the property $\tilde{e}^k(1) = (2\pi i)^k e^k$). The mixed sheaf $G^0_{\mathbb{C}^*}$ is a module over the polynomial algebra $\mathrm{Sym}^\cdot(\mathbb{Q}(1))$—the action of the generator "$2\pi i$" of $\mathbb{Q}(1)$ is "$2\pi i$" $e^k = 2\pi i e^{k+1}$.

4.3.2. Let us define the mixed sheaf $G_{\tilde{X}_D} \in \mathcal{QH}(\tilde{X}_D)$ with $S^\cdot = \mathrm{Sym}^\cdot \mathcal{H}$-module structure. We have a canonical identification of sheaves of $S^\cdot_{\mathcal{O}_{\tilde{X}_D}}$-modules $G_{\mathcal{O}_{\tilde{X}_D}} = S^\cdot_{\mathcal{O}_{\tilde{X}_D}} := \prod S^i_{\mathcal{O}_{\tilde{X}_D}}$ compatible with weight and Hodge filtrations. The connection is given by the formula $\nabla_{G_{\tilde{X}_D}} = \nabla_{S^\cdot} - e\,dz/z$. The

$\mathbb{Q}$-structure $G_{\mathbb{Q}}$ on $G_{\mathbb{C}} := G^{\nabla}_{\mathscr{O}_{\widetilde{X}_D}}$ is characterized by the property that the fiber of $G_{\mathbb{Q}}$ at $z = 1$ coincides with $S^{\cdot}_{\mathbb{Q}} = \prod_{i\geq 0} S^i_{\mathbb{Q}}$. It is easy to see that these formulas actually define a mixed sheaf. The basis $\{e, f\}$ of $\mathscr{H}$ identifies $G_{\mathscr{O}_{\widetilde{X}_D}}$ with $\mathscr{O}_{\widetilde{X}_D}[[e, f]]$. Note that the subsheaf $G^0_{\mathscr{O}_{\widetilde{X}_D}} = \mathscr{O}_{\widetilde{X}_D}[[e]]$ is horizontal and comes from a $\mathbb{Q}$-sublocal system $G^0_{\mathbb{Q}}$; we have an obvious identification of $G^0_{\mathbb{Q}}$ with the pull-back of $G^0_{\mathbb{C}^*\mathbb{Q}}$ by the projection $\widetilde{X}_D = \mathbb{C}^* \times D \to \mathbb{C}^*$.

The action of the $(q)^n$-automorphisms of $\widetilde{X}_D$ lifts canonically to $G_{\widetilde{X}_D}$ by the formula $(q)^{n*}(e^k f^l) = e^k f^l \exp(nf)$; this action of $(q)^{\mathbb{Z}}$ is compatible with a quasi $\mathbb{Q}$-Hodge structure and commutes with the $S^{\cdot}$-action. The involution $[-1]_{\widetilde{X}_D}$, $(z, q) \mapsto (z^{-1}, q)$ also acts on $G_{\widetilde{X}_D}$ by the formula $[-1]_{\widetilde{X}_D}(e^k f^l) = (-1)^{k+l} e^k f^l$. These actions define a quasi-Hodge sheaf G_{X_D} on $X_D = (q)^{\mathbb{Z}} \backslash \widetilde{X}_D$ with an $S^{\cdot}$-module structure, equivariant with respect to the $[-1]_X$-involution.

4.3.3. Lemma. *G_{X_D} is canonically isomorphic to the sheaf G from* 1.2.8.

Proof. It is similar to 3.4.1. First note that $(G_{X_D})_0 = (G_{\widetilde{X}_D})_1 = \prod_{i\geq 0} S^i$ and that our canonical isomorphism $G \to G_{X_D}$ will be uniquely normalized by the property that it coincides with $\mathrm{id}_{\prod S^i}$ on the 0-fibers. Both for G, G_{X_D} we have a canonical isomorphism $G_?/W_{-n-1}G_? = \mathrm{Sym}^n G_?/W_{-2}G_?$; therefore, it suffices to identify $G_{X_D}/W_{-2}G_{X_D}$ with $G/W_{-2}G$. For $(z, q) \in \widetilde{X}_D$ one has $(G/W_{-2}G)_{(\hat{z}, q)} = H_1(X_q, \{0, \hat{z}\}; \mathbb{Q})$ where $\hat{z}$ is the image of z in X_q. The fiber $(G^0_{\widetilde{X}_D}/W_{-4}G^0_{\widetilde{X}_D})_{(z,q)}$ is $H_1(\mathbb{C}^*, \{1, z\}, \mathbb{Q})$. Therefore, we have a canonical embedding of $\mathbb{Q}$-vector spaces $(G^0_{\widetilde{X}_D}/W_{-4}G^0_{\widetilde{X}_D})_{\mathbb{Q}(z,q)} \hookrightarrow (G/W_{-2}G)_{\mathbb{Q}(\hat{z},q)}$ induced by the projection $\mathbb{C}^* \to X_q$. Since $G/W_{-2}G$ is an $S^{\cdot}$-module, the embedding extends to a morphism $(G_{\widetilde{X}_D}/W_{-2}G_{\widetilde{X}_D})_{\mathbb{Q}(z,q)} \to (G/W_{-2}G)_{\mathbb{Q}(\hat{z},q)}$ of $S^{\cdot}$-modules. One checks easily that this isomorphism is compatible with the Hodge structures and is $(q)^{\mathbb{Z}}$-equivariant. We get the desired isomorphism $G_{X_D}/W_{-2}G_{X_D} \xrightarrow{\sim} G/W_{-2}G$. □

4.3.4. Remark. Consider the embeddings $D \hookrightarrow \mathbb{C}^*_q$, $\widetilde{X}_D = \mathbb{C}^* \times D \hookrightarrow \mathbb{C}^*_z \times \mathbb{C}^*_q$. One identifies canonically $\mathbb{C}^*_q$ with the punctured tangent space at the parabolic point $q = 0$ and $\mathbb{C}^*_z \times \mathbb{C}^*_q$ with the punctured normal bundle to the degenerate Neron fiber $\mathbb{C}^*_z$ over the parabolic point. We have the corresponding specialization functors $\mathrm{Sp}_{q=0}$. According to the definition of $\mathrm{Sp}_{q=0}$ we have canonical isomorphisms of $\mathbb{Q}$-local systems $\mathscr{H}_{D\mathbb{Q}} \xrightarrow{\sim} \mathrm{Sp}_{q=0}\mathscr{H}|_D$, $G_{\widetilde{X}_D\mathbb{Q}} \xrightarrow{\sim} \mathrm{Sp}_{q=0}\, G|_{\widetilde{X}_D}$. These isomorphisms are compatible with the Hodge filtrations but not with the weight filtrations; the filtration W_0^{Sp} induced by the weight filtration on the specialized sheaves is given by the formulas

$W_0^{\mathrm{Sp}}\mathscr{H}_{D\mathbb{Q}} = \mathscr{H}_{D\mathbb{Q}}$, $W_{-1}^{\mathrm{Sp}}\mathscr{H}_{D\mathbb{Q}} = W_{-2}^{\mathrm{Sp}}\mathscr{H}_{D\mathbb{Q}} = 2\pi i\mathbb{Q} \hookrightarrow \mathscr{H}_{D\mathbb{Q}}$, $W_{-3}^{\mathrm{Sp}}\mathscr{H}_{D\mathbb{Q}} = 0$, and $W_0^{\mathrm{Sp}}G_{\widetilde{X}_D\mathbb{Q}}$ is the tensor product of the weight filtrations on $G^0_{\widetilde{X}_D}$ and $W_0^{\mathrm{Sp}}\mathscr{H}_D$ (i.e., $W_n^{\mathrm{Sp}}G_{\mathscr{O}_{\widetilde{X}_D}} = \prod_{2k\geq -n}\mathscr{O}_{\widetilde{X}_D}e^k f^l$).

4.4. The elliptic polylogarithm functions. In this section we define a q-averaged version of classical polylogarithm functions; cf. **[Z2]**.

4.4.1. For $m \geq 0$ put

$$\log_m(z, q) := \sum_{\substack{k, l\geq 0\\ k+1\geq m-l\geq 0}} \frac{m!(\log z)^{k+1-m+l}(\log q)^{m-l}}{l!(m-l)!(k+1-m+l)!}e^k f^l .$$

In the formula we choose the branches of $\log z$, $\log q$ that vanish at $z = 1$, $q = 1$ respectively. This is a multi-valued section of $G(1)_{\mathscr{O}_{\widetilde{X}_D}} = G_{\mathscr{O}_{\widetilde{X}_D}}$, and

$$\log_0(z) = \sum \frac{(\log z)^{k+1}}{(k+1)!}e^k$$

is actually a section of $G^0_{\mathscr{O}_{\mathbb{C}^*}} \subset G_{\mathscr{O}_{\widetilde{X}_D}}$. One easily checks that

$$\nabla_G(\log_m(z, q)) = f^m\frac{dz}{z} + mf^{m-1}\frac{dq}{q}$$

and that $\log_m$ is a well-defined (single-valued) section of $G(1)_{\mathscr{O}_{\widetilde{X}_D}}/G(1)_{\mathbb{Q}}$. So, by 4.1.2, these $\log_m$ define extensions of $\mathbb{Q}(0)_{\widetilde{X}_D}$ by $G(1)_{\widetilde{X}_D}$. It is easy to see that

$$(q)^{n*}\log_m = \sum_{a\geq 0}\frac{n^a}{a!}\log_{m+a}, \qquad [-1]^*\log_m = (-1)^{m+1}\log_m .$$

Note that for any n, almost all $\log_m$'s lie in $W_{-n}G(1)$. Therefore, for any $a_m \in \mathbb{Q}$ the series $\sum a_m \log_m$ converges to a section of $G(1)_{\mathscr{O}_{\widetilde{X}_D}}/G(1)_{\mathbb{Q}}$.

4.4.2. Consider the Euler polylogarithm section Li of $G^0_{\mathscr{O}_{\mathbb{C}^*}} = G^0(1)_{\mathscr{O}_{\mathbb{C}^*}}$ defined for $|z| < 1$ by the formula

$$Li(z) = \sum_{k\geq 0} Li_{k+1}(z)e^k = \sum_{\substack{k\geq 0\\ n\geq 1}}\frac{z^n}{n^{k+1}}e^k .$$

Then Li extends to $\mathbb{C}^*\backslash\{1\}$ as a multi-valued holomorphic section, and $\mathscr{L}i := Li \bmod G^0(1)_{\mathbb{Q}}$ is a well-defined (single-valued) section of $G^0(1)_{\mathscr{O}_{\mathbb{C}^*}}/G^0(1)_{\mathbb{Q}}$ on $\mathbb{C}^*\backslash\{1\}$. One has $\nabla\mathscr{L}i = -d\log(1-z)\cdot 1$ and $[-1]^*_{\mathbb{C}^*}\mathscr{L}i = \mathscr{L}i + \log_0$ (these sections have the same derivative, hence they differ by a global section of $G(1)_{\mathbb{C}}/G(1)_{\mathbb{Q}} = (\mathbb{C}/\mathbb{Q})\otimes G(1)_{\mathbb{Q}}$, and any such section is zero).

4.4.3. We shall consider Li as a multi-valued section of $G^0(1)_{\mathscr{O}_{\widetilde{X}_D}} \subset G(1)_{\mathscr{O}_{\widetilde{X}_D}}$ over $\widetilde{X}_D \setminus \{z=1\}$ and $\mathscr{L}i$ as a section of the corresponding quotient sheaf. Note that for $a \in \mathbb{Z}$ the series

$$\sum_{n\geq a} (q)^{n*} Li = \sum_{n\geq a, k\geq 0} Li_{k+1}(q^n z) e^k \exp(nf)$$

converges absolutely on $\widetilde{X}_D \setminus \bigcup_{j\geq a}\{z=q^{-n}\}$ if we demand that for almost all summands we take Li to be the principal branch of the polylogarithm (defined for $|z|<1$ by the series from 4.4.2; for given z one has $|q^j z|<1$ for almost all $j\geq 0$). Clearly $\sum_{n\geq a}(q)^{n*}\mathscr{L}i := \sum_{n\geq a}(q)^{n*}Li \bmod G(1)_{\mathbb{Q}}$ is a well-defined section of $G(1)_{\mathscr{O}_{\widetilde{X}_D}}/G(1)_{\mathbb{Q}}$ over $\widetilde{X}_D \setminus \bigcup_{j\geq a}\{z=q^{-n}\}$.

4.4.4. Consider now the series $\sum_{n\in\mathbb{Z}}(q)^{n*}\mathscr{L}i$ of sections of $G(1)_{\mathscr{O}_{\widetilde{X}_D}}/G(1)_{\mathbb{Q}}$ on $\widetilde{U}_D = \widetilde{X}_D \setminus \bigcup_{n\in\mathbb{Z}}\{z=q^n\}$. Note that one has

$$(q)^{-n*}\mathscr{L}i = [-1]_X^*(q)^{n*}(\mathscr{L}i + \log_0),$$

i.e., $\mathscr{L}i(q^{-n}z) = \mathscr{L}i(q^n z^{-1}) + \log_0(q^n z^{-1})$ by 4.4.2; so we may rewrite our series as

$$\begin{aligned}&\sum_{n\geq 0}(q)^{n*}\mathscr{L}i + [-1]_X^*\sum_{n\geq 1}(q)^{n*}\mathscr{L}i + [-1]_X^*\sum_{n\geq 1}(q)^{n*}\log_0\\ &\quad= \sum_{n\geq 0}\mathscr{L}i(q^n z) + \sum_{n\geq 1}\mathscr{L}i(q^n z^{-1}) + \sum_{n\geq 1}\log_0(q^n z^{-1}).\end{aligned}$$

The first two series converge by 4.4.3. The third series diverges, but we may rewrite it by 4.4.1 as

$$\sum_{n\geq 1}\log_0(q^n z^{-1}) = \sum_{\substack{n\geq 1\\ a\geq 0}}\frac{n^a}{a!}\log_a(z^{-1}) = \sum_{a\geq 0}\left(\sum_{n\geq 1} n^a\right)\frac{\log_a(z^{-1})}{a!}$$

and regularize it using ζ-function summation:

$$:\sum_{n\geq 1}\log_0(q^n z^{-1}): := \sum_{a\geq 0}\frac{\zeta(-a)}{a!}\log_a(z^{-1}) = -\sum_{a\geq 0}\frac{B_{a+1}}{(a+1)!}\log_a(z).$$

Here the B_a are the Bernoulli numbers; recall that $B_{a+1}=0$ for even $a>0$. Denote by $\mathscr{L}i^{\text{ell}}$ the regularized sum

$$\mathscr{L}i^{\text{ell}} :=: \sum_{n\in\mathbb{Z}}(q)^{n*}\mathscr{L}i := \sum_{n\geq 0}(q)^{n*}\mathscr{L}i + [-1]_X^*\sum_{n\geq 1}(q)^{n*}\mathscr{L}i - \sum_{a\geq 0}\frac{B_{a+1}}{(a+1)!}\log_a.$$

This is a well-defined section of $G(1)_{\mathscr{O}}/G(1)_{\mathbb{Q}}$ over $\widetilde{U}_D$.

4.4.5. LEMMA. *One has* $(q)^*\mathscr{L}i^{\text{ell}} - \mathscr{L}i^{\text{ell}} = \sum_{m\geq 0}\frac{\log_m}{(m+1)!}$.

PROOF. Straightforward computation. Denote by S_1, S_2, S_3 the summands in the definition of $\mathscr{L}i^{\text{ell}}$. One has $(q)^*S_1 - S_1 = -\mathscr{L}i(z)$, $(q)^*S_2 -$

$S_2 = \mathscr{L}i(z^{-1})$, and

$$(q)^* S_3 - S_3 = -\sum_{a\geq 0} \frac{B_{a+1}}{(a+1)!}(\log_a(qz) - \log_a(z))$$
$$= -\sum_{a,b\geq 1} \frac{B_a}{a!b!} \log_{a+b-1}(z)$$
$$= -\sum_{m\geq 1} \log_m(z) \left(\sum_{\substack{a,b\geq 1\\ a+b=m+1}} \frac{B_a}{a!b!} \right) = \sum_{m\geq 1} \frac{\log_m(z)}{(m+1)!}.$$

Hence 4.4.5 follows from 4.4.2. □

4.5. The auxiliary quasi-Hodge extensions.

4.5.1. By 4.1.2 the section $\mathscr{L}i^{\text{ell}}$ defines an extension

$$0 \to G(1)_{\widetilde{U}_D} \to \mathscr{F}^{\text{ell}} \to \mathbb{Q}(0)_{\widetilde{U}_D} \to 0$$

of quasi-$\mathbb{Q}$-Hodge sheaves on $\widetilde{U}$. Let

$$0 \to G(1)_{\widetilde{U}_D} \to \widetilde{\mathscr{F}}^{\text{ell}} \to p^*_{\widetilde{U}_D} \prod_{i\geq 0} S^i \to 0$$

be the extension deduced from $p^*_{\widetilde{U}_D} S^{\cdot} \otimes \mathscr{F}^{\text{ell}}$ by the push-out via the multiplication map $p^*_{\widetilde{U}} S^{\cdot} \otimes G(1)_{\widetilde{U}_D} \to G(1)_{\widetilde{U}_D}$. This is naturally an $S^{\cdot}$-module. Finally, let

$$0 \to G(1)_{\widetilde{U}_D} \to \widetilde{\mathscr{P}}' \to p^*_{\widetilde{U}_D} I \to 0$$

be the pull-back of $\widetilde{\mathscr{F}}^{\text{ell}}$ by $I = \prod_{i\geq 1} S^i \hookrightarrow S^{\cdot}$.

4.5.2. Now let us define the Koszul extension

$$0 \to G(1)_{\widetilde{X}_D} \to \widetilde{\mathscr{P}}'' \to p^*_{\widetilde{X}_D} I \to 0$$

of quasi-$\mathbb{Q}$-Hodge sheaves with $S^{\cdot}$-action on $\widetilde{X}_D$. First consider the usual Koszul short exact sequence $0 \to S^{\cdot}(1) \to S^{\cdot} \otimes \mathscr{H} \to I \to 0$ of quasi-Hodge $S^{\cdot}$-modules on D. We have a canonical identification of the short exact sequence of $S^{\cdot}_{\mathscr{O}_{\widetilde{X}_D}}$-modules $(0 \to G(1)_{\mathscr{O}_{\widetilde{X}_D}} \to \widetilde{\mathscr{P}}''_{\mathscr{O}_{\widetilde{X}_D}} \to I_{\mathscr{O}_{\widetilde{X}_D}} \to 0)$ with $\mathscr{O}_{\widetilde{X}_D} \otimes$ (usual Koszul sequence) such that the identification $G(1)_{\mathscr{O}_{\widetilde{X}_D}} = S^{\cdot}(1)_{\mathscr{O}_{\widetilde{X}_D}}$ coincides with the one of 4.3.2. The identification $\widetilde{\mathscr{P}}''_{\mathscr{O}_{\widetilde{X}_D}} = S^{\cdot} \otimes \mathscr{H}_{\mathscr{O}_{\widetilde{X}_D}}$ is compatible with the Hodge and weight filtrations. The connection $\nabla_{\widetilde{\mathscr{P}}''}$ is $\nabla_{S^{\cdot}\otimes\mathscr{H}} - \nu$ where ν is the $\text{End}_{S^{\cdot}_{\mathscr{O}_{\widetilde{X}_D}}}(S^{\cdot} \otimes \mathscr{H}_{\mathscr{O}_{\widetilde{X}_D}}) = (S^{\cdot} \otimes \text{End}\,\mathscr{H})_{\mathscr{O}_{\widetilde{X}_D}}$-valued 1-form defined by the formulas $\nu(1 \otimes e) = 0$, $\nu(1 \otimes f) = -\lambda\, dz/z$, where $\lambda := e \otimes f - f \otimes e$ is the generator of $S(1)_{\mathscr{O}}$ (i.e., one has $\nabla_{\widetilde{\mathscr{P}}''}(1 \otimes e) = 0$, $\nabla_{\widetilde{\mathscr{P}}''}(1 \otimes f) = -1 \otimes e\,dq/q - \lambda dz/z$). The $\mathbb{Q}$-structure $\widetilde{\mathscr{P}}''_{\mathbb{Q}} \subset \widetilde{\mathscr{P}}''_{\mathbb{C}} = \widetilde{\mathscr{P}}''^{\nabla}_{\mathscr{O}}$

is uniquely determined by the property that the fiber of $\widetilde{\mathscr{P}}''_{\mathbb{Q}}$ at $z = 1$ coincides with $(S^{\cdot} \otimes \mathscr{H})_{\mathbb{Q}}$. One checks easily that these formulas actually define an extension of quasi-Hodge sheaves.

4.6. The elliptic polylogarithm sheaf. Define the extension of quasi-Hodge $S^{\cdot}$-modules

$$0 \to G(1)_{\widetilde{U}_D} \to \widetilde{\mathscr{P}} \to p^*_{\widetilde{U}_D} I \to 0$$

on $\widetilde{U}_D$ as the Baer sum of the $\widetilde{\mathscr{P}}'$- and $\widetilde{\mathscr{P}}''$-extensions. The extension $0 \to \mathrm{gr}^W_{\cdot} G(1)_{\widetilde{U}_D} \to \mathrm{gr}^W_{\cdot} \widetilde{\mathscr{P}} \to \mathrm{gr}^W_{\cdot} p^*_{\widetilde{U}_D} \to 0$ is the usual Koszul extension from 4.5.2. To see this note that the extension $\mathrm{gr}^W_{\cdot} \mathscr{P}'$ splits canonically (since $0 \to \mathrm{gr}^W_{\cdot} G(1) \to \mathrm{gr}^W_{\cdot} \mathscr{F}^{\mathrm{ell}} \to \mathbb{Q}(0) \to 0$ splits) and that $\mathrm{gr}^W_{\cdot} \widetilde{\mathscr{P}}''$ coincides with the usual Koszul extension (since ν has negative weight).

In particular, we may consider $\widetilde{\mathscr{P}}$ as an extension in the category $\mathscr{QH}_{\widetilde{U}_D}(U)$ where U is the complement of zero on a universal elliptic curve (see 4.1).

4.6.1. PROPOSITION. *$\widetilde{\mathscr{P}}$ is canonically isomorphic to the elliptic polylogarithm extension from* 1.3.5, 1.3.13. *In particular,* $\widetilde{\mathscr{P}} \in \mathscr{H}(U)_{\mathbb{Q}} \subset \mathscr{QH}_{\widetilde{U}_D}(U)$.

PROOF. First let us check that $\widetilde{\mathscr{P}} \in \mathscr{QH}(U_D) \subset \mathscr{QH}_{\widetilde{U}}(U_D)$, i.e., that one has a canonical $(q)^{\mathbf{Z}}$-action on $\widetilde{\mathscr{P}}$. First consider the $(q)^{\mathbf{Z}}$-action on the $S^{\cdot}_{\mathscr{O}_{\widetilde{X}_D}}$-module $\widetilde{\mathscr{P}}''_{\mathscr{O}_{\widetilde{X}_D}} = S^{\cdot} \otimes \mathscr{H}_{\mathscr{O}_{\widetilde{X}_D}}$ defined by the formulas $(q)^{n*}(1 \otimes f) = 1 \otimes f$, $(q)^{n*}(\lambda) = \exp(nf)\lambda$. Then

$$(q)^{n*}(1 \otimes e) = \exp(nf) \otimes e + \frac{1 - \exp(nf)}{f} e \otimes f$$

and the $(q)^{\mathbf{Z}}$-action on $G(1) \subset \widetilde{\mathscr{P}}''$ coincides with the one from 4.3.2 and is the identity on the quotient $I_{\mathscr{O}_{\widetilde{X}_D}}$. This action also preserves the Hodge and weight filtrations but does not preserve the connection. A straightforward computation shows that for $\varphi \in S^{\cdot}$, $h \in \mathscr{H}$ one has

$$(q)^* \nabla_{\widetilde{\mathscr{P}}''}(\varphi \otimes h) - \nabla_{\widetilde{\mathscr{P}}''}(q)^*(\varphi \otimes h) = -\sum_{n \geq 0} \frac{1}{(n+1)!} \frac{dz}{z} + \frac{(n+1)}{(n+2)!} \frac{dq}{q} f^n \varphi h \lambda.$$

Now consider a $(q)^{\mathbf{Z}}$-action of $\mathscr{F}^{\mathrm{ell}}_{\mathscr{O}_{\widetilde{U}_D}}$ that coincides with the usual action on the submodule $G(1)_{\mathscr{O}_{\widetilde{U}_D}}$ and the quotient module $\mathscr{O}_{\widetilde{U}_D}$ and preserves the weight and Hodge filtrations. Since $\mathscr{F}^{\mathrm{ell}}_{\mathscr{O}_{\widetilde{U}_D}} = F^0 \mathscr{F}^{\mathrm{ell}} \oplus G(1)_{\mathscr{O}_{\widetilde{U}_D}}$, such an action exists and is unique. It extends by $S^{\cdot}$-linearity to a $(q)^{\mathbf{Z}}$-action $\widetilde{\mathscr{P}}'_{\mathscr{O}_{\widetilde{U}_D}}$. We define the $(q)^{\mathbf{Z}}$-action on $\widetilde{\mathscr{P}}_{\mathscr{O}_{\widetilde{U}_D}}$ as the Baer sum of the above actions. It preserves the weight and Hodge filtrations and is horizontal by 4.4.5, 4.4.1, and the above computation. Since

$$\mathrm{Hom}_{S^{\cdot}_{\mathbb{Q}}}(I_{\mathbb{Q}}, (\mathbb{C}/\mathbb{Q}) \underset{\mathbb{Q}}{\otimes} G_{\mathbb{Q}}(1)) = (\mathbb{C}/\mathbb{Q}) \underset{\mathbb{Q}}{\otimes} G_{\mathbb{Q}}(1)$$

has no sections on $\widetilde{U}_D$, we see that this action also preserves the $\mathbb{Q}$-structure. Hence $\widetilde{\mathscr{P}}$ is a $(q)^{\mathbb{Z}}$-equivariant quasi-Hodge sheaf; i.e., $\widetilde{\mathscr{P}} \in \mathscr{QH}(U_D)$.

The restriction $\widetilde{\mathscr{P}}_q$ of $\widetilde{\mathscr{P}}$ to any fiber X_q is a Hodge sheaf. To prove this, one needs to check the asymptotic conditions at $z = 1$. But in a neighbourhood of $z = 1$ $\widetilde{\mathscr{P}}_q$ is the Baer sum of some extension regular at $z = 1$ and an extension derived from the classical polylogarithm $\mathscr{L}i$ by the standard operations; see 4.5. This implies our assertion.

By 1.3.4 the extension $\widetilde{\mathscr{P}}_q$ is canonically isomorphic to the elliptic polylogarithm extension from 1.3.5. Since this fiberwise isomorphism preserves the $\mathbb{Q}$-structure, it is horizontal, i.e., it is actually an isomorphism of quasi Hodge sheaves on U_D. □

4.7. REMARK. The pull-back of $\widetilde{\mathscr{P}}$ with respect to $\mathscr{H} \hookrightarrow I$ gives an explicit construction of the $\mathbb{Q}$-Hodge version of the sheaf $\mathscr{P}$ from 1.3.13. It is easy to see that the specialization of $\mathscr{P}$ to $q = 0$ is essentially the classical polylogarithm sheaf (cf. 1.5), and we see that $\mathscr{P}$ could be reconstructed from $\mathrm{Sp}_{q=0}\mathscr{P}$ via a "despecialization" construction; first change the weight filtration and then make an appropriately regularized q-averaging (cf. 4.3.4).

4.8. In the rest of this section we sketch some formulas for the periods of the elliptic polylogarithm sheaf. Recall that the period matrix expresses a holomorphic basis compatible with the Hodge filtration via a (horizontal) $\mathbb{Q}$-basis compatible with the weight filtration. For example, a period matrix of the classical polylog sheaf is

$$\begin{pmatrix} 1 & & & \\ \Lambda_1(z) & 1 & & 0 \\ \Lambda_2(z) & \frac{\log(z)}{2\pi i} & 1 & \\ \Lambda_3(z) & \frac{1}{2}\frac{\log^2(z)}{2\pi i} & \frac{\log(z)}{2\pi i} & 1 \\ \cdots & \cdots & \cdots & \end{pmatrix}$$

where the

$$\Lambda_k(z) := \frac{1}{(k-1)!}\int_0^z \left(\frac{\log t}{2\pi i}\right)^{k-1} \frac{d\log(1-t)}{2\pi i}$$

are versions of Euler's polylogarithms (we call them *Debye polylogarithms* since, up to a constant, $\Lambda_k(e^{\xi})$ coincide with the functions

$$\int \xi^{n-1} d\xi/(1-e^{-\xi})$$

that appear naturally in solid state physics). One has

$$Li_k(z) = \sum_{k\geq i\geq 1} \frac{(\log z)^{k-i}}{(k-i)!}(-2\pi i)^i \Lambda_i(z),$$

$$\Lambda_k(z) = (-2\pi i)^{-k} \sum_{k\geq i\geq 1} \frac{(-\log z)^{k-i}}{(k-i)!} Li_i(z).$$

We define the elliptic Debye polylogarithm functions of index (m, n) as a ζ-regularization of the divergent series $\frac{1}{m!}\sum_{j\in\mathbb{Z}} j^m\Lambda_n(zq^j)$:

$$\Lambda_{m,n}(\xi, \tau) = \frac{1}{m!}\left\{\sum_{j\geq 0} j^m\Lambda_n(q^j z) + (-1)^{m+n+1}\sum_{j\geq 1} j^m\Lambda_n(q^j z^{-1}) + (-1)^{m+n}\sum_{n\geq k\geq 0}\frac{(-\xi)^{n-k}\tau^k}{(n-k)!k!}\zeta(-m-k) + (-1)^{m+n+1}\frac{B_n}{n!}\zeta(-m)\right\}.$$

Here $\xi = \frac{1}{2\pi i}\log z$, $\tau = \frac{1}{2\pi i}\log q$, and we regularize the series using the functional equation

$$\Lambda_k(z) + (-1)^k\Lambda_k(z^{-1}) = \frac{1}{k!}\left\{\left(\frac{\log z}{2\pi i}\right)^k - (-1)^k B_k\right\}$$

and ζ-regularization: $:\sum_{k\geq 1} k^a :: = \zeta(-a)$ (cf. 6.4.4). These $\Lambda_{m,n}$ are multi-valued holomorphic functions of (ξ, τ), $\operatorname{Im}\tau > 0$.

Note that $\Lambda_{0,1}(\xi, \tau) = \log\theta[{}^{1/2}_{1/2}](\xi, \tau) - \log\eta(\tau)$. The functions $\Lambda_{m,n}$ have the following transformation properties:

$$\Lambda_{m,n}(\xi + 1, \tau) = \sum_{n-1\geq i\geq 0}\frac{1}{i!}\Lambda_{m,n-i}(\xi, \tau) + \frac{B_{m+1}}{(m+1)!(n-1)!},$$

$$\Lambda_{m,n}(\xi + \tau, \tau) = \sum_{m\geq i\geq 0}\frac{(-1)^i}{i!}\Lambda_{m-i,n}(\xi, \tau) + (-1)^{m+1}\sum_{n\geq l\geq 0}\frac{\xi^{n-l}\tau^l}{(n-l)!(m+l+1)!} + (-1)^{m+n}\frac{B_{m+1}}{n!(m+1)},$$

$$\Lambda_{m,n}(-\xi, \tau) = (-1)^{m+n+1}\Lambda_{m,n}(\xi, \tau) + (1 + (-1)^{m+n})(-1)^{n+1}\frac{B_n B_{m+1}}{n!(m+1)!} + \delta_{m,0}(-1)^{n+1},$$

$$\Lambda_{m,n}(\xi, \tau + 1) = \sum_{n-1\geq i\geq 0}\binom{m+i}{m}\Lambda_{m+i,n-i}(\xi, \tau) + r_{m,n} \text{ where } r_{m,n}\in\mathbb{Q},$$

$$\Lambda_{m,n}\left(\xi/\tau, -\frac{1}{\tau}\right) = (-1)^{n-1}\Lambda_{n-1,m+1}(\xi, \tau) + \frac{\xi^{m+n+1}}{\tau^n(m+n+1)!} + t_{m,n} \quad \text{where } t_{m,n}\in\mathbb{Q}.$$

Also one has $(n-1)d\Lambda_{m,n} = (m+1)\tau d\Lambda_{m+1,n-1}(\xi, \tau) + \xi d\Lambda_{m,n-1}(\xi, \tau)$. A period matrix for the sheaf G is the matrix L of the linear transformation $\exp(-\xi t_1 - \tau t_1\partial_{t_2})$ acting on the space $\mathbb{C}[[t_1, t_2]]$ with respect to the basis

$t_1^a t_2^b$. A period matrix for the sheaf $\mathcal{F}^{\text{ell}}$ from 4.5 is

$$\begin{pmatrix} 1 & 0 \\ \sum \Lambda_{m,n} t_1^{n-1} t_2^m & L \end{pmatrix}.$$

5. Semisimple motives generated by an elliptic curve and their absolute cohomology groups

Below for a regular scheme S we denote by $H^i_{\mathcal{M}}(S, \mathbb{Q}(j))$ the absolute motivic cohomology group defined via K-theory —this is the p^j-eigenspace of the Adams operator Ψ^p acting on $K_{2j-i}(S) \otimes \mathbb{Q}$. In particular, when S is a smooth scheme over a field, $H^{2i}_{\mathcal{M}}(S, \mathbb{Q}(i))$ is the Chow group $\mathrm{CH}^i(S)_{\mathbb{Q}}$ of codimension i cycles modulo rational equivalence. For a morphism $f: S_1 \to S_2$ we have the pull-back map $f^*: H^i_{\mathcal{M}}(S_2, \mathbb{Q}(j)) \to H^i_{\mathcal{M}}(S_1, \mathbb{Q}(j))$. If f is a proper morphism and $d = \dim S_2 - \dim S_1$, then one has the Gysin map $f_*: H^i_{\mathcal{M}}(S_1, \mathbb{Q}(j)) \to H^{i+2d}_{\mathcal{M}}(S_2, \mathbb{Q}(j+d))$.

In this section we describe a motivic decomposition of the absolute motivic cohomology of a power of our elliptic curve X. First we shall give a definition of a tensor category $\mathcal{M}^{(X)}$ of motives generated by X. This is a semisimple Tannakian $\mathbb{Q}$-category with Tate twists, and for any abelian variety A isogenous to a power of X we have motivic cohomology objects $H^i(A) \in \mathcal{M}^{(X)}$ that behave like the usual cohomology groups. One has the usual array of realization functors on $\mathcal{M}^{(X)}$. In particular, if X is a curve over $\mathbb{C}$, we have a canonical identification of $\mathcal{M}^{(X)}$ with the tensor subcategory of the category of $\mathbb{Q}$-Hodge structures $\mathcal{H}_{\mathbb{Q}}$ generated by the Hodge structure $H^1(X, \mathbb{Q})$. Then we define the absolute cohomology functors $H^i_{\mathcal{M}}$ on $\mathcal{M}^{(X)}$; for any A as above we have the motivic decomposition formula $H^i_{\mathcal{M}}(A, \mathbb{Q}(j)) = \bigoplus H^{i-a}_{\mathcal{M}}(H^a(A)(j))$. The constructions are based on the fact that for the "linear" cycles on A—the ones that are linear combinations of abelian subvarieties—homological equivalence coincides with the rational one.

5.1. Action of isogenies on the absolute cohomology. Below B is a fixed base scheme that we assume to be regular, Noetherian, and connected, $p_X: X \to B$ is an elliptic curve over B, $0 = 0_X \in X(B)$ is the zero section. Put $R = R_X := \operatorname{End} X$, $E = E_X := R \otimes \mathbb{Q}$. Therefore, E is either $\mathbb{Q}$, or an imaginary quadratic field, or a quaternion algebra over $\mathbb{Q}$ (if X is a supersingular curve) and $R \subset E$ is a ring of integers. For $r \in R$ we denote its action on X as $[r]: X \to X$.

Let m be a nonzero integer. Consider the maps

$$H^{\cdot}_{\mathcal{M}}(X, \mathbb{Q}(*)) \underset{[m]^*}{\overset{[m]_*}{\rightleftarrows}} \mathcal{H}^{\cdot}_{\mathcal{M}}(X, \mathbb{Q}(*)).$$

5.1.1. LEMMA. *Both compositions* $[m]_*[m]^*$, $[m]^*[m]_*$ *are multiplication by* m^2.

PROOF. (i) $[m]_*[m]^* = m^2 \operatorname{id}_H$. This is a general fact; for any finite flat degree d morphism $\varphi\colon Y_1 \to Y_2$ of regular schemes one has $\varphi_*\varphi^* = d \cdot \operatorname{id}_{H^{\cdot}_{\mathcal{M}}} \in \operatorname{End} H^{\cdot}_{\mathcal{M}}(Y_2, \mathbb{Q}(*))$ (since $\varphi_*\varphi^*(a) = a\varphi_*\varphi^*(1)$ by the projection formula, and $\varphi_*\varphi^*(1) = \varphi_*(1) = d$).

(ii) $[m]^*[m]_* = m^2 \operatorname{id}_{H^{\cdot}_{\mathcal{M}}}$. Let $X_m = \operatorname{Ker}[m] \subset X$ be the subgroup of points of order m, and let $c(X_m), c(0) \in H^2_{\mathcal{M}}(X, \mathbb{Q}(1)) = \operatorname{Pic}(X)_{\mathbb{Q}}$ be the classes of the divisors X_m and $0 = 0(B) = X_1$, respectively. Clearly $c(X_m) = [m]^*(c(0))$. Also $[m]^*[m]_*$ is the X_m-averaging map; for $\alpha \in H^{\cdot}_{\mathcal{M}}(X, \mathbb{Q}(*))$ one has $[m]^*[m]_*(\alpha) = +_*(\operatorname{pr}_1^*(c(X_m)) \cdot \operatorname{pr}_2^*(\alpha))$, where $\operatorname{pr}_i, +\colon X \times_B X \to X$ are the projections and the sum map, respectively. Therefore, our claim will follow if we show that $[m]^*c(0) = m^2c(0)$. To see this, consider the filtration $P_0 \subset P_1 \subset P_2 = \operatorname{Pic}(X)_{\mathbb{Q}}$, where $P_0 = p_X^*(\operatorname{Pic}(B)_{\mathbb{Q}})$ and P_1 consists of those divisors that have degree 0 along the fibers of p_X so that $P_1/P_0 = X(B)_{\mathbb{Q}}$, $P_2/P_1 = \mathbb{Q}$. The endomorphisms $[n]^*$, $n \in \mathbb{Z}$, preserve this filtration and act as n^i on P_i/P_{i-1}. Therefore, the filtration splits canonically by the (common) eigenspaces of the $[n]^*$; so we have the isomorphism

$$\bigoplus P_i/P_{i-1} = \operatorname{Pic}(B)_{\mathbb{Q}} \oplus X(B)_{\mathbb{Q}} \oplus \mathbb{Q} \xrightarrow{\sim} \operatorname{Pic}(X)_{\mathbb{Q}}.$$

We must show that $c(0) \in P_2/P_1 \subset \operatorname{Pic}(X)_{\mathbb{Q}}$. Clearly $P_0 \oplus P_2/P_1 \subset \operatorname{Pic}(X)_{\mathbb{Q}}$ is the $+1$-eigenspace of the involution $[-1]^*$ and the projector $\operatorname{Pic}(X)_{\mathbb{Q}} \to P_0 = \operatorname{Pic}(B)_{\mathbb{Q}}$ is 0^*. Since $[-1]^*c(0) = c(0)$, it remains to show that $0^*c(0) \in \operatorname{Pic}(B)_{\mathbb{Q}}$ is zero. This is the class of the modular line bundle $T = 0^*\mathcal{T}_{X/B}$, and $T^{\otimes 12}$ is canonically trivialized by the discriminant section Δ. □

Now let A_1, A_2 be abelian schemes over B isogenous to some power $X^n = X \times_B \cdots \times_B X$ (n times) of X, and let $\varphi\colon A_1 \to A_2$ be an isogeny of degree d. Consider the morphisms

$$H^{\cdot}_{\mathcal{M}}(A_1, \mathbb{Q}(*)) \underset{\varphi^*}{\overset{\varphi_*}{\rightleftarrows}} H^{\cdot}_{\mathcal{M}}(A_2, \mathbb{Q}(*)).$$

5.1.2. COROLLARY. *Both the compositions* $\varphi_*\varphi^*$, $\varphi^*\varphi_*$ *are multiplication by* d; *in particular,* φ^* *and* φ_* *are isomorphisms.*

PROOF. As before (see the first part of the proof of 5.1.1) $\varphi_*\varphi^* = d \operatorname{id}_{H^{\cdot}_{\mathcal{M}}}$ by general reasons. To prove that $\varphi^*\varphi_* = d \operatorname{id}_{H^{\cdot}_{\mathcal{M}}}$ it suffices to check that

φ^* is an isomorphism. We already know that φ^* is injective. Choose any isogeny $\chi: X^n \to A_1$; put $m = \deg(\varphi\chi)$. Then the multiplication by the m-endomorphism $[m]_{X^n} \in \operatorname{End} X^n$ kills $\operatorname{Ker}(\varphi\chi)$; hence, we have a commutative diagram of isogenies

$$\begin{array}{ccc} A_1 & \xrightarrow{\varphi} & A_2 \\ \uparrow{\scriptstyle\chi} & & \downarrow{\scriptstyle\psi} \\ X^n & \xrightarrow{[m]_{X^n}} & X^n. \end{array}$$

Note that $[m]^* = \chi^*\varphi^*\psi^*$ is an isomorphism (one has $[m]_{X^m} = [m]_1 \circ \cdots \circ [m]_n$ where $[m]_i$ is multiplication by m along the ith factor; but $[m]_i^*$ is an isomorphism by 5.1.1 applied to X^n considered as an elliptic curve over X^{n-1} via the projection $(x_1, \dots, x_n) \mapsto (x_1, \dots, \hat{x}_i, \dots, x_n)$). Since χ^*, φ^*, and ψ^* are injective, they are isomorphisms. □

5.1.3. Let $Ab^{(X)}$ be the category of abelian schemes over B isogenous to a power of X, and let $Ab^{(X)}_{\mathbb{Q}}$ be the corresponding category of abelian schemes modulo isogeny. Then $Ab^{(X)}$ is an additive category, $Ab^{(X)}_{\mathbb{Q}}$ is a semisimple abelian $\mathbb{Q}$-category, and we have a canonical functor $Ab^{(X)} \to Ab^{(X)}_{\mathbb{Q}}$, $A \mapsto A_{\mathbb{Q}}$. Denote by E-mod, mod-E the categories of finite-dimensional left, resp. right, E-modules. The functor $\operatorname{Hom}(X_{\mathbb{Q}}, \cdot): Ab^{(X)}_{\mathbb{Q}} \to \text{mod-}E$ is an equivalence of categories; the inverse equivalence is $V \mapsto V \otimes X := V \otimes_E X$.

Consider $H^{\cdot}_{\mathscr{M}}(\cdot, \mathbb{Q}(*))$ as a contravariant functor on $Ab^{(X)}$ with values in the category of bigraded $\mathbb{Q}$-algebras. Since, by 5.1.2, the pull-back map for an isogeny is an isomorphism, the absolute cohomology functor extends canonically to $Ab^{(X)}_{\mathbb{Q}}$, i.e., we have the contravariant functor $H^{\cdot}_{\mathscr{M}}(\cdot, \mathbb{Q}(*))$ on $Ab^{(X)}_{\mathbb{Q}}$ together with a canonical identification $H^{\cdot}_{\mathscr{M}}(A, \mathbb{Q}(*)) = H^{\cdot}_{\mathscr{M}}(A_{\mathbb{Q}}, \mathbb{Q}(*))$ for $A \in Ab^{(X)}$.

5.1.4. Any morphism in $Ab^{(X)}$ is proper; therefore, the Gysin maps provide a covariant functoriality for the absolute cohomology-groups on $Ab^{(X)}$. To extend the Gysin functoriality to $Ab^{(X)}_{\mathbb{Q}}$ one must twist the $H^{\cdot}_{\mathscr{M}}$-groups. Namely, consider a rule that assigns to any $A_{\mathbb{Q}} \in Ab^{(X)}_{\mathbb{Q}}$ a $\mathbb{Q}$-line (i.e., a one-dimensional $\mathbb{Q}$-vector space) $\lambda_{A_{\mathbb{Q}}}$ and to any $A \in Ab^{(X)}$ a nonzero element $\lambda(A) \in \lambda_{A_{\mathbb{Q}}}$, such that $\lambda_{A_{\mathbb{Q}}}$ is functorial with respect to isomorphisms in $Ab^{(X)}_{\mathbb{Q}}$, and such that for any isogeny $\varphi: A_1 \to A_2$ in $Ab^{(X)}$ the corresponding isomorphism $\lambda_{\varphi_{\mathbb{Q}}}: \lambda_{A_{1\mathbb{Q}}} \to \lambda_{A_{2\mathbb{Q}}}$ sends $\lambda(A_1)$ to $\deg\varphi\lambda(A_2)$. It is easy to see that such a rule exists and is unique in an obvious sense. Now, as follows from 5.1.2, there is a unique way to provide a covariant Gysin $Ab^{(X)}_{\mathbb{Q}}$-functoriality for the groups $H^{\cdot}_{\mathscr{M}}(A_{\mathbb{Q}}, \mathbb{Q}(*)) \otimes \lambda_{A_{\mathbb{Q}}}$, $A_{\mathbb{Q}} \in Ab^{(X)}_{\mathbb{Q}}$, such that for any morphism $\varphi\colon A_1 \to A_2$ in $Ab^{(X)}$ the diagram

$$\begin{array}{ccc} H^{\cdot}_{\mathcal{M}}(A_{1\mathbb{Q}}, \mathbb{Q}(*)) \otimes \lambda_{A_{1\mathbb{Q}}} & \xrightarrow{\varphi_{\mathbb{Q}*}} & H^{\cdot+2d}_{\mathcal{M}}(A_{2\mathbb{Q}}, \mathbb{Q}(*+d)) \otimes \lambda_{A_2\mathbb{Q}} \\ \uparrow{\scriptstyle \otimes\lambda(A_1)} & & \uparrow{\scriptstyle \otimes\lambda(A_2)} \\ H^{\cdot}_{\mathcal{M}}(A_1, \mathbb{Q}(*)) & \xrightarrow{\varphi_*} & H^{\cdot+2d}_{\mathcal{M}}(A_2, \mathbb{Q}(*+d)) \end{array}$$

commutes; here $d = \dim A_2 - \dim A_1$.

5.1.5. Here is another description of the $\lambda_{A_\mathbb{Q}}$'s. We denote by $\lambda\colon \mathrm{GL}(n, E) \to \mathbb{Q}^*$ the character defined by

$$\lambda(g) := \begin{cases} \det^2(g), & E = \mathbb{Q}, \\ Nm_{E/\mathbb{Q}} \det(g), & E \text{ is an imaginary quadratic field,} \\ \text{Dieudonné determinant of } g, & E \text{ is a quaternion algebra.} \end{cases}$$

There is a unique way to assign to every $V \in$ mod-E a $\mathbb{Q}$-line λ_V functorial with respect to isomorphisms in mod-E, together with trivializations $\mathbb{Q} \xrightarrow{\sim} \lambda_{E^n}$, $n \geq 0$, such that $g \in \mathrm{GL}(n, E) = \mathrm{Aut}(A^n)$ acts on λ_{E^n} as multiplication by $\lambda(g)$. There is a unique functorial isomorphism $\lambda_{V_1 \otimes V_2} = \lambda_{V_1} \otimes \lambda_{V_2}$ compatible with the trivializations (for the standard isomorphisms $E^{n_1} \otimes E^{n_2} \xrightarrow{\sim} E^{n_1 n_2}$; note that the problem of signs does not occur).

Now we have a unique functorial isomorphism

$$\lambda_{V \otimes X_\mathbb{Q}} = \lambda_{X_\mathbb{Q}}^{\otimes n} \otimes \lambda_V,$$

where $n = \dim_E V$, that sends $\lambda(X^n) \in \lambda_{X^n_\mathbb{Q}} = \lambda_{E^n \otimes X_\mathbb{Q}}$ to $\lambda(X)^{\otimes n}$ (with respect to the trivialization $\lambda_{E^n} = \mathbb{Q}$). This isomorphism depends only on $X_\mathbb{Q}$ (and not on X).

5.1.6. Let $b \in B$ be a geometric point, ℓ be a prime $\neq \operatorname{char} b$. We have a canonical isomorphism

$$\lambda_{\mathbb{Q}_\ell} := \lambda \otimes \mathbb{Q}_\ell \xrightarrow{\sim} H^{2n}(A_\mathbb{Q}, \mathbb{Q}_\ell(n))' = (\det H^1(A_\mathbb{Q}, \mathbb{Q}_\ell))'(-n),$$

$n = \dim A_\mathbb{Q}$, that sends $\lambda(A)$ to the trace functional.

5.2. A linear algebra digression. In this subsection we shall assign to any $V \in$ mod-E a certain commutative graded finite-dimensional $\mathbb{Q}$-algebra $C^{\cdot}(V)$; for a geometric meaning of $C^{\cdot}(V)$ see 5.3.2.

We shall consider $\mathrm{Aut}(V)$, $V \in$ mod-E, as the set of $\mathbb{Q}$-points of a reductive algebraic group over $\mathbb{Q}$. In particular, we have a notion of an algebraic action of $\mathrm{Aut}(V)$ on a finite-dimensional F-vector space where F is any field of char 0. For example, the action of $\mathrm{Aut}(V)$ on λ_V is algebraic.

5.2.1. Fix an integer $i \geq 0$. Let C^i be a contravariant functor on mod-E with values in the category of $\mathbb{Q}$-vector spaces. Assume that for any $V \in$ mod-E of dimension i we have an isomorphism $\sigma\colon \lambda'_V \xrightarrow{\sim} C^i(V)$ natural with respect to isomorphisms of such V's. Consider the following properties (where for a morphism $f\colon V_1 \to V_2$ we put $f^* = C^i(f)\colon C^i(V_2) \to C^i(V_1)$):

(a) For any $V \in$ mod-E one has $\dim C^i(V) < \infty$ and the action of $\mathrm{Aut}(V)$ on $C^i(V)$ is algebraic.

(b) For $V \in$ mod-E the $\mathbb{Q}$-vector space $C^i(V)$ is generated by the images of lines $p_U^*: \lambda'_U = C^i(U) \to C^i(V)$ where $p_U: V \to U$ are morphisms with $\dim U = i$.

5.2.2. LEMMA. (i) *Such* (C^i, σ) *exists and is unique* (*up to a unique isomorphism*).

(ii) *For every* $V \in$ mod-E, $C^i(V)$ *is an irreducible* $\mathrm{Aut}(V)$*-module.*

PROOF. Assume that (C^i, σ) exists. Note that $C^i(V) = 0$ if $\dim V < i$, by axiom (b) (since for any morphism $p_U: V \to U$, $\dim U = i$, there exists $g \in \mathrm{Aut}(U)$ with $\lambda(g) \neq 1$ such that $gp_U = p_U$, hence $p_U^* = \lambda(g)p_U^*$ and $p_U^* = 0$). Assume that $\dim V \geq i$. We see that in (b) we may consider only surjective p_U's; for these the map p_U^* is injective (since p_U has a right inverse). Equivalently, $C^i(V)$ is generated by the set of lines $\lambda'_{V/W} \hookrightarrow C^i(V)$, where the $W \subset V$ are codimension i subspaces. For such $W \subset V$ let P_W be the parabolic subgroup of $\mathrm{Aut}(V)$ that preserves W. Since $\mathrm{Aut}(V)$ acts transitively on the set of all W's, we see that $C^i(V)$ is generated as an $\mathrm{Aut}(V)$-module by any line $\lambda'_{V/W}$; this line is preserved by the parabolic P_W that acts on it via the projection $P_W \to \mathrm{Aut}(V/W)$. By (a) we may use the highest weight theory. It says that $C^i(V)$ is an irreducible $\mathrm{Aut}(V)$-module. If $(C^{i\sim}, \sigma^\sim)$ is another such datum, then there exists a unique isomorphism $C^i(V) \xrightarrow{\sim} C^i(V)^\sim$ of $\mathrm{Aut}(V)$-modules that induces the identity on the lines $\lambda_{V/W}$. This isomorphism is functorial (since for any morphism $\varphi: V_1 \to V_2$ we know how $\varphi^*: C^i(V_2) \to C^i(V_1)$ acts on the lines $\lambda'_{V_2/W}$), which proves the unicity statement.

Let us prove the existence. For $V \in$ mod-E we define $C^i(V)$ as the unique irreducible algebraic $\mathrm{Aut}(V)$-module equipped with a system of embeddings $\lambda'_{V/W} \hookrightarrow C^i(V)$ natural with respect to the action of $\mathrm{Aut}(V)$. Such $C^i(V)$ exists by the highest weight theory since P_W acts on $\lambda'_{V/W}$ by a dominant weight. For a morphism $\varphi: V_1 \to V_2$ the corresponding $\varphi^*: C^i(V_2) \to C^i(V_1)$ is uniquely determined by its action on the lines, so it suffices to construct φ^* for injective and surjective φ's.

Let φ be injective. We may assume that $\dim V_1 \geq i$. Let $P_\varphi \subset \mathrm{Aut}\, V_2$ be the parabolic subgroup that preserves $\varphi(V_1)$, and let N_φ be the nilradical of P_φ; then we have the projections $P_\varphi \to P_\varphi/N_\varphi \twoheadrightarrow \mathrm{Aut}(V_1)$. Since $C^i(V_2)$ is irreducible, the space of co-invariants $C^i(V_2)_{N_\varphi}$ is a nonzero irreducible P_φ/N_φ-module. For a subspace $W \subset V_1$ of codimension i, the group N_φ acts transitively on the set of subspaces $W^\sim \subset V_2$ of codimension i such that $W^\sim \cap V_1 = W$. For any such $W^\sim$ we have the isomorphism $V_1/W \xrightarrow{\sim} V_2/W^\sim$. Clearly the composition $\lambda'_{V_1/W} \xrightarrow{\sim} \lambda'_{V_2/W^\sim} \hookrightarrow$

$C^i(V_2) \twoheadrightarrow C^i(V_2)_{N_\varphi}$ depends only on W. The images of the $\lambda'_{V_1/W}$'s generate $C^i(V_2)_{N_\varphi}$ (since the $\lambda'_{V_2/W^\sim}$'s generate $C^i(V_2)$), and the action of P_φ on $C^i(V_2)_{N_\varphi}$ factors through the $\mathrm{Aut}(V_1)$-action (since this is the case on the lines $\lambda'_{V_1/W}$). Therefore, we may identify $C^i(V_2)_{N_\varphi}$ with $C^i(V_1)$, which gives $\varphi^*: C^i(V_2) \twoheadrightarrow C^i(V_1)$.

Let φ be surjective. Let $P_\varphi \subset \mathrm{Aut}(V_1)$ be the parabolic subgroup that preserves $\mathrm{Ker}\,\varphi$; we have the projection $P_\varphi \twoheadrightarrow \mathrm{Aut}(V_2)$. Consider the subspace $C^i(V_1)_\varphi \subset C^i(V_1)$ generated by the lines $\lambda'_{V_1/W}$ for $W \supset \mathrm{Ker}\,\varphi$. This is a P_φ-submodule of $C^i(V_1)$, and the P_φ-action on it factors through the $\mathrm{Aut}(V_2)$-action (since this is the case on the lines). The map $W \mapsto \varphi(W)$ identifies the above W's with the codimension i subspaces of V_2; one has $V_1/W = V_2/\varphi(W)$. Therefore, as above, we may canonically identify $C^i(V_1)_\varphi$, equipped with this system of lines, with $C^i(V_2)$. We get $\varphi^*: C^i(V_2) \hookrightarrow C^i(V_1)$. □

5.2.3. LEMMA. (i) *There exists a unique functorial pairing*

$$\times: C^i(V_1) \otimes C^j(V_2) \to C^{i+j}(V_1 \oplus V_2)$$

such that in case $\dim V_1 = i$, $\dim V_2 = j$ *it coincides with the dual of the isomorphism from* 5.1.5

$$\lambda'_{V_1} \otimes \lambda'_{V_2} \xrightarrow{\sim} \lambda'_{V_1 \oplus V_2}.$$

(ii) *For* $V \in \text{mod-}E$ *consider the multiplication*

$$\cdot: C^i(V) \otimes C^j(V) \to C^{i+j}(V),\ a \cdot b := \delta^*(a \times b),$$

where $\delta: V \to V \oplus V$ *is the diagonal embedding. This multiplication is commutative and associative; hence,* $C^\cdot(V)$ *is a commutative graded* $\mathbb{Q}$*-algebra.*

(iii) $C^\cdot(V)$ *is generated, as* $\mathbb{Q}$*-algebra, by the degree* 1 *component* $C^1(V)$.

(iv) *The pairing* $\cdot: C^i(V) \otimes C^{n-i}(V) \to C^n(V) = \lambda'_V$, $n = \dim V$, *is nondegenerate.*

PROOF. (i) Consider the subspace of $C^{i+j}(V_1 \oplus V_2)$ generated by the lines

$$\lambda'_{V_1 \oplus V_2/W_1 \oplus W_2} = \lambda'_{V_1/W_1 \oplus V_2/W_2} = \lambda'_{V_1/W_1} \otimes \lambda'_{V_2/W_2},$$

where $W_1 \subset V_1$, $W_2 \subset V_2$ are codimension i and j subspaces respectively. By the same highest-weight-theory reasons as above this is an irreducible $\mathrm{Aut}(V_1) \times \mathrm{Aut}(V_2)$-module which may be identified canonically with $C^i(V_1) \otimes C^j(V_2)$; so we get $\times: C^i(V_1) \otimes C^j(V_2) \hookrightarrow C^{i+j}(V_1 \oplus V_2)$.

(ii) Check on the lines.

(iii), (iv) both follow from the irreducibility of the C^i's since the pairing $\cdot: C^i(V) \otimes C^j(V) \to C^{i+j}(V)$ is nonzero for $i + j \le n$. □

5.2.4. REMARKS. (i) Denote by $Q(V)$ the space of those quadratic forms $\nu: V \to \mathbb{Q}$ for which $\nu(ev) = \lambda(e)\nu(v)$ for $e \in E$, $\nu \in V$ (here $\lambda: E^* = \mathrm{GL}(1, E) \to \mathbb{Q}^*$ was defined in 5.1.5). Note that for $\dim V = 1$ we may identify canonically $Q(V)$ with λ'_V: we have a unique functorial isomorphism that sends the form $\nu_1 \in Q(E)$, $\nu_1(1) = 1$, to $1 \in \mathbb{Q} = \lambda'_E$. Therefore, by 5.2.2, we have a canonical functorial identification $C^1(V) = Q(V)$ (the only thing to check is the axiom 5.2.2(b) for the functor $Q(V)$, which is obvious).

(ii) Let F be any field of char 0. In 5.2.1 we may consider the functors C^i_F with values in F-vector spaces that satisfy the same axioms (a), (b) (in the definition of σ one needs to take $\lambda'_V \otimes F$). All the above results remain true in this context. In particular, by unicity, any such functor is canonically isomorphic to $C^i \otimes F$ (since $C^i \otimes F$ obviously satisfies the axioms).

Here is an explicit description of $C^{\cdot}(V)$. Let F be a field of characteristic 0; if E is quaternion algebra, we assume that F splits E. Choose a right $(F \otimes E)$-module T of F-dimension 2; if E is an imaginary quadratic field, we assume that T is a free $(F \otimes E)$-module. Such T is uniquely determined up to an isomorphism. Put $G(T) = \mathrm{Aut}_{F\otimes E}(T)$; so $G(T) = \mathrm{Aut}_F(T) \simeq \mathrm{GL}_2(F)$, for $F = \mathbb{Q}$, $G(T) = (F \otimes E)^*$ if E is imaginary quadratic, and $G(T) = F^*$ if E is a quaternion algebra.

5.2.5. LEMMA. *For any $V \in$ mod-E one has a canonical functorial isomorphism*

$$C^i(V)_F = \left[\Lambda^{2i}_F(T \underset{E}{\otimes} V') \underset{F}{\otimes} \det{}_F^{\otimes - i}(T)\right]^{G(T)} = \mathrm{Hom}_{G(T)}\left(\det{}_F^{\otimes i}(T), \Lambda^{2i}_F(T \underset{E}{\otimes} V')\right)$$

which sends the product on $C^{\cdot}(V)_F$ to the Λ-product.

PROOF. Denote the functor on the right by $C^i(V)^{\sim}_F$. By 5.2.2 we need to establish for any V of dimension i a canonical isomorphism

$$\sigma: \lambda'_V \otimes F \xrightarrow{\sim} C^i(V)^{\sim}_F = \det\left(T \underset{E}{\otimes} V'\right) \otimes \det{}^{\otimes - i}(T),$$

and then to check axiom (b) (since (a) is trivial). Note that on both lines $\mathrm{Aut}(V)$ acts by the same character and for $V = E^n$ both lines are canonically trivialized. Our σ is the unique functorial isomorphism compatible with these trivializations.

To check axiom (b) it suffices to verify that $C^i(V)^{\sim}_F$ is an irreducible $\mathrm{Aut}(V)$-module. This is clear since the action of $G(T) \times \mathrm{Aut}(V)$ on $\Lambda^{\cdot}(T \otimes_E V')$ has simple spectrum (see [M], Chapter 1§4 (4.3)$'$ for the precise character formula). □

5.3. Linear cycles. For an abelian scheme A over B we define $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A)_{\mathbb{Q}} \subset \mathrm{CH}^{\cdot}(A)_{\mathbb{Q}}$ as $\mathbb{Q}$-subspace generated by classes of abelian subschemes in A.

5.3.1. LEMMA. (i) *For any* $A \in \mathscr{A}b^{(X)}$, $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A)_{\mathbb{Q}}$ *coincides with the subalgebra of* $\mathrm{CH}^{\cdot}(A)_{\mathbb{Q}}$ *generated by* $\mathrm{CH}^{1}_{\mathrm{lin}}(A)_{\mathbb{Q}}$.

(ii) *For any morphism* $\varphi: A_1 \to A_2$ *in* $\mathscr{A}b^{(X)}$ *the operators* φ^*, φ_* *between the groups* $\mathrm{CH}^{\cdot}(A_i)_{\mathbb{Q}}$ *preserve the* $\mathrm{CH}^{\cdot}_{\mathrm{lin}}$*-subspaces.*

PROOF. Consider (ii) first. The statement on φ_* is clear (since the image under φ of an abelian subscheme is abelian). If φ is an isogeny then φ_* maps $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_1)_{\mathbb{Q}}$ onto $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_2)_{\mathbb{Q}}$ (since any abelian subscheme in A_2 is the φ-image of an abelian subscheme in A_1); hence by 5.1.2, φ_* and φ^* induce isomorphisms between the spaces $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_i)_{\mathbb{Q}}$. We see that $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A)_{\mathbb{Q}}$ depends only on A modulo isogeny; therefore for any $A_{\mathbb{Q}} \in \mathscr{A}b^{(X)}_{\mathbb{Q}}$ we have a well-defined subspace $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_{\mathbb{Q}}) \subset \mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_{\mathbb{Q}}) := H^{2\cdot}_{\mathscr{M}}(A_{\mathbb{Q}}, \mathbb{Q}(\cdot))$ such that $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A)_{\mathbb{Q}} = \mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_{\mathbb{Q}})$. Now let us check that φ^* preserves $\mathrm{CH}^{\cdot}_{\mathrm{lin}}$ for arbitrary φ. Since any abelian subvariety $A' \subset A_2$ is the preimage of zero by the projection $A_2 \to A_2/A'$, we see (replacing A_2 by A_2/A') that it suffices to show that $\varphi^*(c(0_{A_2})) \in \mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_1)$. Since in $\mathscr{A}b^{(X)}_{\mathbb{Q}}$ any morphism is equivalent to the composition of a coordinate projection $X^{n_1}_{\mathbb{Q}} \overset{\pi}{\twoheadrightarrow} X^{n_2}_{\mathbb{Q}}$ and a coordinate embedding $X^{n_2}_{\mathbb{Q}} \overset{i}{\hookrightarrow} X^{n_3}_{\mathbb{Q}}$, it suffices to check our statement for $\varphi = \pi$ and $\varphi = i$. The first case is obvious; in the second case we have $i^*(c(0_{X^{n_3}})) = 0$ (if $n_2 < n_3$) since the normal bundle to X^{n_2} in X^{n_3} has zero Chern classes (see the proof of 5.1.1).

It remains to prove (i). If $A', A'' \subset A$ are abelian subschemes then $c(A') \cdot c(A'') = \delta^*(c(A' \times_B A''))$ where $\delta: A \hookrightarrow A \times_B A$ is diagonal embedding. Therefore, $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A)_{\mathbb{Q}}$ is a subalgebra of $\mathrm{CH}^{\cdot}(A)_{\mathbb{Q}}$. To prove that it is generated by the degree 1 component, it suffices, as above, to check that $c(0_A)$ is a product of elements from $\mathrm{CH}^{1}_{\mathrm{lin}}(A)_{\mathbb{Q}}$. Since A is isogeneous to X^n, we may assume that $A = X^n$ where the statement is obvious. □

Therefore for any $A_{\mathbb{Q}} \in Ab^{(X)}_{\mathbb{Q}}$ we have a natural subalgebra $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_{\mathbb{Q}}) \subset \mathrm{CH}^{\cdot}(A_{\mathbb{Q}}) = H^{2\cdot}_{\mathscr{M}}(A_{\mathbb{Q}}, \mathbb{Q}(\cdot))$; the subspaces $\mathrm{CH}^{\cdot}_{\mathrm{lin}}(A_{\mathbb{Q}}) \otimes \lambda_{A_{\mathbb{Q}}} \subset H^{2\cdot}_{\mathrm{lin}}(A_{\mathbb{Q}}, \mathbb{Q}(\cdot)) \otimes \lambda_{A_{\mathbb{Q}}}$ are functorial with respect to Gysin maps (see 5.1.4).

5.3.2. PROPOSITION. *For* $V \in$ mod-E *there is a canonical functorial isomorphism of graded* $\mathbb{Q}$*-algebras*

$$\mathrm{CH}^{\cdot}_{\mathrm{lin}}(V \otimes_E X_{\mathbb{Q}}) = C^{\cdot}(V) \otimes \lambda_{X_{\mathbb{Q}}}^{\otimes -\cdot}.$$

PROOF. Put $C^i(V)^{\sim} := \mathrm{CH}^{i}_{\mathrm{lin}}(V \otimes X_{\mathbb{Q}}) \otimes \lambda_{X_{\mathbb{Q}}}^{\otimes i}$. According to 5.2.2 the proposition will follow if we establish a canonical isomorphism $\sigma: \lambda^i_V \overset{\sim}{\to} C^i(V)^{\sim}$ for i-dimensional V's and check the axioms 5.2.1(a), (b) for $(C^{i\sim}, \sigma)$.

To define σ note that $\mathrm{CH}^{\dim A}_{\mathrm{lin}}(A)$ is obviously a 1-dimensional $\mathbb{Q}$-vector space. If $\dim A = 0$, it is canonically trivialized. For arbitrary $A \in \mathscr{A}b^{(X)}_{\mathbb{Q}}$

the Gysin map for the zero projection $A \to 0$ is a canonical isomorphism $\mathrm{CH}^{\dim A}_{\mathrm{lin}}(A) \otimes \lambda_A \xrightarrow{\sim} \mathbb{Q}$. We may rewrite it using the formula from 5.1.5 as $C^i(V)^\sim \otimes \lambda_V \xrightarrow{\sim} \mathbb{Q}$ which provides our σ.

Since any abelian subvariety $A' \subset A$ is the preimage of 0 by the projection $A \to A/A'$, we see that 5.2.1(b) trivially holds. To check 5.2.1(a) it suffices by 5.3.1(i), to treat the case $i = 1$. For $A \in \mathscr{A}b^{(X)}$ one has the standard filtration $P_0 \subset P_1 \subset P_2 = \mathrm{Pic}(A)_\mathbb{Q} = \mathrm{CH}^1(A)_\mathbb{Q}$, $P_0 = \mathrm{Pic}(B)_\mathbb{Q}$, $P_1/P_0 = A(B)_\mathbb{Q}$. This filtration splits canonically. Namely, let $P^\pm$ be the $\pm$-eigenspaces of the $[-1]$-involution; put $P_{(0)} = P_{(0)}$, $P_{(1)} = P^-$, $P_{(2)} := \mathrm{Ker}(0_A^*\colon \mathrm{Pic}(A)_\mathbb{Q} \to \mathrm{Pic}(B)_\mathbb{Q}) \cap P^+$. Then the splitting is $P_{(0)} \oplus P_{(1)} \oplus P_{(2)} \xrightarrow{\sim} \mathrm{Pic}(A)_\mathbb{Q}$ (see the proof of 5.1.1 where the case $\dim A = 1$ was considered). Note that $\mathrm{CH}^1_{\mathrm{lin}}(A)_\mathbb{Q} \subset P_{(2)}$. If $b \in B$ is a geometric point and ℓ is a prime $\neq \mathrm{char}\, b$ then we have the Chern class map

$$c_{b,\mathbb{Q}_\ell}\colon \mathrm{Pic}(A) \otimes \mathbb{Q}_\ell \to \mathrm{Pic}(A_b) \otimes \mathbb{Q}_\ell \xrightarrow{c_1} H^2(A_b, \mathbb{Q}_\ell(1)) = [\Lambda^2 H^1(A_b, \mathbb{Q}_\ell)](1).$$

One knows that $c_{b,\mathbb{Q}_\ell}$ kills P_1 and that the induced map $(P_2/P_1) \otimes \mathbb{Q}_\ell \to [\Lambda^2 H^1(A, \mathbb{Q}_\ell)](1)$ is injective. Therefore, $c_{b,\mathbb{Q}_\ell}$ is injective on $\mathrm{CH}^1_{\mathrm{lin}}(A) \otimes \mathbb{Q}_\ell$. Hence $\dim \mathrm{CH}^1_{\mathrm{lin}}(A)_\mathbb{Q} < \infty$ and the action of $\mathrm{Aut}(A_\mathbb{Q})$ on $\mathrm{CH}^1_{\mathrm{lin}}(A)_\mathbb{Q}$ is algebraic (since so is the action on $H^1(A_b, \mathbb{Q}_\ell)$), and we get 5.1.1(a) for $C^{1\sim}$. $\square$

Let now, as above, b be a geometric point of B, $\ell \neq \mathrm{char}\, b$. For $A_\mathbb{Q} \in \mathscr{A}b^{(X)}_\mathbb{Q}$ let $c_{b,\mathbb{Q}_\ell}$ be the composition $\mathrm{CH}^i_{\mathrm{lin}}(A_\mathbb{Q}) \otimes \mathbb{Q}_\ell \to \mathrm{CH}^i(A_{\mathbb{Q}b}) \to H^{2i}(A_{\mathbb{Q}b}, \mathbb{Q}_\ell(i))$ of the restriction and the $\mathbb{Q}_\ell$-class maps. Put $T_{\mathbb{Q}_\ell} = H^1(X_{\mathbb{Q}b}, \mathbb{Q}_\ell)$—this is a right $(\mathbb{Q}_\ell \otimes E)$-module that satisfies the conditions of 5.2.5 (for $F = \mathbb{Q}_\ell$).

5.3.3. Corollary. (i) *For $V \in$ mod-E the diagram*

$$\begin{array}{ccc}
\mathrm{CH}^i_{\mathrm{lin}}(V \underset{E}{\otimes} X_\mathbb{Q}) \otimes \mathbb{Q}_\ell & \xrightarrow{c_{b,\mathbb{Q}_\ell}} & H^{2i}(V \underset{E}{\otimes} X_{\mathbb{Q}b}, \mathbb{Q}_\ell(i)) \\
(5.3.2)\Vert & & \Vert \\
 & & \Lambda^{2i} H^1(V \underset{E}{\otimes} X_{\mathbb{Q}b}, \mathbb{Q}_\ell)(i) \\
C^i(V) \otimes \lambda_{X_\mathbb{Q}}^{\otimes -i} \otimes \mathbb{Q}_\ell & & \Vert \\
 & & \Lambda^{2i}(T_{\mathbb{Q}_\ell} \underset{E}{\otimes} V')(i) \\
(5.1.6)\Vert & & \uparrow \\
C^i(V)_{\mathbb{Q}_\ell} \otimes \det^{\otimes i}(T_{\mathbb{Q}_\ell})(i) & \underset{(5.2.5)}{=} & [\Lambda^{2i}(T_{\mathbb{Q}_\ell} \underset{E}{\otimes} V') \otimes \det^{\otimes -i}(T_{\mathbb{Q}_\ell})]^{G(T_{\mathbb{Q}_\ell})} \otimes \det^{\otimes i}(T_{\mathbb{Q}_\ell})(i)
\end{array}$$

commutes.

(ii) *The map* $c_{b,\mathbb{Q}_\ell}$ *is injective.*

PROOF. (i) The diagram is functorial; therefore, it suffices (say, by 5.2.1(b)) to check its commutativity for $i = \dim V$, where it is obvious.

(ii) follows from (i). □

5.3.4. REMARKS. (i) If b is a $\mathbb{C}$-point, we may use in 5.3.3 the usual Betti $\mathbb{Q}$-cohomology instead of $\mathbb{Q}_\ell$-ones.

(ii) The statement 5.3.3(ii) should hold for any abelian scheme (not necessary a product of elliptic curves). This would follow immediately from general conjectures on mixed motives **[B1]**.

5.3.5. Let $\varphi\colon A_1 \to A_2$ be a morphism in $\mathscr{A}b^{(X)}$. For $\ell_1 \in \mathrm{CH}^i(A_1)$, $\ell_2 \in \mathrm{CH}^j(A_2)$ one has the projection formula $\varphi_*(\ell_1\varphi^*\ell_2) = \varphi_*(\ell_1)\ell_2$. This implies (take $i+j = \dim A_1$) that the map φ_* on $\mathrm{CH}^{\cdot}_{\mathrm{lin}}$ is the transpose of φ^* with respect to the intersection pairing $\mathrm{CH}^i_{\mathrm{lin}}(A)_{\mathbb{Q}} \times \mathrm{CH}^{n-i}_{\mathrm{lin}}(A)_{\mathbb{Q}} \to \mathrm{CH}^n_{\mathrm{lin}}(A)_{\mathbb{Q}} = \mathbb{Q}$, $n = \dim A$. Note that this pairing is nondegenerate by 5.3.2, 5.2.3(iv). Another way to say this is to define for a morphism $\psi\colon V_1 \to V_2$ in mod-E the map $\psi_*\colon C^{\cdot}(V_1) \otimes \lambda_{V_1} \to C^{\cdot -d}(V_2) \otimes \lambda_{V_2}$, $d = \dim V_2 - \dim V_1$, as the transpose of ψ^* with respect to the pairings 5.2.3(iv). Then the isomorphism 5.3.2 sends ψ_* to the corresponding Gysin map on the $\mathrm{CH}^{\cdot}_{\mathrm{lin}}$-groups (we identify the twists using 5.1.5).

5.4. The category $\mathscr{M}^{(X)}$. We are ready to define our category of motives $\mathscr{M}^{(X)}$. This will be done in four steps; we need some auxiliary categories.

(i) The category $\mathscr{C}$. Its objects coincide with that of $\mathscr{A}b^{(X)}_{\mathbb{Q}}$; for $A_{\mathbb{Q}} \in \mathscr{A}b^{(X)}_{\mathbb{Q}}$ we denote by $[A_{\mathbb{Q}}]$ the corresponding object of $\mathscr{C}$. The Hom's in $\mathscr{C}$ are graded $\mathbb{Q}$-vector spaces of linear correspondences: one has

$$\mathrm{Hom}^i_{\mathscr{C}}([A_{2\mathbb{Q}}], [A_{1\mathbb{Q}}]) := \mathrm{CH}^{i+n_2}_{\mathrm{lin}}(A_{1\mathbb{Q}} \times A_{2\mathbb{Q}}) \otimes \lambda_{A_{2\mathbb{Q}}}, \; n_2 = \dim A_{2\mathbb{Q}}.$$

The composition

$$\mathrm{Hom}^i_{\mathscr{C}}([A_{2\mathbb{Q}}], [A_{1\mathbb{Q}}]) \times \mathrm{Hom}^j_{\mathscr{C}}([A_{3\mathbb{Q}}], [A_{2\mathbb{Q}}]) \xrightarrow{\circ} \mathrm{Hom}^{i+j}_{\mathscr{C}}([A_{3\mathbb{Q}}], [A_{1\mathbb{Q}}])$$

is $f \circ g := p_{13*}(p_{12}^*(f) \cdot p_{23}^*(g))$, where $p_{ij}\colon A_{1\mathbb{Q}} \times A_{2\mathbb{Q}} \times A_{3\mathbb{Q}} \to A_{i\mathbb{Q}} \times A_{j\mathbb{Q}}$ are the projections. Denote by $\mathscr{C}^0 \subset \mathscr{C}$ the subcategory with the same objects as $\mathscr{C}$ and $\mathrm{Hom}_{\mathscr{C}^0} = \mathrm{Hom}^0_{\mathscr{C}}$. We have a canonical contravariant functor $[\;]\colon \mathscr{A}b^{(X)}_{\mathbb{Q}} \to \mathscr{C}^0$ that sends $A_{\mathbb{Q}}$ to $[A_{\mathbb{Q}}]$ and $\varphi\colon A_{1\mathbb{Q}} \to A_{2\mathbb{Q}}$ to $[\varphi] :=$ (class of the graph of φ) $\in \mathrm{Hom}^0_{\mathscr{C}}([A_{2\mathbb{Q}}], [A_{1\mathbb{Q}}])$.

(ii) The category $\mathscr{C}^{\sim}$. Its objects are pairs $(A_{\mathbb{Q}}, i)$, $A_{\mathbb{Q}} \in \mathscr{A}b^{(X)}_{\mathbb{Q}}$, $i \in \mathbb{Z}$; we denote such objects as $[A_{\mathbb{Q}}](i)$. One has $\mathrm{Hom}_{\mathscr{C}^{\sim}}([A_{2\mathbb{Q}}](i_2), [A_{1\mathbb{Q}}](i_1)) := \mathrm{Hom}^{i_1-i_2}_{\mathscr{C}}([A_{2\mathbb{Q}}], [A_{1\mathbb{Q}}])$ with obvious composition law. We have a canonical fully faithful embedding $\mathscr{C}^0 \hookrightarrow \mathscr{C}^{\sim}$, $[A_{\mathbb{Q}}] \mapsto [A_{\mathbb{Q}}](0)$.

(iii) The category $\mathscr{C}^{\sim}$ is almost a $\mathbb{Q}$-linear additive category; it only lacks direct sums. We define the category $\mathscr{C}^{\approx}$ as a universal additive category that contains $\mathscr{C}^{\sim}$. Its objects are formal finite direct sums of objects in $\mathscr{C}^{\sim}$; the Hom's are defined accordingly.

(iv) We define $\mathscr{M}^{(X)}$ as the pseudo-abelian envelope of $\mathscr{C}^{\approx}$. Therefore, its objects are pairs (F, p), where $F \in \mathscr{C}^{\approx}$ and $p \in \operatorname{End} F$ is a projector; the Hom's are defined in the usual way. We have a canonical contravariant functor $\mathscr{A}b_{\mathbb{Q}}^{(X)\circ} \to \mathscr{M}^{(X)}$, $A_{\mathbb{Q}} \mapsto [A_{\mathbb{Q}}] = [A_{\mathbb{Q}}](0)$.

5.4.1. Lemma. *For any pseudo-abelian $\mathbb{Q}$-linear category $\mathscr{N}$ the category of $\mathbb{Q}$-linear functors $\mathscr{M}^{(X)} \to \mathscr{N}$ is equivalent to those of $\mathbb{Q}$-linear functors $\mathscr{C}^{\sim} \to \mathscr{N}$.* □

5.4.2. Let us define the standard *realization functors*. For a geometric point $b \in B$ and $\ell \neq \operatorname{char} b$ we have the *ℓ-adic realization functor* $r_{\mathbb{Q}_\ell b}\colon \mathscr{M}^{(X)} \to \operatorname{Vect}_{\mathbb{Q}_\ell}$, $M \mapsto r_{\mathbb{Q}_\ell b}(M) = M_{\mathbb{Q}_\ell b}$. By 5.4.1 we need to define it on $C^{\sim}$. We put $[A_{\mathbb{Q}}](i)_{\mathbb{Q}_\ell b} := \bigoplus_j H^j(A_{\mathbb{Q}b}, \mathbb{Q}_\ell)(i)$. The functoriality is the action of correspondences on ℓ-adic cohomology:

$$\begin{aligned}
\operatorname{Hom}_{\mathscr{C}^{\sim}}([A_{2\mathbb{Q}}](i_2), [A_{1\mathbb{Q}}](i_1)) &= \operatorname{CH}_{\mathrm{lin}}^{i_1-i_2+n_2}(A_{1\mathbb{Q}} \times A_{2\mathbb{Q}}) \otimes \lambda_{A_{2\mathbb{Q}}} \\
&\overset{c_{b,\mathbb{Q}_\ell}}{\to} H^{2(i_1-i_2+n_2)}(A_{1\mathbb{Q}b} \times A_{2\mathbb{Q}b}, \mathbb{Q}_\ell(i_1 - i_2 + n_2)) \otimes H^{2n_2}(A_{2\mathbb{Q}b}, \mathbb{Q}_\ell(n_2))' \\
&= \bigoplus_j \operatorname{Hom}(H^j(A_{2\mathbb{Q}b}, \mathbb{Q}_\ell)(i_2), H^{2(i_1-i_2)+j}(A_{1\mathbb{Q}b}, \mathbb{Q}_\ell)(i_1)) \\
&\subset \operatorname{Hom}([A_{2\mathbb{Q}}](i_2)_{\mathbb{Q}_\ell b}, [A_{1\mathbb{Q}}](i_1)_{\mathbb{Q}_\ell b}).
\end{aligned}$$

When b varies, the $r_{\mathbb{Q}_\ell b}(M)$ form a $\mathbb{Q}_\ell$-local system on B (assuming $\ell \notin \operatorname{char} B$).

A $\mathbb{C}$-point b of B defines the *Betti realization functor* $r_{Bb}\colon \mathscr{M}^{(X)} \to \operatorname{Vect}_{\mathbb{Q}}$ which lifts to the *Hodge realization functor* $r_{\mathscr{H}_{\mathbb{Q}}b}\colon \mathscr{M}^{(X)} \to \mathscr{H}_{\mathbb{Q}}$; if S is a scheme over $\mathbb{C}$, the $r_{\mathscr{H}_{\mathbb{Q}}b}(M)$ are fibers of a variation of Hodge structure $r_{\mathscr{H}_{\mathbb{Q}}}(M)$ We may also consider the *de Rham realization functor* r_{DR}, etc. The construction of these functors repeats the one for $r_{\mathbb{Q}_\ell}$.

Consider now the group $G(T_{\mathbb{Q}_\ell}) \subset \operatorname{GL}(T_{\mathbb{Q}_\ell})$ from 5.3.3, 5.2.5. It acts canonically on the functor $r_{\mathbb{Q}_\ell b}$; i.e., $r_{\mathbb{Q}_\ell b}$ lifts to the functor $r_{\mathbb{Q}_\ell b}^G\colon \mathscr{M}^{(X)} \to$ (algebraic $G(T_{\mathbb{Q}_\ell})$-modules). By 5.4.1 it suffices to define the $G(T_{\mathbb{Q}_\ell})$-action on the $\mathbb{Q}_\ell$-vector spaces $[A_{\mathbb{Q}}](i)_{\mathbb{Q}_\ell b} = [X_{\mathbb{Q}} \otimes_E V](i)_{\mathbb{Q}_\ell b} = \Lambda^{\cdot}(T_{\mathbb{Q}_\ell} \otimes_E V) \otimes \mathbb{Q}_\ell(i)$ (here $A_{\mathbb{Q}} = X_{\mathbb{Q}} \otimes_E V$). This is the restriction of the standard $\operatorname{GL}(T_{\mathbb{Q}_\ell})$-action on the exterior algebra multiplied by the character $\det^{-i}$, one checks in a moment that this $G(T_{\mathbb{Q}_\ell})$-action is natural with respect to morphisms in $\mathscr{C}^{\sim}$. If b is a $\mathbb{C}$-point, we may repeat the above construction for r_{Bb} (see 5.3.4) to get the functor $r_{Bb}^G\colon \mathscr{M}^{(X)} \to$ (algebraic $G(T_{\mathbb{Q}})$-modules).

5.4.3. LEMMA. (i) *The category* $\mathscr{M}^{(X)}$ *is a semisimple abelian* $\mathbb{Q}$*-category.*

(ii) *The functors* $r^G_{\mathbb{Q}_\ell b}: \mathscr{M}^{(X)} \otimes \mathbb{Q}_\ell \to$ (*algebraic* $G(T_{\mathbb{Q}_\ell})$*-modules*), $r^G_{Bb}: \mathscr{M}^{(X)} \to$ (*algebraic* $G(T)$*-modules*) *are equivalences of categories. In particular,* $r_{\mathbb{Q}_\ell b}$, r_{Bb} *are faithful functors.*

(iii) *If* B *is a scheme of finite type over* $\mathbb{Z}$ *or* $\mathbb{Q}$ *then the functor* $r_{\mathbb{Q}_\ell}: \mathscr{M}^{(X)} \otimes \mathbb{Q}_\ell \to$ (*lisse* $\mathbb{Q}_\ell$*-sheaves on* B) *is fully faithful.*

PROOF. (i) follows from (ii) since $G(T_{\mathbb{Q}_\ell})$, $G(T_{\mathbb{Q}})$ are reductive groups.

(ii) By 5.3.3, 5.3.4 the functors $r^F_{\mathbb{Q}_\ell b}$, r^G_{Bb} induce isomorphisms for the Hom's between the objects $[A_{\mathbb{Q}}](i)$; hence $r^G_{\mathbb{Q}_\ell b}$, r^G_{Bb} are fully faithful functors. Since $G(T)$ is reductive and any irreducible representation of $G(T)$ occurs in some $r^G_b([A_{\mathbb{Q}}](i))$, we are done.

(iii) follows from (ii) since $G(T_{\mathbb{Q}_\ell})$ coincides with the Zariski closure of the image of the Galois group acting on $T_{\mathbb{Q}_\ell}$; similarly $G(T_{\mathbb{Q}})$ is the corresponding Hodge group. □

5.4.4. Our $\mathscr{M}^{(X)}$ has a natural tensor category structure. By (a variant of) 5.4.1 it suffices to define the tensor product for objects $[A_{\mathbb{Q}}(i)]$ where it is

$$[A_{1\mathbb{Q}}](i_1) \otimes [A_{2\mathbb{Q}}](i_2) := [A_{1\mathbb{Q}} \times A_{2\mathbb{Q}}](i_1 + i_2).$$

The associativity and commutativity constraints are the obvious ones. Clearly $r^G_{\mathbb{Q}_\ell b}$, r^G_{Bb} are equivalences of tensor categories; so $\mathscr{M}^{(X)}$ is a Tannakian $\mathbb{Q}$-category.

One has a canonical invertible object $\mathbb{Q}(1) := [0](1)$. For $M \in \mathscr{M}^{(X)}$, $i \in \mathbb{Z}$, we put $M(i) := M \otimes \mathbb{Q}(1)^{\otimes i}$.

5.4.5. The objects of $\mathscr{M}^{(X)}$ have a natural grading by weights $M = \bigoplus M_w$. To define this grading for any M it suffices, by 5.4.1, to define it on objects $[A_{\mathbb{Q}}](i)$. These objects carry a $\mathbb{Q}^*$-action; let $n \in \mathbb{Q}^*$ act as $n^{-i}[n_A]^*$. By 5.4.3(ii) we have the decomposition $[A_{\mathbb{Q}}](i) = \bigoplus [A_{\mathbb{Q}}](i)_w$ where $[A_{\mathbb{Q}}](i)_w$ is the component on which $n \in \mathbb{Q}^*$ acts as n^w. The equivalences 5.4.3(ii) send this weight grading to the grading by weights for the multiplicative subgroup of homotheties $\subset G(T)$. One has $(M \otimes N)_w = \bigoplus_{a+b=w} M_a \otimes N_b$.

For $A_{\mathbb{Q}} \in \mathscr{A}b^{(X)}_{\mathbb{Q}}$ put $H^j(A) := [A_{\mathbb{Q}}]_j$, so $[A_{\mathbb{Q}}] = \bigoplus H^j(A_{\mathbb{Q}})$. The realization functors $r_{\mathbb{Q}_\ell b}$, r_{Bb} send $H^j(A_{\mathbb{Q}})$ to $H^j(A_{\mathbb{Q}}, \mathbb{Q}_\ell)$, $H^j(A_{\mathbb{Q}}, \mathbb{Q})$ respectively. Note that $[A_{\mathbb{Q}}]$ is a ring object of $\mathscr{M}^{(X)}$ with the multiplication $[A_{\mathbb{Q}}] \otimes [A_{\mathbb{Q}}] \to [A_{\mathbb{Q}}]$ defined as the pull-back for the diagonal map $A_{\mathbb{Q}} \to A_{\mathbb{Q}} \times A_{\mathbb{Q}}$. By 5.4.3(ii) we see that this multiplication defines an isomorphism (of ring objects) $\Lambda^{\cdot} H^1(A_{\mathbb{Q}}) \xrightarrow{\sim} [A_{\mathbb{Q}}]$, i.e., $H^j(A_{\mathbb{Q}}) = \Lambda^j H^1(A_{\mathbb{Q}})$.

5.5. Absolute cohomology groups. For $j \in \mathbb{Z}$ let us define a $\mathbb{Q}$-linear functor $\tilde{H}^j_{\mathscr{M}}: \mathscr{M}^{(X)} \to \mathrm{Vect}_{\mathbb{Q}}$ as follows. By 5.4.1 it suffices to define $\tilde{H}^j_{\mathscr{M}}$ on $\mathscr{C}^{\sim}$; put $\tilde{H}^j_{\mathscr{M}}([A_{\mathbb{Q}}](i)) := H^{j+2i}_{\mathscr{M}}(A_{\mathbb{Q}}, \mathbb{Q}(i))$. The morphisms in $\mathscr{C}^{\sim}$ act on

$\widetilde{H}^j_{\mathscr{M}}$ as correspondences; for a morphism $\varphi \in \mathrm{Hom}_{\mathscr{C}^\sim}([A_{2\mathbb{Q}}](i_2), [A_{1\mathbb{Q}}](i_1)) = \mathrm{CH}^{n_2+i_1-i_2}_{\mathrm{lin}}(A_{1\mathbb{Q}} \times A_{2\mathbb{Q}}) \subset H^{2(n_2+i_1-i_2)}_{\mathscr{M}}(A_{1\mathbb{Q}} \times A_{2\mathbb{Q}}, \mathbb{Q}(n_2 + i_1 - i_2))$, $\alpha \in \widetilde{H}^j_{\mathscr{M}}([A_{2\mathbb{Q}}](i))$, one has $\varphi(\alpha) := \mathrm{pr}_{1*}(\varphi\, \mathrm{pr}_2^*(\alpha))$.

The $\mathbb{Q}$-linear *absolute cohomology* functor $H^j_{\mathscr{M}}: \mathscr{M}^{(X)} \to \mathrm{Vect}_{\mathbb{Q}}$ is

$$H^j_{\mathscr{M}}(M) := \bigoplus_w \widetilde{H}^{j+w}_{\mathscr{M}}(M_w).$$

For $A \in \mathscr{A}b^{(X)}$ the weight decomposition of $[A_{\mathbb{Q}}]$ provides a canonical *motivic decomposition*

$$H^j_{\mathscr{M}}(A, \mathbb{Q}(i)) = \bigoplus_a H^{j-a}_{\mathscr{M}}(H^a(A_{\mathbb{Q}})(i)) = \bigoplus_a H^{j-a}_{\mathscr{M}}(\Lambda^a H^1(A_{\mathbb{Q}})(i)).$$

The a-component of the above decomposition coincides with the n^a-eigenspace of the operators $[n]^*_A$, $n \in \mathbb{Z}$, acting on $H^j_{\mathscr{M}}(A, \mathbb{Q}(i))$.

6. A motivic version

In this section we shall give a motivic, in the K-theory sense, construction of the elliptic polylogarithm extension. We start with another construction of the class of this extension in the sheaf framework, and then we shall show how to lift it to the K-version of absolute motivic cohomology.

6.1. The polylogarithm classes. Let $p_X: X \to B$ be our family of elliptic curves, $j: U = X \backslash 0(B) \hookrightarrow X$ be the embedding. For $n \geq 0$ put $X^{n+1} = X \times_B \cdots \times_B X$ ($n+1$ times), $j^{n+1}: U^{n+1} \hookrightarrow X^{n+1}$; for a subset $I \subset \{1, \dots, n+1\}$ let $i^I: X^I \hookrightarrow X^{n+1}$ be the subscheme defined by the equations $X_j = 0$ for $j \notin I$. Let $\sigma^{n+1}: X^{n+1} \to X$ be the projection $\sigma^{n+1}(x_1, \dots, x_{n+1}) = x_1 + \cdots + x_{n+1}$. If $I \neq \varnothing$, then $\sigma^{n+1} i^I: X^I \to X$ is a smooth proper fibration. We have the usual action of the symmetric group $\sum_{n+1}$ on X^{n+1} that interchanges the I^I's and fixes σ^{n+1}. Put $\lambda := \det(\mathbb{Q}^{n+1})$—this is a standard $\mathbb{Q}$-line on which $\sum_{n+1}$ acts via the character sgn. If M is an object of an abelian $\mathbb{Q}$-category equipped with a $\sum_{n+1}$-action, we put $M_{\mathrm{sgn}} := (M \otimes \lambda)^{\Sigma_{n+1}}$—this is the component of M on which $\sum_{n+1}$ acts via sgn.

As in §1 we put

$$\mathscr{H} := R^1 p_{X*} F(1)_X, \qquad S^i := \mathrm{Sym}^i \mathscr{H}.$$

6.1.1. Lemma. (i) *One has* $[R^i \sigma^{n+1}_* F(n)_{X^{n+1}}]_{\mathrm{sgn}} = 0$ *for* $i \neq n$; *there is a canonical isomorphism* $[R^n \sigma^{n+1}_* F(n)_{X^{n+1}}]_{\mathrm{sgn}} = p_X^* S^n$.

(ii) *One has a canonical isomorphism*

$$R^n j^{n+1}_* F(n)_{U^{n+1}} = \bigoplus_{\substack{I \subset \{1, \dots, n+1\} \\ |I| = n+1-a}} i^I_* F(n-a)_{X^I}.$$

Proof. (i) is a standard calculation. The isomorphism $p_X^* S^n \xrightarrow{\sim} [R^n \sigma^{n+1}_* F(n)_{X^{n+1}}]_{\mathrm{sgn}}$ is given by the formula $\gamma^n \mapsto r_1^*\gamma \cup r_2^*\gamma \cup \cdots \cup r_n^*\gamma$, $\gamma \in \mathscr{H}$, where $r_i \colon X^{n+1} \to X$ is the ith projection. The complement of U in X is a divisor with normal crossings, so (ii) is a standard isomorphism given by successive residues. □

6.1.2. Put $G_n^{b,a} := [R^b\sigma_*^{n+1}R^a j_*^{n+1}F(n)_{U^{n+1}}]_{\mathrm{sgn}}$. We see that all these sheaves on X are zero but the following ones: $G_n^{0,n+1} = 0_*F_B(-1)$ (the skyscraper sheaf supported at the zero point of X) and $G_n^{b,n-b} = p_X^*S^b$ for $0 \le b \le n$. This implies that the complex

$$[R(\sigma^{n+1}j^{n+1})_*F(n)_{U^{n+1}}]_{\mathrm{sgn}}[n] = [R\sigma_*^{n+1}Rj_*^{n+1}F(n)_{U^{n+1}}]_{\mathrm{sgn}}[n]$$

is supported in degrees 0 and 1. The first cohomology is $0_*F(-1)$. Denote the zero cohomology by $G^{(n)}$. This is a lisse sheaf on X; the canonical filtration τ on $Rj_{j*}^{n+1}F(n)_{U^{n+1}}[n]$ defines on $G^{(n)}$ a filtration $G^{(n)} = G^{(n)0} \supset G^{(n)1} \supset \cdots \supset G^{(n)n+1} = 0$ with successive quotients $G^{(n)i}/G^{(n)i+1}$ equal to $p_X^*S^i$, $i = 0, \dots, n$. The residue map at $X^n = X^{\{1,\dots,n\}} \subset X^{n+1}$ defines a canonical projection $G^{(n)} \twoheadrightarrow G^{(n-1)}$ that identifies $G^{(n-1)}$ with $G^{(n)}/G^{(n)n} = G^{(n)}/p_X^*S^n$. Put $G := \varprojlim G^{(n)}$—this is a lisse sheaf equipped with filtration $G^i := \varprojlim_n G^{(n)i}$ one has $G^{(n)} = G/G^{n+1}$, hence $\operatorname{gr}^n G = p_X^*S^n$.

6.1.3. For a nonzero integer a prime to characteristics of B consider the multiplication by a isogenies $[a]_X \colon X \to X$ $[a]_{X^{n+1}} \colon X^{n+1} \to X^{n+1}$. One has $[a]_X^{-1}(U) \subset U$, $[a]_{X^{n+1}}^{-1}(U^{n+1}) \subset U^{n+1}$, and $[a]_X\sigma^{n+1} = \sigma^{n+1}[a]_{X^{n+1}}$. Hence we have the corresponding relative trace morphism $\mathrm{tr}_a^{(n)} \colon R(\sigma^{n+1}j^{n+1})_*F(n)_{U^{n+1}} \to [a]_X^*R(\sigma^{n+1}j^{n+1})_*F(n)_{U^{n+1}}$. Precisely, consider the commutative diagram

$$\begin{array}{ccccc}
[a]_{X^{n+1}}^{-1}(U^{n+1}) & \xrightarrow{\varphi_U} & \widetilde{U}^{n+1} & \longrightarrow & U^{n+1} \\
\downarrow & & \downarrow & & \downarrow \\
X^{n+1} & \xrightarrow{\varphi} & \widetilde{X}^{n+1} & \xrightarrow{\psi} & X^{n+1} \\
\downarrow{\scriptstyle \sigma^{n+1}} & & \downarrow{\scriptstyle \widetilde{\sigma}^{n+1}} & & \downarrow{\scriptstyle \sigma^{n+1}} \\
X & = & X & \xrightarrow{[a]_X} & X
\end{array}$$

where the right squares are Cartesian and $\psi\varphi = [a]_{X^{n+1}}$. Then $\mathrm{tr}_a^{(n)}$ is the composition

$$R(\sigma^{n+1}|_{U^{n+1}})_*F(n)_{U^{n+1}} \to R(\sigma^{n+1}|_{[a]_{X^{n+1}}^{-1}(U^{n+1})})_*F(n)_{[a]_{X^{n+1}}^{-1}(U^{n+1})}$$
$$\xrightarrow{\mathrm{tr}\,\varphi_U} R(\sigma^{n+1}|_{\widetilde{U}^{n+1}})_*F(n)_{\widetilde{U}^{n+1}} \xrightarrow{\sim} [a]_X^*R(\sigma^{n+1}j^{n+1})_*F(n)_{U^{n+1}}$$

where the last isomorphism comes from base change. Since it commutes with the $\sum_{n+1}$-action, we get the corresponding morphism $\mathrm{tr}_{aG}^{(n)} \colon G \to [a]_X^*G^{(n)}$; and since the trace commutes with residues, we see that the $\mathrm{tr}_{aG}^{(n)}$ are compatible, i.e., we have a morphism $\mathrm{tr}_{aG} = \varprojlim \mathrm{tr}_{aG}^{(n)} \colon G \to [a]_X^*G$ compatible with the filtration $G^{\cdot}$. All these trace maps for different a's commute.

6.1.4. LEMMA. *One has* $\operatorname{gr}^n \operatorname{tr}_{aG} = a^n \operatorname{id}_{p_X^* S^n}$.

PROOF. Clear. □

6.1.5. LEMMA. *The filtered prosheaf* $G = G^0 \supset G^1 \supset \cdots$ *equipped with the isomorphisms* $\operatorname{gr}^n G = p_X^* S^n$ *is canonically isomorphic to the filtered prosheaf* $G^{\cdot}$ *from* 1.2.3.

PROOF. By 1.2.1(a) we need to check the differential d_1 in the spectral sequence and to show that at point 0 there is a splitting $F_B = G_0/G_0^1 \to G_0$. The check is easy; the splitting follows, for example, from 6.1.4 since the filtration of G_0 splits according to eigenspaces of tr_{aG}'s. □

6.1.6. We may rewrite tr_{aG} as a morphism $[a]_{X*}G \to G$; hence, we have the corresponding trace endomorphisms tr_a acting on $Rp_{U_*}G_U$ (defined as a composition $Rp_{U_*}G_U \to Rp_{[a]_X^{-1}(U)*}G_{[a]_X^{-1}(U)} = Rp_{U_*}([a]_{X*}G)_U \overset{\operatorname{tr}_{aG}}{\to} Rp_{U*}G_U$). Similarly, we have $\operatorname{tr}_a^{(n)} \in \operatorname{End} Rp_{U*}G_U^{(n)}$ and $\operatorname{tr}_a = \varprojlim \operatorname{tr}_a^{(n)}$.

6.1.7. Put $U_0^{n+1} := (\sigma^{n+1})^{-1}(U) \cap U^{n+1} = U^{n+1} \backslash (\sigma^{n+1})^{-1}(0)$. This is an open $\sum_{n+1}$-invariant subspace of U^{n+1} such that $[a]_{X^{n+1}}^{-1}(U_0^{n+1}) \subset U_0^{n+1}$. Therefore, $Rp_{U_0^{n+1}*}F_{U_0^{n+1}}$ carries an action of $\sum_{n+1}$ and endomorphisms $\operatorname{tr}_{[a]_{X^{n+1}}}$ that mutually commute. In particular, we have the complex $[Rp_{U_0^{n+1}*}F_{U_0^{n+1}}]_{\mathrm{sgn}}$ with $\operatorname{tr}_{[a]_{X^{n+1}}}$-action. By 6.1.2 one has a canonical isomorphism

$$[Rp_{U_0^{n+1}*}F_{U_0^{n+1}}]_{\mathrm{sgn}}(n)[n] = Rp_{U*}G_U^{(n)} \tag{6.1.8}$$

that identifies $\operatorname{tr}_{[a]_{X^{n+1}}}$ with $\operatorname{tr}_a^{(n)}$.

Consider the action of the involution $[-1]_X$ on $Rp_{X*}F_X$. One has $[Rp_{X*}F_X]_- = R^1p_{X*}F[-1] = \mathscr{H}'[-1]$ where the index $-$ means anti-invariants of $[-1]_X$-action. We get a canonical isomorphism

$$H^{2+n}_{\mathrm{abs}}(X \times_B U_0^{n+1}, F(n+1))_{-,\mathrm{sgn}} = \operatorname{Ext}^1_U(p_U^*\mathscr{H}, G_U^{(n)}(1)). \tag{6.1.9}$$

Here the index $(-, \mathrm{sgn})$ means a subspace on which $[-1]_X \times \operatorname{id}_{U_0^{n+1}}$ and $\operatorname{id}_X \times \sum_{n+1}$ act as -1 and sgn respectively. The isomorphism is the composition $H^{2+n}_{\mathrm{abs}}(X \times U_0^{n+1}, F(n+1))_{-,\mathrm{sgn}} \overset{\sim}{\to} H^1_{\mathrm{abs}}(B, \mathscr{H}' \otimes Rp_{U*}G_U^{(n)}(1)) \overset{\sim}{\to} \operatorname{Ext}^1_U(p_U^*\mathscr{H}, G_U^{(n)}(1))$ where the first arrow is the Künneth isomorphism and the second is the adjunction map.

Assume we have a vector space M equipped with an action of mutually commuting operators tr_a, where $a \neq 0$ is an integer prime to characteristics of B. Put $M^i = \{m \in M : \operatorname{tr}_a m = a^i m\} \subset M$.

6.1.10. LEMMA. (i). *One has canonical isomorphisms*

$$[R^1p_{U*}G_U^{(n)}(1)]^1 = \mathscr{H}, \qquad [R^ip_{U*}G_U^{(n)}(1)]^1 = 0$$

for $i \neq 1$, *compatible with projections* $G^{(n)} \to G^{(n+1)}$.

(ii) *The residue at* $X^n = X^{\{1,\dots,n\}} \subset X^{n+1}$ *defines isomorphisms*

$$\cdots \xrightarrow{\sim} H^{n+i+1}_{\text{abs}}(U_0^{n+1}, F(n+1))^1_{\text{sgn}} \xrightarrow{\sim} H^{n+1}_{\text{abs}}(U_0^n, F(n))^1_{\text{sgn}}$$
$$\xrightarrow{\sim} \cdots \xrightarrow{\sim} H^{i+1}_{\text{abs}}(U_0^1, F(1))^1_{\text{sgn}} = H^{i+1}_{\text{abs}}(U, F(1))^1 = H^i_{\text{abs}}(B, \mathscr{H}).$$

PROOF. (i) We have computed $R^i p_{U*} G_U^{(n)}(1)$ in 1.3.2. So $R^0 p_{U*} G_U^{(n)}(1) = R^0 p_{U*} \operatorname{gr}^n G_U(1) = S^n(1)$, and tr_a acts on it as multiplication by a^{n+2} (see 6.1.4; note that tr_a acts on $R^0 p_{U*} F_U = R^0 p_{X*} F_X$ as multiplication by a^2). The construction of the isomorphism $R^1 p_{U*} G_U^{(n)}(1) = I/I^{n+2} = \prod_{1 \le i \le n+1} S^i$ from 1.3.2 shows, by compatibility of trace maps and residues, that tr_a acts on $S^i \subset R^1 p_{U*} G_U^{(n)}(1)$ as multiplication by a^i. This implies (i), and (ii) follows from (i) and 6.1.8. □

Note that $H^\cdot_{\text{abs}}(U, F(1))^1$ coincides with the subspace $H^\cdot_{\text{abs}}(U, F(1))_- \subset H^\cdot_{\text{abs}}(U, F(1))$ of anti-invariants of the involution $[-1]_X$.

Consider the base change of our picture by $p_X : X \to B$. The space $X \times U_0^{n+1} = X \times_B U_0^{n+1}$ carries an involution $[-1]_X$ that acts in each cohomology space commuting with the $\sum_{n+1}$- and $\operatorname{tr}_{[a]_{X^{n+1}}}$-actions. According to 6.1.10(ii) we have the isomorphisms

(6.1.11)
$$\cdots \xrightarrow{\sim} H^{2+n}_{\text{abs}}(X \times U_0^{n+1}, F(n+1))^1_{-,\text{sgn}} \xrightarrow{\sim} \cdots \xrightarrow{\sim} H^2_{\text{abs}}(X \times U, F(1))^1_{-,\text{sgn}}$$
$$= H^2_{\text{abs}}(X \times U, F(1))_{-,-}.$$

Note also that the Künneth formula provides the isomorphism $H^2_{\text{abs}}(X \times U, F(1))_{-,-} = \operatorname{End} \mathscr{H}$. This endomorphism identifies $\operatorname{cl}\Delta_{-,-} :=$ $(-,-)$-component of the class of diagonal (which is the same as the Kuenneth $(1,1)$-component) with $\operatorname{id}_{\mathscr{H}}$.

6.1.12. Denote by $\mathscr{P}^{(n)} \in H^{2+n}_{\text{abs}}(X \times U_0^{n+1}, F(n+1))^1_{-,\text{sgn}}$ the classes that correspond to $\operatorname{cl}\Delta_{-,-}$ via 6.1.11. The isomorphism 6.1.9 identifies $\mathscr{P}^{(n)}$ with a compatible system of extension classes. Put

$$\mathscr{P} = \varprojlim \mathscr{P}^{(n)} \in \operatorname{Ext}^1_U(p_U^* \mathscr{H}, G_U(1)).$$

By 1.3.14 this $\mathscr{P}$ coincides with the class of the elliptic polylogarithm extension from 1.3.13. Our aim is to show that the classes $\mathscr{P}^{(n)}$ come from canonical classes in absolute motivic cohomology.

6.2. Eigenvalues of tr_a on absolute motivic cohomology. We need some preliminary information about the action of trace operators on the absolute motivic cohomology of our varieties.

Consider the scheme X^{n+1} equipped with the $\sum_{n+1}$-action; let $X^{(n)} \subset X^{n+1}$ be the kernel of the sum map $\sigma^{n+1} : X^{n+1} \to X$.

6.2.1. LEMMA. *One has canonical isomorphisms*

$$H^i_{\mathscr{M}}(X^{(n)}, \mathbb{Q}(j))_{\mathrm{sgn}} = H^{i-n}_{\mathscr{M}}(\mathrm{Sym}^n H^1(X)(j)),$$
$$H^i_{\mathscr{M}}(X^{n+1}, \mathbb{Q}(j))_{\mathrm{sgn}} = H^{i-n}_{\mathscr{M}}(\mathrm{Sym}^n H^1(X)(j)) \oplus H^{i-n-1}_{\mathscr{M}}(\mathrm{Sym}^{n+1} H^1(X)(j) \oplus \mathrm{Sym}^{n-1} H^1(X)(j-1)) \oplus H^{i-n-2}_{\mathscr{M}}(\mathrm{Sym}^n H^1(X)(j-1)).$$

PROOF. Let us compute the motives $(\Lambda^{\cdot} H^1(X^{n+1}))_{\mathrm{sgn}}$, $(\Lambda^{\cdot} H^1(X^{(n)}))_{\mathrm{sgn}}$. One has

$$\Lambda^{\cdot} H^1(X^{n+1}) = \Lambda^{\cdot}(H^1(X)^{n+1}) = (\Lambda^{\cdot} H^1(X))^{\otimes n+1} = (\mathbb{Q}\oplus H^1(X)\oplus\mathbb{Q}(-1))^{\otimes n+1},$$

and the elementary representation theory of $\sum_{n+1}$ says that $(\Lambda^{\cdot} H^1(X^{n+1}))_{\mathrm{sgn}}$ lies in the sum of those components $(\Lambda^{i_1} H^1(X)) \otimes \cdots \otimes (\Lambda^{i_{n+1}} H^1(X))$ where at most one of the indices i_j equals zero and at most one of them equals two. Therefore, $(\Lambda^n H^1(X^{n+1}))_{\mathrm{sgn}} = \mathrm{Sym}^n H^1(X)$, $(\Lambda^{n+1} H^1(X^{n+1}))_{\mathrm{sgn}} = \mathrm{Sym}^{n+1} H^1(X) \oplus \mathrm{Sym}^{n-1} H^1(X)(-1)$, $(\Lambda^{n+2} H^1(X^{n+1}))_{\mathrm{sgn}} = \mathrm{Sym}^n H^1(X)(-1)$, and $(\Lambda^j H^1(X^{n+1}))_{\mathrm{sgn}} = 0$ for $j \neq n, n+1, n+2$. On the other hand, $\Lambda^{\cdot} H^1(X^{n+1}) = (\Lambda^{\cdot} H^1(X^{(n)})) \otimes \Lambda^{\cdot} H^1(X)$ (since $H^1(X^{n+1}) = H^1(X^{(n)}) \oplus H^1(X)$ because of the $\sum_{n+1}$-invariant isomorphism

$$X^{(n)}_{\mathbb{Q}} \times X_{\mathbb{Q}} \xrightarrow{\sim} X^{n+1}_{\mathbb{Q}}, \qquad ((t_1, \dots, t_{n+1}), t) \mapsto (t_1 + \frac{t}{n+1}, \dots, t_{n+1} + \frac{t}{n+1})).$$

Hence $(\Lambda^{\cdot} H^1(X^{n+1}))_{\mathrm{sgn}} = (\Lambda^{\cdot} H^1(X^{(n)}))_{\mathrm{sgn}} \otimes \Lambda^{\cdot} H^1(X)$; so $\Lambda^n(H^1(X^{(n)}))_{\mathrm{sgn}} = \mathrm{Sym}^n H^1(X)$ and $(\Lambda^j H^1(X^{(n)}))_{\mathrm{sgn}} = 0$ for $j \neq n$. Now our lemma follows from the motivic decomposition of the absolute cohomology groups (see 5.5). □

Now consider the action of the multiplication by a isogenies $[a]_{X^{(n)}}$, $[a]_{X^{n+1}}$, where $a \in \mathbb{Z}$, $a \neq 0$:

6.2.2. COROLLARY. (i) *The operators* $[a]_{X^{(n)}*}$ *and* $[a]^*_{X^{(n)}}$ *act on* $H^i_{\mathscr{M}}(X^{(n)}, \mathbb{Q}(j))_{\mathrm{sgn}}$ *as multiplication by* a^n.

(ii) *The operators* $[a]^*_{X^{n+1}}$, $[a]_{X^{n+1}*}$ *preserve the above three-term decomposition of* $H^i_{\mathscr{M}}(X^{n+1}, \mathbb{Q}(j))_{\mathrm{sgn}}$ *and act on its terms as multiplication by* a^n, a^{n+1}, a^{n+2} *and* a^{n+2}, a^{n+1}, a^n *respectively.*

6.2.3. Put $X^{n+1}_0 := X^{n+1} \backslash X^{(n)}$. Since $X^{(n)}$ is a retract of X^{n+1}, the long exact cohomology sequence of the pair (X^{n+1}, X^{n+1}_0) reduces to the short exact sequences

$$0 \to H^{i-2}_{\mathscr{M}}(X^{(n)}, \mathbb{Q}(j-1)) \to H^i_{\mathscr{M}}(X^{n+1}, \mathbb{Q}(j)) \to H^i_{\mathscr{M}}(X^{n+1}_0, \mathbb{Q}(j)) \to 0,$$

where the first arrow is the Gysin map. The group $\sum_{n+1}$ and the operators $[a]_{X^{n+1}*}$ act on this sequence. The computations in the proof of 6.2.1 show that the Gysin map $\Lambda^n H^1(X^{(n)})(-1) \to \Lambda^{n+2} H^1(X^{n+1})$ induces an isomorphism between the sgn-components. Therefore,

$$(\Lambda^{\cdot} H^1(X^{n+1}))_{\mathrm{sgn}}/(\Lambda^{\cdot} H^1(X^{(n)})_{\mathrm{sgn}}(-1) \\ = (\Lambda^n H^1(X^{n+1}))_{\mathrm{sgn}} \oplus (\Lambda^{n-1} H^1(X^{n+1}))_{\mathrm{sgn}},$$

which implies

6.2.4. COROLLARY. *The operators* $[a]_{X^{n+1}*}$ *act on* $H^i_{\mathcal{M}}(X_0^{n+1}, \mathbb{Q}(j))_{\mathrm{sgn}}$ *with eigenvalues* a^{n+2}, a^{n+1}.

6.3. The motivic polylogarithm classes. Consider now the absolute motivic cohomology groups of the schemes $U_0^{n+1} = X_0^{n+1} \cap U^{n+1}$ from 6.1.7. These groups carry the mutually commuting actions of $\sum_{n+1}$ and the operators $\mathrm{tr}_{[a]_{X^{n+1}}}$. We also have the residue along $X^n \subset X^{n+1}$ maps $H^{\cdot}_{\mathcal{M}}(U_0^{n+1}, \mathbb{Q}(*)) \to H^{\cdot-1}_{\mathcal{M}}(U_0^n, \mathbb{Q}(*-1))$ that commute with the $\sum_n \subset \sum_{n+1}$-action and the trace operators. They induce maps between the sgn-components.

6.3.1. LEMMA. *One has a canonical spectral sequence that converges to* $H^{\cdot}_{\mathcal{M}}(U_0^{n+1}, \mathbb{Q}(*))_{\mathrm{sgn}}$ *with the first term* $E_1^{p,q}$ *equal to*

$$H^{2p+q}_{\mathcal{M}}(X_0^{n+p+1}, \mathbb{Q}(p+*))_{\mathrm{sgn}}$$

for $0 \geq p \geq -n$ *and zero otherwise. The trace operators act naturally on this spectral sequence, and* $\mathrm{tr}_{[a]_{X^{n+1}}}$ *acts on the above* $E_1^{p,q}$ *as* $\mathrm{tr}_{[a]_{X^{n+p+1}}}$.

PROOF. The complement $X_0^{n+1} \setminus U_0^{n+1}$ is a divisor with normal crossing in X_0^{n+1}. The corresponding strata (intersections of components of the divisor) are precisely the X_0^I's for nonempty subsets $I \subset \{1, \ldots, n+1\}$. We get a standard spectral sequence that converges to $H^{\cdot}_{\mathcal{M}}(U_0^{n+1}, \mathbb{Q}(*))$ with the first term $\widetilde{E}_1^{p,q}$ equal to $\bigoplus_{I:|I|=n+p+1} H^{2p+q}_{\mathcal{M}}(X_0^I, \mathbb{Q}(p+*))$ for $0 \geq p \geq -n$ and zero otherwise. The group $\sum_{n+1}$ and the trace operators act on this spectral sequence. The desired spectral sequence is the sgn-component of $\widetilde{E}_r^{p,q}$. □

6.3.2. LEMMA. (i). *There is a unique grading on* $H^{\cdot}_{\mathcal{M}}(U_0^{n+1}, \mathbb{Q}(*))_{\mathrm{sgn}}$ *and the spectral sequence of* 6.3.1 *such that the* j*-component of the grading is the (generalized) eigenspace of* $\mathrm{tr}_{[a]_{X^{n+1}}}$ *with eigenvalue* a^j, *for any* $a \in \mathbb{Z}$.

(ii) *For given* j *the* j*-component of* $E_1^{p,q}$ *vanishes if* $p \neq j-n-1, j-n-2$.

PROOF. (i) follows from 6.2.2 and 6.3.1, (ii) is 6.2.4. □

6.3.3. Consider the 1-component $H^{\cdot}_{\mathcal{M}}(U_0^{n+1}, \mathbb{Q}(*))^1_{\mathrm{sgn}}$. By 6.3.2(ii), 6.3.1 the 1-component of the spectral sequence vanishes unless $p = -n$; i.e., we have $H^{\cdot}_{\mathcal{M}}(U_0^{n+1}, \mathbb{Q}(*))^1_{\mathrm{sgn}} = H^{\cdot-n}_{\mathcal{M}}(X_0^1, \mathbb{Q}(*-n))^1_{\mathrm{sgn}} = H^{\cdot-n}_{\mathcal{M}}(U, \mathbb{Q}(*-n))^1_{\mathrm{sgn}}$. According to the construction of the spectral sequence this isomorphism is the composition of residues along the divisors (last coordinate $= 0$); i.e., we have a sequence of isomorphisms

$$\cdots \xrightarrow{\sim} H^{\cdot+n}_{\mathscr{M}}(U_0^{n+1}, \mathbb{Q}(*+n))^1_{\mathrm{sgn}} \underset{\sim}{\xrightarrow{\mathrm{Res}}} H^{\cdot+n-1}_{\mathscr{M}}(U_0^n, \mathbb{Q}(*+n-1))^1_{\mathrm{sgn}}$$
$$\underset{\sim}{\xrightarrow{\mathrm{Res}}} \cdots \underset{\sim}{\xrightarrow{\mathrm{Res}}} H^{\cdot}_{\mathscr{M}}(U, \mathbb{Q}(*))^1_{\mathrm{sgn}} = H^{\cdot}_{\mathscr{M}}(U, \mathbb{Q}(*))^1.$$

6.3.4. Now we may proceed exactly as at the end of §6.1. Make a base change by $\rho_X: X \to B$, and consider the groups $H^{2+n}_{\mathscr{M}}(X \times U_0^{n+1}, \mathbb{Q}(n+1))^1_{\mathrm{sgn}}$. The corresponding residue maps identify these groups by 6.3.3. The involutions $[-1] \times \mathrm{id}_{U_0^{n+1}}$ act on these groups in a compatible way; denote by the lower index the corresponding anti-invariants. We get the isomorphisms

$$\cdots \underset{\sim}{\xrightarrow{\mathrm{Res}}} H^{2+n}_{\mathscr{M}}(X \times U_0^{n+1}, \mathbb{Q}(n+1))^1_{-,\mathrm{sgn}}$$
$$\underset{\sim}{\xrightarrow{\mathrm{Res}}} H^{2+n-1}_{\mathscr{M}}(X \times U_0^n, \mathbb{Q}(n))^1_{-,\mathrm{sgn}}$$
$$\underset{\sim}{\xrightarrow{\mathrm{Res}}} \cdots \underset{\sim}{\xrightarrow{\mathrm{Res}}} H^2_{\mathscr{M}}(X \times U, \mathbb{Q}(1))^1_{-,\mathrm{sgn}}$$
$$= H^2_{\mathscr{M}}(X \times U, \mathbb{Q}(1))_{-,-} = \mathrm{CH}^1(X \times U) \otimes \mathbb{Q}_{-,-},$$

where $-,-$ means the component on which both $[-1]\times \mathrm{id}_U$ and $\mathrm{id}_X \times [-1]$ act as multiplication by -1. Let $\mathscr{P}^{(0)}_{\mathscr{M}} \in H^2_{\mathscr{M}}(X \times U, \mathbb{Q}(1))_{-,-}$ be the $(-,-)$ component of the class of the diagonal $U \hookrightarrow X \times U$ (which is $\frac{1}{2}$ (class of the diagonal minus class of the antidiagonal)), and let $\mathscr{P}^{(n)}_{\mathscr{M}} = \mathscr{P}^{(n)}_{\mathscr{M}}(X)$ be the corresponding elements in $H^{2+n}_{\mathscr{M}}(X \times U_0^{n+1}, \mathbb{Q}(n+1))^1_{-,\mathrm{sgn}}$ such that $\mathrm{Res}\,\mathscr{P}^{(n)}_{\mathscr{M}} = \mathscr{P}^{(n-1)}_{\mathscr{M}}$.

6.3.5. Definition. The classes $\mathscr{P}^{(n)}_{\mathscr{M}} = \mathscr{P}^{(n)}_{\mathscr{M}}(X)$ are called *motivic elliptic polylogarithm classes.*

According to 6.1.11, 6.1.12 the regulator map $H^{\cdot}_{\mathscr{M}}(?, \mathbb{Q}(*)) \to H^{\cdot}_{\mathrm{abs}}(?, F(*))$ maps $\mathscr{P}^{(n)}_{\mathscr{M}}$ to $\mathscr{P}^{(n)}$ (since the regulator commutes with residues). The $\mathscr{P}^{(n)}_{\mathscr{M}}$ have the usual functorial properties:

6.3.6. Lemma. (i) *The classes $\mathscr{P}^{(n)}_{\mathscr{M}}$ are compatible with base change (for our family of elliptic curves).*

(ii) *If $\varphi: X \to X'$ is an isogeny of degree m then the trace map $\mathrm{tr}_{\varphi^{n+2}}$: $H^{2+n}_{\mathscr{M}}(X \times U_0^{n+1}, \mathbb{Q}(n+1)) \to H^{2+n}_{\mathscr{M}}(X' \times U_0'^{n+1}, \mathbb{Q}(n+1))$ sends $\mathscr{P}^{(n)}_{\mathscr{M}}(X)$ to $m\mathscr{P}^{(n)}_{\mathscr{M}}(X')$.*

Proof. (i) is clear since all the above construction was compatible with base change. Since $\mathrm{tr}_{\varphi^{n+2}}$ commutes with the action of all the standard operators, it maps $H^{\cdot}_{\mathscr{M}}(X \times U_0^{n+1}, \mathbf{Q}(*))^1_{-,\mathrm{sgn}}$ to $H^{\cdot}_{\mathscr{M}}(X' \times U_0'^{n+1}, \mathbb{Q}(*))^1_{-,\mathrm{sgn}}$. Therefore, (ii) follows since $\mathrm{tr}_{\varphi^{n+2}}$ commutes with the residues and obviously $\mathrm{tr}_{\varphi^2}(\mathscr{P}^{(0)}_{\mathscr{M}}(X)) = m\mathscr{P}^{(0)}_{\mathscr{M}}(X')$. □

6.4. Eisenstein classes. Let $x \in X(B)$ be a nonzero torsion point, i.e., $x \in U(B)$ and $[N]x = 0$ for a certain nonzero integer N. Put $X^{(n)}_{(x)} := (\sigma^{n+1})^{-1}(x) \subset X^{n+1}$, $U^{(n)}_{(x)} := X^{(n)}_{(x)} \cap U^{n+1} \subset U_0^{n+1}$, and denote by $\mathscr{P}^{(n)}_{\mathscr{M}x} \in H^{2+n}_{\mathscr{M}}(X \times U^{(n)}_{(x)}, \mathbb{Q}(n+1))_{-,\mathrm{sgn}}$ the restriction of $\mathscr{P}^{(n)}_{\mathscr{M}}$ to the subspace $X \times U^{(n)}_{(x)} \subset X \times U_0^{n+1}$. Let us show that similar to §2 this class defines a collection of Eisenstein classes $\mathscr{E}^{i+2}_{\mathscr{M}x} \in H^{i+1}_{\mathscr{M}}(X^{(i)}, \mathbb{Q}(i+1))_{\mathrm{sgn}}$.

Consider the multiplication by $N+1$ isogeny $[N+1]_{X^{n+1}}$. It preserves $(\sigma^{n+1})^{-1}(x)$ and $X^{n+1}\backslash U^{n+1}$; therefore, we have the corresponding trace operator $\mathrm{tr}_{[N+1]_{U^{(n)}_{(x)}}} \in \mathrm{End}\, H^{\cdot}_{\mathscr{M}}(U^{(n)}_{(x)}, \mathbb{Q}(*))$ that commutes with the $\sum_n$-action.

6.4.1. Lemma. *One has a canonical isomorphism*

$$H^{\cdot}_{\mathscr{M}}(U^{(n)}_{(x)}, \mathbb{Q}(*))_{\mathrm{sgn}} = H^{\cdot}_{\mathscr{M}}(X^{(n)}, \mathbb{Q}(*))_{\mathrm{sgn}} \oplus H^{\cdot-1}_{\mathscr{M}}(X^{(n-1)}, \mathbb{Q}(*-1))_{\mathrm{sgn}} \oplus \cdots \oplus H^{\cdot-n+1}_{\mathscr{M}}(X^{(1)}, \mathbb{Q}(*-n+1))_{\mathrm{sgn}}$$

such that $\mathrm{tr}_{[N+1]_{X^{n+1}}}$ *acts on* $H^{\cdot-i}_{\mathscr{M}}(X^{(n-i)}, \mathbb{Q}(*-i))_{\mathrm{sgn}}$ *as multiplication by* $(N+1)^{n-i}$.

Proof. Just as in 6.3.1 we have a canonical spectral sequence that converges to $H^{\cdot}_{\mathscr{M}}(U^{(n)}_{(x)}, \mathbb{Q}(*))_{\mathrm{sgn}}$ with the first term

$$E_1^{p,q} = H^{2p+q}_{\mathscr{M}}(X^{(n+p)}_{(x)}, \mathbb{Q}(*+p))_{\mathrm{sgn}}$$

for $0 \geq p \geq 1-n$ and zero otherwise. The operator $\mathrm{tr}_{[N+1]_{U^{(n)}_{(x)}}}$ acts on this spectral sequence; the action on the above first term coincides with $\mathrm{tr}_{[N+1]_{X^{(n+p)}_{(x)}}}$. Note that a translation by x along some coordinate identifies $X^{(a)}_{(x)}$ with $X^{(a)}$. The corresponding isomorphism $H^{\cdot}_{\mathscr{M}}(X^{(a)}_{(x)}, \mathbb{Q}(*)) = H^{\cdot}_{\mathscr{M}}(X^{(a)}, \mathbb{Q}(*))$ does not depend on a choice of the coordinate (the two choices differ by translation by a finite-order point of $X^{(a)}$ that acts trivially on $H^{\cdot}_{\mathscr{M}}(X^{(a)}, \mathbb{Q}(*))$ by, e.g., 5.1.1); hence, it commutes with the $\sum_{a+1}$-action. It transforms $\mathrm{tr}_{[N+1]_{X^{(a)}}}$ to $\mathrm{tr}_{[N+1]_{X^{(a)}_{(x)}}}$. Therefore, we have a canonical isomorphism $E_1^{p,q} = H^{2p+q}_{\mathscr{M}}(X^{(n+p)}, \mathbb{Q}(*+p))_{\mathrm{sgn}}$ and, according to 6.2.2(i), $\mathrm{tr}_{[N+1]_{U^{(n)}_{(x)}}}$ acts on $E_1^{p,q}$ as multiplication by $(N+1)^{n+p}$. We see that the spectral sequence degenerates, and the decomposition of $H^{\cdot}_{\mathscr{M}}(U^{(n)}_{(x)}, \mathbb{Q}(*))_{\mathrm{sgn}}$ by eigenspaces of $\mathrm{tr}_{[N+1]_{U^{(n)}_{(x)}}}$ gives 6.4.1. □

6.4.2. Let us make a base change by $p_X : X \to B$ and apply 6.4.1. We get a canonical isomorphism

$$H^{2+n}_{\mathscr{M}}(X \times U^{(n)}_{(x)}, \mathbb{Q}(n+1))_{-,\mathrm{sgn}} = \bigoplus_{n \geq j \geq 0} H^{2+j}_{\mathscr{M}}(X \times X^{(j)}, \mathbb{Q}(j+1))_{-,\mathrm{sgn}}.$$

Let $\pi_j\colon H^{2+j}_{\mathscr{M}}(X \times X^{(j)}, \mathbb{Q}(j+1))_{-,\mathrm{sgn}} \to H^j_{\mathscr{M}}(X^{(j-1)}, \mathbb{Q}(j))_{\mathrm{sgn}}$ be the projection defined by the formula $\pi_j = \alpha^*_{j-1} p_{j*}$ where $p_j\colon X \times X^{(j)} \to X^j$ is the projection $p_j(x, (x_1, \dots, x_{j+1})) = (x + x_1, \dots, x + x_j)$, $x \in X$, $(x_1, \dots, x_{j+1}) \in X^{(j)}$, and $\alpha_{j-1}\colon X^{(j-1)} \hookrightarrow X^j$ is the standard embedding.

6.4.3. Definition. The class $\mathscr{E}^{i+2}_{\mathscr{M}x} \in H^{i+1}_{\mathscr{M}}(X^{(i)}, \mathbb{Q}(i+1))_{\mathrm{sgn}}$ is the image of the $H^{i+3}_{\mathscr{M}}$-component of $\mathscr{P}^{(n)}_{\mathscr{M}x}$, $n > i \geq 0$, by the projection π_{i+1}.

It is easy to see this class does not depend on the choice of $n > i$.

6.4.4. Note that we may apply the above arguments to absolute cohomology groups $H^{\cdot}_{\mathrm{abs}}$ of some theory of mixed sheaves. One has

$$\begin{aligned} H^{2+n}_{\mathrm{abs}}(X \times U^{(n)}_{(x)}, F(n+1))_{-,\mathrm{sgn}} &= H^{\cdot}_{\mathrm{abs}}(B, (Rp_{X*}F)_- \otimes (Rp_{U^{(n)}_{(x)}*}F)_{\mathrm{sgn}}(n+1)) \\ &= H^1(B, \mathscr{H} \otimes (R^n p_{U^{(n)}_{(x)}*}F(n))_{\mathrm{sgn}}). \end{aligned}$$

The same arguments as in 6.4.1 tell us that the mixed sheaf $(R^n p_{U^{(n)}_{(x)}*}F(n))_{\mathrm{sgn}}$ splits; so we have a canonical isomorphism

$$(R^n p_{U^{(n)}_{(x)}*}F(n))_{\mathrm{sgn}} = \bigoplus_{n \geq j \geq 0} (R^j p_{X^{(j)}*}F(j))_{\mathrm{sgn}} = \bigoplus_{n \geq j \geq 0} \mathrm{Sym}^j \mathscr{H}.$$

Now $\mathscr{H} \otimes \mathrm{Sym}^j \mathscr{H} = \mathrm{Sym}^{j+1} \mathscr{H} \oplus (\mathrm{Sym}^{j-1} \mathscr{H})(1)$, and the projection π_j corresponds to the projection on the second multiple. We already know that the components of elliptic polylogarithm class $\mathscr{P}_x$ lie in $(\mathrm{Sym}^{j-1} \mathscr{H})(1)$; see 2.1.5. Therefore, they coincide with their projections by π_j. Finally the standard functoriality of regulator map implies

6.4.5. Lemma. *The image of* $\mathscr{E}^{i+2}_{\mathscr{M}x}$ *by the regulator coincides with the Eisenstein class* $\mathscr{E}^{i+2}_x$ *from* 2.1.5.

6.5. Concluding Remark. One may write down an explicit formula for $\mathscr{P}^{(n)}_{\mathscr{M}}$ in terms of Milnor n-symbols supported on some divisors on $X \times U^{n+1}_0$. We shall publish it elsewhere.

References

[B1] A. Beilinson, *Height pairing between algebraic cycles*, Contemp. Math., vol. 67, Amer. Math. Soc., Providence, RI, 1987, pp. 1–24.

[B2] ——, *Higher regulators of modular curves*, Contemp. Math., vol. 55, part I, Amer. Math. Soc., Providence, RI, 1986, pp. 1–34.

[BBD] A. Beilinson, J. Bernstein, and P. Deligne, *Faisceaux pervers*, Astérisque **100** (1982).

[BD] A. Beilinson and P. Deligne, *Motivic polylogarithm and Zagier's conjecture*, preprint.

[Bl] S. Bloch, *Higher regulators, algebraic K-theory and zeta functions of elliptic curves*, preprint 1979.

[D] P. Deligne, *Le groupe fondamental de la droite projective moins trois points*, Galois Groups over $\mathbb{Q}$ (Y. Ihara, K. Ribet, and J.-P. Serre, eds.), Springer-Verlag, New York, 1989, pp. 80–290.

[J] U. Jannsen, *Mixed motives and algebraic K-theory*, Lecture Notes in Math., vol. 1400, Springer-Verlag, New York, 1990.

[K] M. Kashiwars, *A study of a variation of mixed Hodge structure*, Publ. Res. Inst. Math. Sci. **22** (1986), 991–1024.
[M] I. Macdonald, *Symmetric functions and Hall polynomials*, Clarendon Press, Oxford, 1979.
[S1] M. Saito, *Mixed Hodge modules*, Publ. Res. Inst. Math. Sci. **26** (1990), 221–333.
[S2] ——, *On the formalism of mixed sheaves*, preprint RIMS 1991.
[W] W. Weil, *Elliptic functions according to Eisenstein and Kronecker*, Springer-Verlag, New York, 1976.
[Z1] D. Zagier, *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields*, Proceedings of the Texel conference on Arithmetical Algebraic Geometry, Birkhäuser, Basel, 1991.
[Z2] ——, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. **286** (1990), 613–624.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA

LANDAU INSTITUTE OF THEORETICAL PHYSICS, MOSCOW

p-adic and Characteristic *p* Theory

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Iwasawa Theory and p-adic Deformations of Motives

RALPH GREENBERG

1. Tate motives

Almost ten years ago, Mazur and Wiles proved a fundamental conjecture of Iwasawa that gives a precise link between the ideal class groups of certain towers of cyclotomic fields and the critical values of Dirichlet L-functions. For the Riemann zeta function, this result can be considered as a statement about Tate motives. It has become quite clear in recent years that Iwasawa's conjecture can be viewed in a far more general context. Our purpose in this paper is to state a rather general conjecture (in §4) which we will lead up to in several steps. We begin by describing Iwasawa's conjecture (the Mazur-Wiles theorem) and then a natural reformulation involving p-adic deformations of Tate motives. After that, we shall be able to describe more easily the remainder of the paper.

Let $n \in \mathbb{Z}$. For each prime ℓ, let $\chi_\ell \colon G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{Z}_\ell^*$ denote the ℓth cyclotomic character, which gives the action of $G_{\mathbb{Q}}$ on the ℓ-power roots of unity μ_{ℓ^∞}. The ℓ-adic cohomology of the Tate motive $\mathbb{Q}(n)$ is $\mathbb{Q}_\ell(n)$. This is a one-dimensional $\mathbb{Q}_\ell$-vector space on which $G_{\mathbb{Q}}$ acts by χ_ℓ^n. The L-function for the motive $\mathbb{Q}(n)$ is $L(s, \mathbb{Q}(n)) = \zeta(s+n)$, where $\zeta(s)$ denotes the Riemann zeta function. The value $L(0, \mathbb{Q}(n)) = \zeta(n)$ is a critical value in the sense of Deligne if either $n \geq 2$ and is even or $n \leq -1$ and is odd. Let C denote the set of such n's. If $n \in C$ is negative (and hence odd), then $L(0, \mathbb{Q}(n)) = -B_m/m \in \mathbb{Q}$, where $m = 1-n$ and B_m is the mth Bernoulli number.

Let p be an odd prime. Let $m_1, m_2 \geq 2$ be such that $m_1, m_2 \not\equiv 0 \pmod{p-1}$. The Kummer congruences state that if $m_1 \equiv m_2 \pmod{(p-1)p^t}$,

---

1991 *Mathematics Subject Classification.* Primary 11R23; Secondary 14G10.
Supported, in part, by a National Science Foundation grant.
This paper is in final form and no version of it will be submitted for publication elsewhere.

$t \geq 0$, then

$$(1) \qquad (1-p^{m_1-1})\frac{B_{m_1}}{m_1} \equiv (1-p^{m_2-1})\frac{B_{m_2}}{m_2} \pmod{p^{t+1}\mathbb{Z}_p}.$$

The two sides of this congruence are p-integral (i.e, in $\mathbb{Q} \cap \mathbb{Z}_p$). $\mathbb{Z}_p$ denotes the p-adic integers. For each i, $0 \leq i < p-1$, let $C_i = \{n \in C \mid n \equiv i \pmod{p-1}\}$. Fix an odd i, $i \neq 1$. The above congruences imply that the function $n \to (1-p^{-n})L(0, \mathbb{Q}(n))$ from C_i to $\mathbb{Q}$ is continuous for the p-adic absolute value. Since C_i is dense in $\mathbb{Z}_p$, one sees easily that there will be a unique continuous function $L_p^{(i)}(s)\colon \mathbb{Z}_p \to \mathbb{Z}_p$ that satisfies the interpolation property $L_p^{(i)}(n) = (1-p^{-n})L(0, Q(n))$ for all $n \in C_i$. Note that $(1-p^{-n})$ is the Euler factor for p in $\zeta(s)$, evaluated at $s = n$. $L_p^{(i)}(s)$ is the Kubota-Leopoldt p-adic L-function which is usually denoted by $L_p(s, \omega^j)$, where ω is a certain $\mathbb{Z}_p$-valued Dirichlet character, described below, and $j \equiv 1-i \pmod{p-1}$, $0 \leq j < p-1$. The reason for this notation comes from a more complete statement of the interpolation property, giving the values of $L_p^{(i)}(n)$ for all $n \in \mathbb{Z}$, $n \leq 0$, which we shall also describe below. For $i = 1$ (i.e., $j = 0$), the values $(1-p^{-n})L(0, \mathbb{Q}(n))$ are no longer p-integral for $n \in C_i$, but nevertheless a continuous $\mathbb{Q}_p$-valued function $L_p^{(1)}(s) = L_p(s, \omega^0)$ defined on $\mathbb{Z}_p$, except at $s = 1$, satisfying the above interpolation property, still exists. Kubota and Leopoldt prove that $L_p^{(i)}(s) = L_p(s, \omega^j)$ is analytic for $s \in \mathbb{Z}_p$, except for a simple pole at $s = 1$ if $j = 0$. (See **[I1]** for a complete description of these functions.)

Let $K_\infty = \mathbb{Q}(\mu_{p^\infty})$. Then $\chi = \chi_p$ gives a canonical isomorphism

$$\chi\colon \operatorname{Gal}(K_\infty/\mathbb{Q}) \xrightarrow{\sim} \mathbb{Z}_p^* = \mu_{p-1} \times (1 + p\mathbb{Z}_p).$$

The character ω is obtained by composing χ with the projection to μ_{p-1}. If $\Delta = \operatorname{Gal}(\mathbb{Q}(\mu_p)/\mathbb{Q})$, then ω induces the usual isomorphism $\Delta \cong \mu_{p-1} \cong (\mathbb{Z}/p\mathbb{Z})^*$. We let κ denote the composition of χ with projection to $1 + p\mathbb{Z}_p$. Then κ induces an isomorphism $\Gamma \xrightarrow{\sim} 1 + p\mathbb{Z}_p$, where $\Gamma = \operatorname{Gal}(\mathbb{Q}_\infty/\mathbb{Q})$ and $\mathbb{Q}_\infty$ is the so-called cyclotomic $\mathbb{Z}_p$-extension of $\mathbb{Q}$. We have $\Gamma \cong \mathbb{Z}_p$ and so $\mathbb{Q}_\infty = \bigcup_{r \geq 0} \mathbb{Q}_r$, where $\mathbb{Q}_r$ is the unique subfield of $\mathbb{Q}(\mu_{p^{r+1}})$ of degree p^r over $\mathbb{Q}$. $\operatorname{Gal}(\mathbb{Q}_r/\mathbb{Q}) \cong \mathbb{Z}_p/p^r\mathbb{Z}_p$ is cyclic. We have a canonical isomorphism $\operatorname{Gal}(K_\infty/\mathbb{Q}) \cong \Delta \times \Gamma$. Let $\Lambda = \mathbb{Z}_p[[\Gamma]] = \varprojlim \mathbb{Z}_p[\operatorname{Gal}(\mathbb{Q}_r/\mathbb{Q})]$. Λ is the completed group algebra over $\mathbb{Z}_p$ as defined in the theory of pro-p groups. It is a compact $\mathbb{Z}_p$-algebra and has the following property: Let $\phi\colon \Gamma \to \overline{\mathbb{Q}}_p^*$ be any continuous homomorphism, where $\overline{\mathbb{Q}}_p$ is an algebraic closure of $\mathbb{Q}_p$. Then ϕ can be uniquely extended to a continuous $\mathbb{Z}_p$-algebra homomorphism $\phi\colon \Lambda \to \overline{\mathbb{Q}}_p$. More explicitly, let $\gamma_0 \in \Gamma$ generate a dense subgroup of Γ (e.g., $\gamma_0 = \kappa^{-1}(1+p)$). Let $T_0 = \gamma_0 - 1 \in \Lambda$. Then it turns out that $\Lambda = \mathbb{Z}_p[[T_0]]$. (See **[Wa]**.) If $\phi \in \operatorname{Hom}_{\text{cont}}(\Gamma, \overline{\mathbb{Q}}_p^*)$, then $\phi(\gamma_0)$ is a principal unit in some finite extension $\mathbb{Q}_p(\phi(\gamma_0))$ of $\mathbb{Q}_p$. One obtains the $\mathbb{Z}_p$-algebra homomorphism $\phi\colon \Lambda \to \mathbb{Z}_p[\phi(\gamma_0)] \subseteq \overline{\mathbb{Q}}_p$ by defining $\phi(f(T_0)) = f(\phi(\gamma_0)-1)$ for any $f(T_0) \in \Lambda$. The convergence is obvious. Let $\mathscr{L}$ denote the field of fractions of Λ. If $\theta = fg^{-1} \in \mathscr{L}$, then one can define $\phi(\theta) = \phi(f)\phi(g)^{-1}$ for all $\phi \in \mathrm{Hom}_{\mathrm{cont}}(\Gamma, \overline{\mathbb{Q}}_p^*)$ such that $\phi(g) \neq 0$. Notice that for any $s \in \mathbb{Z}_p$, one can define an element $\phi = \kappa^s$ in $\mathrm{Hom}_{\mathrm{cont}}(\Gamma, \overline{\mathbb{Q}}_p^*)$ by $\phi(\gamma) = \kappa(\gamma)^s$ for all $\gamma \in \Gamma$.

Iwasawa proved that $L_p^{(i)}(s) = \kappa^s(\theta_i)$ for all $s \in \mathbb{Z}_p$ (except $s = 1$ if $i = 1$), where θ_i is an explicitly described element of $\mathscr{L}$. (See **[I1,** Chapter 6].) For $i \neq 1$, $\theta_i = f_i \in \Lambda$, but if $i = 1$ $\theta_1 = f_1 g_1^{-1}$, where $f_1 \in \Lambda^*$ and $g_1 = \gamma_0 - \kappa(\gamma_0)$. The fact that $L_p^{(i)}(s)$ is analytic except for a pole at $s = 1$ if $i = 1$ follows from this. The interpolation property for $L_p^{(i)}(s)$ states that for all $n \in C_i$

(2) $$L_p^{(i)}(n) = \kappa^n(\theta_i) = (1 - p^{-n})L(0, \mathbb{Q}(n)).$$

The element θ_i is clearly determined by this property because

$$\bigcap_{n \in C_i} \ker(\kappa^n\colon \Lambda \to \mathbb{Z}_p) = 0.$$

This is a consequence of the Weierstrass Preparation Theorem.

Let L_∞ denote the maximal abelian extension of K_∞ that is unramified and such that $Y_\infty = \mathrm{Gal}(L_\infty/K_\infty)$ is a pro-p group. L_∞ is just the union of the p-Hilbert class fields of $\mathbb{Q}(\mu_{p^r})$, $r \geq 0$, and so the structure of Y_∞ is clearly related to the structure of the p-primary subgroups of the ideal class groups of the number fields $\mathbb{Q}(\mu_{p^r})$. Now $\mathrm{Gal}(K_\infty/\mathbb{Q}) = \Delta \times \Gamma$ acts on Y_∞ in a natural way (by inner automorphisms). We have

$$Y_\infty = \bigoplus_{i=0}^{p-2} Y_\infty^{\omega^i},$$

where $Y_\infty^{\omega^i} = \{y \in Y_\infty \mid \delta(y) = \omega^i(\delta)y \text{ for all } \delta \in \Delta\}$. Then $Y_\infty^{\omega^i}$ is a $\mathbb{Z}_p$-module on which Γ acts continuously and so one can regard $Y_\infty^{\omega^i}$ as a Λ-module. It turns out to be a finitely generated, torsion Λ-module. (See **[I2]** or **[Wa]**.) Since Λ is a UFD, height-one prime ideals in Λ are principal. The characteristic ideal of any finitely generated, torsion Λ-module (as defined in **[Wa]**) will be principal. The following theorem was conjectured by Iwasawa and proved by Mazur and Wiles **[M-W1]**. A later, more accessible proof using Kolyvagin's notion of Euler systems was found by Rubin **[R1]**.

THEOREM. *Let i be odd, $0 \leq i < p-1$. Then the characteristic ideal of $Y_\infty^{\omega^i}$ is generated by f_i.*

Here f_i is the numerator of the element $\theta_i \in \mathscr{L}$ characterized by (2). The denominator is $g_i = 1$ if $i \neq 1$, and $g_1 = \gamma_0 - \kappa(\gamma_0)$ as before. We want also to point out that g_i is a generator of a characteristic ideal, namely for the torsion Λ-module $\mathbb{Z}_p(1)^{\omega^i}$.

We will now describe a "dual" version of the Mazur-Wiles theorem. Let j be even, $0 \le j < p-1$. Let M_∞ denote the maximal abelian pro-p extension of K_∞, which is unramified except at primes above p. (Hence $L_\infty \subseteq M_\infty$.) Let $X_\infty = \mathrm{Gal}(M_\infty/K_\infty)$. Then $X_\infty^{\omega^j}$ is again a finitely generated, torsion Λ-module. (For odd j, the Λ-rank is 1.) One can relate the structures of $X_\infty^{\omega^j}$ and $Y_\infty^{\omega^i}$ as Λ-modules (where $i + j \equiv 1 \pmod{p-1}$) by an argument using Kummer theory and class field theory (the so-called Spiegelungssatz) and then one obtains the following equivalent version of the above theorem. If $n \in C_j$ (so that $n \ge 2$ now), then $L(0, \mathbb{Q}(n)) = \zeta(n)$ is a critical value and one has $\zeta(n)/\Omega_n = \zeta(1-n) \in \mathbb{Q}$, where $\Omega_n = (2\pi)^n/2(n-1)!$. This follows from the functional equation for $\zeta(s)$. The same congruences (1) imply the existence of a p-adic L-function, which we denote by $L_p^{(j)}(s)$, satisfying the interpolation property

$$L_p^{(j)}(n) = \kappa^n(\theta_j) = (1 - p^{n-1})L(0, \mathbb{Q}(n))/\Omega_n \tag{3}$$

for all $n \in C_j$, where θ_j is a certain element of $\mathscr{L}$. We have $L_p^{(j)}(s) = \kappa^s(\theta_j)$ for $s \in \mathbb{Z}_p$, except possibly $s = 0$. It is clear that actually $L_p^{(j)}(s) = L_p^{(i)}(1-s) = \kappa^{1-s}(\theta_i)$ and hence θ_j could easily be determined from θ_i. One finds that $\theta_j = f_j g_j^{-1}$, where f_j generates the characteristic ideal of $X_\infty^{\omega^j}$ and g_j generates the characteristic ideal of $\mathbb{Z}_p(0)^{\omega^j}$.

We shall now describe a reformulation of the Mazur-Wiles theorem in the language of p-adic deformations. This will lead us to a rather broad generalization of Iwasawa's original conjecture.

Let $g \in G_{\mathbb{Q}}$. If $n_1 \equiv n_2 \pmod{(p-1)p^t}$, then $\chi^{n_1}(g) \equiv \chi^{n_2}(g) \pmod{p^{t+1}\mathbb{Z}_p}$. Thus, $\mathbb{Z}_p(n_1)/p^{t+1}\mathbb{Z}_p(n_1) \cong \mathbb{Z}_p(n_2)/p^{t+1}\mathbb{Z}_p(n_2)$ as $G_{\mathbb{Q}}$-modules. We can "interpolate" the χ^n's in the following way. Fix i, $0 \le i < p-1$. Now χ^n factors through $\mathrm{Gal}(K_\infty/\mathbb{Q}) = \Delta \times \Gamma$. If $n \equiv i \pmod{p-1}$, then $\chi^n|_\Delta = \omega^i$. Also $\chi^n|_\Gamma = \kappa^n$ and so $\chi^n = \omega^i\kappa^n$. We define $\tilde{\chi}$: $\mathrm{Gal}(K_\infty/\mathbb{Q}) \to \Lambda^*$ by $\tilde{\chi}(g) = \omega^i(g)(g|_{\mathbb{Q}_\infty})$. Of course, $\tilde{\chi}$ depends on the value of i. Here $\omega^i(g) \in \mu_{p-1} \subseteq \mathbb{Z}_p^* \subseteq \Lambda^*$ is a "constant" power series and $g|_{\mathbb{Q}_\infty} \in \Gamma \subseteq \Lambda^*$. Then $\tilde{\chi}$ is a continuous homomorphism making the following diagram commutative:

$$\begin{array}{ccc} G_{\mathbb{Q}} & \xrightarrow{\ \tilde{\chi}\ } & \Lambda^* \\ & \searrow^{\chi^n} & \downarrow{\scriptstyle \kappa^n} \\ & & \mathbb{Z}_p^* \end{array}$$

for all $n \in \mathbb{Z}$, $n \equiv i \pmod{p-1}$. The map $\kappa^n: \Lambda \to \mathbb{Z}_p$ is the continuous $\mathbb{Z}_p$-algebra homomorphism obtained from $\kappa^n: \Gamma \to 1 + p\mathbb{Z}_p$. The set of nonzero continuous $\mathbb{Z}_p$-algebra homomorphisms from Λ to $\overline{\mathbb{Q}}_p$ (called "specializations") will be denoted by $\mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$. The elements $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$ correspond to elements $\phi \in \mathrm{Hom}_{\mathrm{cont}}(\Gamma, \overline{\mathbb{Q}}_p^*)$, as mentioned before. If $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$, then $\phi(\Lambda)$ is a closed subring of the ring of integers of the finite extension $\mathbb{Q}_p(\phi(\gamma_0))$ of $\mathbb{Q}_p$. If $\phi(\Lambda) \subseteq \mathbb{Z}_p$, then $\phi = \kappa^s$ for some $s \in \mathbb{Z}_p$. (We are assuming p is odd. If $p = 2$, this is not quite correct.) In the terminology of **[M2]**, $\tilde{\chi}: G_{\mathbb{Q}} \to \mathrm{GL}_1(\Lambda) = \Lambda^*$ is a deformation of $\chi^n: G_{\mathbb{Q}} \to \mathrm{GL}_1(\mathbb{Z}_p) = \mathbb{Z}_p^*$ for any $n \equiv i \pmod{p-1}$. $\tilde{\chi}$ is a deformation of the representation into $\mathrm{GL}_1(\mathbb{Z}/p\mathbb{Z})$ given by ω^i. $\tilde{\chi}$ is characterized by the interpolation property $\kappa^n \circ \tilde{\chi} = \chi^n$ for all $n \in \mathbb{Z}$, $n \equiv i \pmod{p-1}$.

If $\phi = \kappa^n$, where $n \in \mathbb{Z}$ is arbitrary, then $\kappa^n \circ \tilde{\chi} = \chi^n \omega^{i-n}$. We will fix embeddings $\sigma_p: \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$ and $\sigma_\infty: \overline{\mathbb{Q}} \to \mathbb{C}$. Then $\sigma_p^{-1}(\omega^{i-n})$ determines an Artin motive $M(\sigma_p^{-1}(\omega^{i-n}))$ which is a motive with coefficients contained in the field $\mathbb{Q}(\mu_{p-1})$. The representation $\chi^n\omega^{i-n}$ of $G_{\mathbb{Q}}$ arises from the p-adic cohomology of the motive $M_\phi = \mathbb{Q}(n) \otimes M(\sigma_p^{-1}(\omega^{i-n}))$ together with the embedding into $\mathbb{Q}_p$ of $\mathbb{Q}(\mu_{p-1})$ induced by σ_p. We now denote the θ_i from (2) or (3) by $\theta_{\tilde{\chi}}$. For $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$, we put $L_p(\phi, \tilde{\chi}) = \phi(\theta_{\tilde{\chi}})$ (if defined). Thus $L_p(\phi, \tilde{\chi})$ satisfies the interpolation property

$$L_p(\phi, \tilde{\chi}) = \phi(\theta_{\tilde{\chi}}) = \sigma_p(\sigma_\infty^{-1}(c_\phi L(0, M_\phi))),$$

where $\phi = \kappa^n$ with $n \in \mathbb{Z}$ is such that $L(0, M_\phi)$ is a critical value. The complex number $L(0, M_\phi)$ is $L(n, \xi)$, where ξ is the Dirichlet character $\sigma_\infty(\sigma_p^{-1}(\omega^{n-i}))$. If i is odd, this value is critical when $n \leq 0$. If i is even, it is critical for $n \geq 1$. For the interpolation factor c_ϕ, we have $c_{\kappa^n} = (1 - \xi(p)p^{-n})$ if $n \leq 0$ and $c_{\kappa^n} = (1 - \xi^{-1}(p)p^{n-1})\,\mathrm{cond}(\xi)^n/\Omega_n\tau(\xi)$ if $n \geq 1$, where $\mathrm{cond}(\xi) = 1$ or p is the conductor of ξ, $\tau(\xi)$ is the Gaussian sum, and $\Omega_n = (2\pi i)^n/2(n-1)!$ as before. For $n \in C_i$, we obtain the interpolation properties (2) or (3).

To reformulate the Mazur-Wiles theorem, we will associate a "Selmer group" to the deformation $\tilde{\chi}$. Let A be a torsion abelian group on which $G_F = \mathrm{Gal}(\overline{\mathbb{Q}}/F)$ acts continuously (with the discrete topology on A), where F is any subfield of $\overline{\mathbb{Q}}$. The general idea of a Selmer group takes the following form. For every prime ν of F, we define a subgroup L_ν of $H^1(D_\nu, A)$, where D_ν is the decomposition group in G_F for some prime of $\overline{\mathbb{Q}}$ lying over ν. If $\nu | l$, the definition of L_ν usually depends on whether A has l-torsion or not. The associated Selmer group is

$$S_A(F) = \bigcap_\nu \ker(H^1(F, A) \to H^1(D_\nu, A)/L_\nu), \tag{4}$$

where ν runs over all primes of F.

Let $\widehat{\Lambda}$ denote the Pontryagin dual of Λ, $\widehat{\Lambda} = \mathrm{Hom}_{\mathrm{cont}}(\Lambda, \mathbb{Q}_p/\mathbb{Z}_p)$. If $\lambda \in \Lambda$ and $\alpha \in \widehat{\Lambda}$, we define $\lambda\alpha \in \widehat{\Lambda}$ by $\lambda\alpha(x) = \alpha(\lambda x)$ for $x \in \Lambda$. This makes $\widehat{\Lambda}$ into a discrete Λ-module. If I is any ideal of Λ, then let $\widehat{\Lambda}[I] = \{\alpha \in \widehat{\Lambda} \mid \lambda\alpha = 0 \text{ for all } \lambda \in I\}$. Then $\widehat{\Lambda}[I]$ is simply $I^{\perp} = (\Lambda/I)^{\wedge}$. Let $\Lambda(\widetilde{\chi})$ denote the free Λ-module of rank 1 on which $G_{\mathbb{Q}}$ acts by $\widetilde{\chi}$. Let $A(\widetilde{\chi}) = \Lambda(\widetilde{\chi}) \otimes_{\Lambda} \widehat{\Lambda}$, which is just the Λ-module $\widehat{\Lambda}$ on which $g \in G_{\mathbb{Q}}$ acts by multiplication by $\widetilde{\chi}(g) \in \Lambda^*$. We denote $A(\widetilde{\chi})$ more briefly by $\widetilde{A}$. $\widetilde{A}$ is a discrete torsion p-group on which $G_{\mathbb{Q}}$ acts. If I is any ideal of Λ, then we define $\widetilde{A}[I]$ as above. If $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$, let $P_\phi = \ker(\phi)$ (which is in $\mathrm{Spec}(\Lambda)$ and has height one). If $\phi = \kappa^n$ and $n \equiv i \pmod{p-1}$, then the reduction of $\widetilde{\chi}$ modulo P_ϕ gives the Galois action on $\Lambda(\widetilde{\chi})/P_\phi\Lambda(\widetilde{\chi}) \cong \mathbb{Z}_p(n)$. We have $\widetilde{A}[P_\phi] \cong \mathbb{Q}_p(n)/\mathbb{Z}_p(n)$ as a $G_{\mathbb{Q}}$-module.

Let ν be a prime of $\mathbb{Q}$. Let D_ν and I_ν denote the decomposition and inertia subgroups of $G_{\mathbb{Q}}$ for some prime of $\overline{\mathbb{Q}}$ over ν. If $\nu \neq p$, we define $L_\nu = H^1_{\mathrm{unr}}(D_\nu, \widetilde{A}) = \ker(H^1(D_\nu, \widetilde{A}) \to H^1(I_\nu, \widetilde{A}))$. Thus $\ker(H^1(\mathbb{Q}, \widetilde{A}) \to H^1(D_\nu, \widetilde{A})/L_\nu)$ consists of cocycle classes σ such that $\sigma|_{I_\nu}$ is the trivial class. But I_ν acts trivially on $\widetilde{A}$ for $\nu \neq p$ and ∞. For $\nu = \infty$, $D_\nu = I_\nu$ has order 2 and the local cohomology groups vanish, since we assume p is odd. For finite $\nu \neq p$, $H^1(I_\nu, \widetilde{A}) = \mathrm{Hom}(I_\nu, \widetilde{A})$ and hence $\sigma|_{I_\nu}$ trivial means σ is identically zero on I_ν. Also Hom(,) will always mean continuous homomorphisms.

If $\nu = p$, we define L_ν as follows. For any $n \in \mathbb{Z}$, define

$$F^+\mathbb{Z}_p(n) = \begin{cases} \mathbb{Z}_p(n) & \text{if } n \geq 1, \\ 0 & \text{if } n \leq 0. \end{cases} \tag{5}$$

Define $F^+\Lambda(\widetilde{\chi})$ as the unique Λ-submodule of $\Lambda(\widetilde{\chi})$ with the property that

$$\phi(F^+\Lambda(\widetilde{\chi})) = F^+\mathbb{Z}_p(n) \tag{6}$$

for $\phi = \kappa^n$ and all $n \in C_i$. Recall that $n \in C_i$ is equivalent to saying that $n \equiv i \pmod{p-1}$ and that $L(0, \mathbb{Q}(n))$ is a critical value. One sees easily that

$$F^+\Lambda(\widetilde{\chi}) = \begin{cases} \Lambda(\widetilde{\chi}) & \text{if } i \text{ is even}, \\ 0 & \text{if } i \text{ is odd}. \end{cases}$$

We define $F^+\widetilde{A} = (F^+\Lambda(\widetilde{\chi})) \otimes_{\Lambda} \widehat{\Lambda}$, which is either $\widetilde{A}$ itself or 0. For $\nu = p$, we define $L_\nu = \ker(H^1(D_\nu, \widetilde{A}) \to H^1(I_\nu, \widetilde{A}/F^+\widetilde{A}))$. The Selmer group $S_{\widetilde{A}}(\mathbb{Q})$ is defined by (4) with these L_ν's. Since $\widetilde{A}$ is a Λ-module, so is $H^1(\mathbb{Q}, \widetilde{A})$. $S_{\widetilde{A}}(\mathbb{Q})$ is a Λ-submodule. The Λ-module structure comes entirely from that on $\widetilde{A}$.

We can relate the Λ-module structure of $S_{\widetilde{A}}(\mathbb{Q})$ to that of X_∞ or Y_∞ in the following way. Let $G_\infty = \mathrm{Gal}(K_\infty/\mathbb{Q})$. The action of $G_{\mathbb{Q}}$ on $\widetilde{A}$

factors through G_∞. The restriction map $H^1(\mathbb{Q}, \widetilde{A}) \to H^1(K_\infty, \widetilde{A})^{G_\infty}$ is an isomorphism. The kernel is $H^1(G_\infty, \widetilde{A})$. To see that this vanishes, it is enough to show that $H^1(\Gamma, \widetilde{A}) = 0$. But this is $\widetilde{A}/(\gamma_0 - 1)\widetilde{A}$ and $\gamma_0 - 1$ acts on $\widetilde{A}$ by multiplication by $\widetilde{\chi}(\gamma_0) - 1 = \gamma_0 - 1 \in \Lambda$, which is nonzero. The vanishing follows because if $\lambda \in \Lambda$, $\lambda \neq 0$, then we have $\lambda\widehat{\Lambda} = \ker(\text{mult. by } \lambda\colon \Lambda \to \Lambda)^\perp = \widehat{\Lambda}$. (That is, $\widehat{\Lambda}$ is a divisible Λ-module.) The cokernel of the restriction map is contained in $H^2(G_\infty, \widetilde{A})$, which also vanishes because Γ is a free pro-p group and so $H^2(\Gamma, \widetilde{A}) = 0$. Thus $H^1(\mathbb{Q}, \widetilde{A}) \cong H^1(K_\infty, \widetilde{A})^{G_\infty} = \mathrm{Hom}_{G_\infty}(\mathrm{Gal}(K_\infty^{\mathrm{ab}}/K_\infty), \widetilde{A})$. For $\nu \neq p$, we have $I_\nu \subseteq \mathrm{Gal}(\overline{\mathbb{Q}}/K_\infty)$ and the local triviality conditions at such ν assert that $S_{\widetilde{A}}(\mathbb{Q})$ restricts to a subgroup of $\mathrm{Hom}_{G_\infty}(X_\infty, \widetilde{A}) = \mathrm{Hom}_\Gamma(X_\infty^{\omega^i}, \widetilde{A})$, where we make use of the fact that $G_\infty = \Delta \times \Gamma$ and Δ acts on $\widetilde{A}$ by ω^i. Now if i is even and $\nu = p$, then $L_\nu = H^1(D_\nu, \widetilde{A})$ and the local triviality condition is vacuous. If i is odd, then $F^+\widetilde{A} = 0$ and so $L_p = H^1_{\mathrm{unr}}(D_p, \widetilde{A})$. Let π be the prime of K_∞ above p. Let $I_\pi = I_p \cup G_{K_\infty}$ = the inertia subgroup for a prime of $\overline{\mathbb{Q}}$ over π. Exactly as above, one sees that the restriction map $H^1(I_p, \widetilde{A}) \to H^1(I_\pi, \widetilde{A}) = \mathrm{Hom}(I_\pi, \widetilde{A})$ is injective. Consequently,

$$S_{\widetilde{A}}(\mathbb{Q}) \cong \begin{cases} \mathrm{Hom}_\Gamma(X_\infty^{\omega^i}, \widetilde{A}) & \text{if } i \text{ is even}, \\ \mathrm{Hom}_\Gamma(Y_\infty^{\omega^i}, \widetilde{A}) & \text{if } i \text{ is odd}. \end{cases} \tag{7}$$

In both cases, $X_\infty^{\omega^i}$ or $Y_\infty^{\omega^i}$ is a finitely generated, torsion Λ-module. The isomorphism (7) is one of Λ-modules, where the Λ-module structure on $\mathrm{Hom}_\Gamma(\ , \widetilde{A})$ is inherited from that on $\widetilde{A}$.

Let X be any finitely generated, torsion Λ-module. Let $S = \widehat{X} = \mathrm{Hom}(X, \mathbb{Q}_p/\mathbb{Z}_p)$. ($\mathrm{Hom}$, Hom_Γ, Hom_Λ, etc. will always refer to continuous homomorphisms.) We make S a Λ-module by $\lambda s(x) = s(\lambda x)$ for $\lambda \in \Lambda$, $s \in S$, and $x \in X$. (Note that this differs from the action of Λ on S induced by the usual action of $\Gamma \subseteq \Lambda$ on X, which is $\gamma s(x) = s(\gamma^{-1}x)$ for $\gamma \in \Gamma$.) We define the characteristic ideal of the Λ-module S to be precisely the characteristic ideal of $X = \widehat{S}$ as a Λ-module. We say that S is Λ-cotorsion. We want to compare $\mathrm{Hom}_\Gamma(X, \widetilde{A})$ with S as Λ-modules. Now $\gamma \in \Gamma$ acts on $\widetilde{A}$ by multiplication by $\widetilde{\chi}(\gamma) = \gamma \in \Lambda^*$. We see that as Λ-modules $\mathrm{Hom}_\Gamma(X, \widetilde{A}) = \mathrm{Hom}_\Lambda(X, \widetilde{A}) = \mathrm{Hom}_\Lambda(X, \widehat{\Lambda})$. This can be identified as a Λ-module with $\mathrm{Hom}_\Lambda(\Lambda, \widehat{X}) = \widehat{X} = S$. Thus the characteristic ideal of $\mathrm{Hom}_\Gamma(X, \widetilde{A})$ is the same as that of S (and X). For each i, $0 \le i < p-1$, we see from (7) that $S_{\widetilde{A}}(\mathbb{Q})$ is Λ-cotorsion and its characteristic ideal is the same as that of $X_\infty^{\omega^i}$ if i is even or $Y_\infty^{\omega^i}$ if i is odd. This characteristic ideal is generated by f_i, with the notation as before. Also, one easily sees that the ideal (g_i) is the characteristic ideal of

$H^0(\mathbb{Q}, \widetilde{A}) = \widetilde{A}^{G_\infty}$ if i is even and that of $H^0(\mathbb{Q}, \widetilde{A}^*)$ if i is odd. Here we define $\widetilde{A}^* = \mathrm{Hom}(\Lambda(\widetilde{\chi}), \mathbb{Q}_p/\mathbb{Z}_p(1)) = \Lambda(\chi\widetilde{\chi}^{-1}) \otimes_\Lambda \widehat{\Lambda}$, where χ is the cyclotomic character, $\chi\colon \mathrm{Gal}(K_\infty/\mathbb{Q}) \to \mathbb{Z}_p^* \subseteq \Lambda^*$. (The values are constant power series.)

To summarize, for each i, $0 \le i < p-1$, we can define $\widetilde{\chi}\colon G_\mathbb{Q} \to \mathrm{GL}_1(\Lambda)$, which is characterized by the property that for $\phi = \kappa^n$, $n \in C_i$, $\phi\circ\widetilde{\chi}\colon G_\mathbb{Q} \to \mathrm{GL}_1(\mathbb{Z}_p)$ gives the action of $G_\mathbb{Q}$ on the p-adic cohomology of the Tate motive $\mathbb{Q}(n)$. There exists an element $\theta_{\widetilde{\chi}} \in \mathscr{L}$ which is characterized by the property that for $\phi = \kappa^n$, $n \in C_i$, $\phi(\theta_{\widetilde{\chi}}) = c_\phi L(0, \mathbb{Q}(n))$ (which is rational), for suitably defined interpolation factors c_ϕ. The Mazur-Wiles theorem can be equivalently reformulated as follows.

THEOREM. *$\theta_{\widetilde{\chi}} = f/gh$, where f, g, and h generate the characteristic ideals of the Λ-modules $S_{\widetilde{A}}(\mathbb{Q})$, $H^0(\mathbb{Q}, \widetilde{A})$, and $H^0(\mathbb{Q}, \widetilde{A}^*)$, respectively.*

Beginning with the results of Kubota-Leopoldt and Iwasawa from the 1960s described above, there has been considerable work on the general problem of constructing p-adic analogues of the complex L-functions attached to motives. For the motives arising from a modular elliptic curve defined over $\mathbb{Q}$ and also from modular forms of arbitrary weight, the corresponding p-adic L-functions were constructed by Mazur and Swinnerton-Dyer **[M-SwD]** and by Manin (e.g., **[Ma]**) in the early 1970s. One should also see **[M-T-T]** for a more comprehensive treatment of these cases. The literature on this topic is quite extensive and we refer the reader to **[C-P]** and **[P]** (and the references found there) for further discussion. Coates and Perrin-Riou formulate a very general conjecture about the existence and properties of p-adic L-functions in **[C-P]**. The "one-variable" p-adic L-functions which they consider can be viewed as those corresponding to the cyclotomic deformations of motives discussed in §3.

A p-adic deformation of a motive M can be viewed as an analytic family of p-adic representations of $G_\mathbb{Q}$ which includes the p-adic cohomology $H_p(M)$. A p-adic L-function can be thought of as a function whose domain is such a family. For various noncyclotomic deformations, p-adic L-functions have also been constructed. One finds important special cases in **[Kz1, Kz2]**, **[K]**, **[G-S]**, and in **[H3, H4]** (and several other papers of Hida referred to there). We will discuss such p-adic L-functions in §§2 and 4.

Deformations of p-adic representations have been studied by Hida, Mazur, and Wiles **[H-1, H-2]**, **[M2]**, **[M-W2]**, **[W1, W2]**. The idea of formulating a "main conjecture" in this context has seemed natural for a long time. Such a conjecture is described already by Mazur and Tilouine in a certain case **[M-T]**. Our own thoughts (which we explain in this paper) have been influenced by discussions over the years with many people, especially Mazur, Coates, Tilouine, and Hida, to all of whom we are grateful. We will try to motivate the general conjecture stated in §4 by discussing some other special cases that contain features not yet present in the case of Tate motives.

In §2, we consider several p-adic deformations associated to the motive arising from an elliptic curve defined over $\mathbb{Q}$. Section 3 concerns the cyclotomic deformations of a motive M, reformulating the conjecture stated in **[G2, G3]**, although under a somewhat less restrictive hypothesis. The general conjecture of §4 should be viewed with some tentativeness. For one thing, in the cases that are now well understood, the deformation rings are rather simple: formal power series rings in one or several variables. Also, one major flaw in this conjecture is that a sufficiently precise (conjectural) definition of p-adic L-functions is lacking. We shall make some remarks about this later.

One purpose in formulating such conjectures is that it provides a possible approach to studying the connection between the critical values of L-functions and the arithmetic properties of motives. In the case of Dirichlet L-functions and the Hasse-Weil L-function for an elliptic curve, this approach has certainly been quite fruitful. It has contributed to progress on the Birch-Tate conjecture, some of Lichtenbaum's conjectures, and the conjecture of Birch and Swinnerton-Dyer. Possibly one can also hope eventually for some progress on the far more general Bloch-Kato conjecture from this point of view.

2. Elliptic curves

Let E be an elliptic curve defined over $\mathbb{Q}$. Let p be a prime such that E has good, ordinary reduction modulo p. Assuming the Weil-Shimura-Taniyama conjecture for E, Mazur and Swinnerton-Dyer have constructed a p-adic L-function $L_p(s, E)$, defined for all $s \in \mathbb{Z}_p$ **[M-SwD]**. In the language of **[M-SwD]**, this p-adic L-function corresponds to a $\mathbb{Z}_p$-valued measure on $\Gamma = \mathrm{Gal}(\mathbb{Q}_\infty/\mathbb{Q})$. Equivalently, there exists an element $\theta_E \in \Lambda = \mathbb{Z}_p[[\Gamma]]$ such that $L_p(s, E) = \kappa^{s-1}(\theta_E)$ for all $s \in \mathbb{Z}_p$. Let M be the motive over $\mathbb{Q}$ arising from E, so that the p-adic cohomology of M is $V_p(E) = T_p(E) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$, where $T_p(E)$ is the Tate module for E. The complex L-function attached to M is $L(z, M) = L(1+z, E)$, where $L(z, E)$ is the usual Hasse-Weil L-function for $E/\mathbb{Q}$. The only critical value is $L(0, M)$. However, if ϕ is any Dirichlet character, then $L(0, M, \phi^{-1}) = L(0, M \otimes M(\phi))$ is a critical value. Here $M(\phi)$ is the Artin motive attached to ϕ, when we regard ϕ as an Artin character by class field theory. The above complex L-value corresponds to considering ϕ as $\mathbb{C}$-valued by our fixed embedding $\sigma_\infty : \overline{\mathbb{Q}} \to \mathbb{C}$. For brevity, we write $M_\phi = M \otimes M(\phi)$.

Let Ω_E denote the real period of E. Then $L(0, M_\phi)/\Omega_E$ is an algebraic number. By using our fixed embedding $\sigma_p : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$, we can also regard ϕ as a $\overline{\mathbb{Q}}_p$-valued Dirichlet or Artin character. Then θ_E (and hence $L_p(s, E)$) is characterized by the interpolation property $\phi(\theta_E) = \sigma_p \circ \sigma_\infty^{-1}(e_\phi L(0, M_\phi)/\Omega_\phi)$

for all ϕ that factor through Γ (i.e., $\phi \in \widehat{\Gamma}$, where $\widehat{\Gamma} = \mathrm{Hom}(\Gamma, \mu_{p^\infty}) \subseteq \mathrm{Hom}(\Gamma, \overline{\mathbb{Q}}_p)$, continuous homomorphisms as always). Here $\Omega_\phi = \Omega_E \tau(\phi^{-1})$, where $\tau(\phi^{-1}) \in \mathbb{C}$ is the Gaussian sum for ϕ. To describe $e_\phi \in \mathbb{C}$, let α and β be the eigenvalues of the arithmetic Frobenius for p (on $T_l(E)$, $l \neq p$) embedded into $\mathbb{C}$ by σ_∞. Let α be the unit eigenvalue when embedded into $\mathbb{Q}_p$ by $\sigma_p \circ \sigma_\infty^{-1}$. Then $e_\phi = (1 - \alpha^{-1})(1 - \alpha^{-1})$ if $\phi = \phi_0$ (the trivial character) and $e_\phi = \alpha^n$ if ϕ has conductor p^n, $n > 0$.

In **[M]**, Mazur formulates a conjecture relating θ_E to the arithmetic behavior of E in the tower $\mathbb{Q}_\infty = \bigcup \mathbb{Q}_n$. Let $S_E(\mathbb{Q}_\infty)_{\mathrm{class}}$ denote the p-primary subgroup of the classical Selmer group for E over $\mathbb{Q}_\infty$. The natural action of Γ on $S_E(\mathbb{Q}_\infty)_{\mathrm{class}}$ makes it a discrete Λ-module. Then $X_E = \widehat{S}_E(\mathbb{Q}_\infty)_{\mathrm{class}}$ is a finitely generated Λ-module. If X_E is Λ-torsion, we say that $S_E(\mathbb{Q}_\infty)_{\mathrm{class}}$ is Λ-cotorsion and, as in §1, define its characteristic ideal to be that of X_E. If X_E is not Λ-torsion, we let (0) be its characteristic ideal. Assume p is odd. Here is Mazur's conjecture.

CONJECTURE 2.1. *$S_E(\mathbb{Q}_\infty)_{\mathrm{class}}$ is Λ-cotorsion and has characteristic ideal generated by θ_E.*

Rubin has proved the conjecture in the case where E has complex multiplication **[R3]**. Also, in partial support of the conjecture in general, Rohrlich has proved that $\theta_E \neq 0$ **[Ro]**.

One can reformulate this conjecture in a form involving a p-adic deformation of $T_p(E)$. Let $\tilde{\kappa}: \Gamma \to GL_1(\Lambda) = \Lambda^*$ be the inclusion map. Note that if $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$, then $\phi \circ \tilde{\kappa} = \phi|_\Gamma$, which we shall also write simply as ϕ. Thus we identify $\mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$ and $\mathrm{Hom}(\Gamma, \overline{\mathbb{Q}}_p^*)$. We let $\Lambda(\tilde{\kappa})$ denote the free, rank-one Λ-module on which Γ acts via $\tilde{\kappa}$. (It is the "regular representation" of Γ over $\mathbb{Z}_p$.) We let $\rho_0: G_\mathbb{Q} \to \mathrm{Aut}_{\mathbb{Z}_p}(T_p(E))$ give the action of $G_\mathbb{Q}$ on $T_p(E)$. Let $\tilde{\rho}_0: G_\mathbb{Q} \to \mathrm{Aut}_\Lambda(T_p(E) \otimes_{\mathbb{Z}_p} \Lambda)$ be defined by letting $G_\mathbb{Q}$ act by ρ_0 on $T_p(E)$ and trivially on Λ. Then $\tilde{\rho}_0$ is a "constant deformation" of ρ_0. Choose a basis for $T_p(E)$ as a $\mathbb{Z}_p$-module. We then get a basis for $T_p(E) \otimes_{\mathbb{Z}_p} \Lambda$ as a Λ-module. For every $s \in \mathbb{Z}_p$, $\kappa^s: \Lambda \to \mathbb{Z}_p$ induces a homomorphism $\kappa^s: \mathrm{GL}_2(\Lambda) \to \mathrm{GL}_2(\mathbb{Z}_p)$. Identifying $\tilde{\rho}_0$ and ρ_0 with corresponding matrix-valued representations, we have $\kappa^s \circ \tilde{\rho}_0 = \rho_0$ for all $s \in \mathbb{Z}_p$. We let $\tilde{\rho} = \tilde{\rho}_0 \otimes \tilde{\kappa}$. The underlying $G_\mathbb{Q}$-module is $\widetilde{T}_p(E) = T_p(E) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa})$, which is a free Λ-module of rank two. For each n, $\kappa^n \circ \tilde{\rho}$ gives the action of $G_\mathbb{Q}$ on the reduction modulo P_{κ^n}: $\widetilde{T}_p(E)/P_{\kappa^n}\widetilde{T}_p(E) \cong T_p(E) \otimes \mathbb{Z}_p(\kappa^n)$. For $n \equiv 0 \pmod{p-1}$, this is the Tate twist $T_p(E)(n)$. Let $A = E_{p^\infty} = T_p(E) \otimes (\mathbb{Q}_p/\mathbb{Z}_p)$. Define $\widetilde{A} = \widetilde{T}_p(E) \otimes_\Lambda \widehat{\Lambda}$. Then $\widetilde{A}$ is a discrete $G_\mathbb{Q}$-module which is isomorphic to $\widehat{\Lambda} \times \widehat{\Lambda}$ as a Λ-module. Note that $A = \widetilde{A}[P_{\kappa^0}]$, where $P_{\kappa^0} = (T_0) =$ the so-called augmentation ideal in Λ.

We identify $G_{\mathbb{Q}_p}$ with the decomposition group D_p in $G_{\mathbb{Q}}$ for some prime of $\overline{\mathbb{Q}}$ lying over p. Let $\overline{E}$ denote the reduction of E modulo p. We have a surjective reduction homomorphism $T_p(E) \to T_p(\overline{E})$ of $G_{\mathbb{Q}_p}$-modules. Since $\overline{E}$ is ordinary, $T_p(\overline{E})$ is a $\mathbb{Z}_p$-module of rank one. Let $F^+T_p(E)$ be the kernel of the reduction homomorphism. If I_p denotes the inertia subgroup of $D_p = G_{\mathbb{Q}_p}$, then I_p acts trivially on $T_p(E)/F^+T_p(E)$ and by the cyclotomic character χ on $F^+T_p(E)$. We define $F^+\widetilde{T}_p(E) = F^+T_p(E) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa})$. This is a free Λ-submodule of $\widetilde{T}_p(E)$ of rank one, invariant under $G_{\mathbb{Q}_p}$. The quotient $\widetilde{T}_p(E)/F^+\widetilde{T}_p(E)$ is also a free Λ-module of rank one. We define $F^+\widetilde{A} = F^+\widetilde{T}_p(E) \otimes_\Lambda \widehat{\Lambda}$. We can then define a Selmer group $S_{\widetilde{A}}(\mathbb{Q})$ as in (4), where we take $L_\nu \subseteq H^1(D_\nu, \widetilde{A})$ for every prime ν of $\mathbb{Q}$ just as before:

$$(8) \qquad L_\nu = \begin{cases} \ker(H^1(D_\nu, \widetilde{A}) \to H^1(I_\nu, \widetilde{A})) & \text{if } \nu \neq p, \\ \ker(H^1(D_\nu, \widetilde{A}) \to H^1(I_\nu, \widetilde{A}/F^+\widetilde{A})) & \text{if } \nu = p. \end{cases}$$

$S_{\widetilde{A}}(\mathbb{Q})$ is a discrete Λ-module.

The classical Selmer group $S_E(\mathbb{Q}_\infty)_{\text{class}}$ can be described solely in terms of the Galois module $A = E_{p^\infty}$. Let $F^+A = F^+T_p(E) \otimes (\mathbb{Q}_p/\mathbb{Z}_p) \subseteq A$. One can then define $S_A(\mathbb{Q}_\infty)$, again as in (4). It turns out that $S_E(\mathbb{Q}_\infty)_{\text{class}} = S_A(\mathbb{Q}_\infty)$. For a simple proof of this important observation, see **[G3]**. It also turns out that $S_A(\mathbb{Q}_\infty) \cong S_{\widetilde{A}}(\mathbb{Q})$ as Λ-modules. We will give the simple proof of this in a much more general context in §3. The element θ_E can also be characterized in terms of the deformation $\tilde{\rho}$. If $\phi \in \widehat{\Gamma}$, then $\phi \circ \tilde{\rho}\colon G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ is the representation of $G_{\mathbb{Q}}$ on $V_p(E) \otimes_{\mathbb{Q}_p} W_\phi$, where W_ϕ is the one-dimensional $\overline{\mathbb{Q}}_p$-vector space on which $G_{\mathbb{Q}}$ acts by ϕ. This $\overline{\mathbb{Q}}_p$-representation arises from the p-adic cohomology of $M_\phi = M \otimes M(\phi)$ together with our fixed embedding $\sigma_p\colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$. We will now denote θ_E by $\theta_{\tilde{\rho}}$. The interpolation property is

$$(9) \qquad \phi(\theta_{\tilde{\rho}}) = \sigma_p \circ \sigma_\infty^{-1}(e_\phi L(0, M_\phi)/\Omega_\phi),$$

for all $\phi \in \widehat{\Gamma}$, where e_ϕ, Ω_ϕ are as before. The fact that

$$\bigcap_{\phi \in \widehat{\Gamma}} \ker(\phi\colon \Lambda \to \overline{\mathbb{Q}}_p) = \{0\}$$

shows that $\theta_{\tilde{\rho}} \in \Lambda$ is indeed uniquely determined by (9). Mazur's conjecture then states that $S_{\widetilde{A}}(\mathbb{Q})$ is Λ-cotorsion and $\theta_{\tilde{\rho}}$ generates its characteristic ideal.

We want to make a remark about $H^0(\mathbb{Q}, \widetilde{A})$. We have

$$\begin{aligned} \widetilde{A} &= (T_p(E) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa})) \otimes_\Lambda \widehat{\Lambda} \\ &= \mathrm{Hom}_{\mathbb{Z}_p}(\Lambda(\tilde{\kappa}^{-1}), \mathbb{Q}_p/\mathbb{Z}_p) \otimes_{\mathbb{Z}_p} T_p(E) = \mathrm{Hom}_{\mathbb{Z}_p}(\Lambda(\tilde{\kappa}^{-1}), E_{p^\infty}) \end{aligned}$$

as $G_{\mathbb{Q}}$-modules. Now let $B = E_{p^\infty}^{G_{\mathbb{Q}_\infty}}$ = the p-torsion on $E(\mathbb{Q}_\infty)$, which is easily shown to be a finite group. Since the action of $G_{\mathbb{Q}}$ on $\Lambda(\tilde{\kappa}^{-1})$ factors through Γ, we have $H^0(\mathbb{Q}, \widetilde{A}) = \widetilde{A}^{G_{\mathbb{Q}}} = \mathrm{Hom}_\Gamma(\Lambda(\tilde{\kappa}^{-1}), B)$. We can regard both B and $\Lambda(\tilde{\kappa}^{-1})$ as $\Lambda = \mathbb{Z}_p[[\Gamma]]$-modules, using the action of Γ on these $\mathbb{Z}_p$-modules. Then $\Lambda(\tilde{\kappa}^{-1}) \cong \Lambda$ and $H^0(\mathbb{Q}, \widetilde{A}) = \mathrm{Hom}_\Lambda(\Lambda, B) \cong B$ as a Λ-module. (This is slightly inaccurate. One must let Λ act through the involution described in §3. This makes no difference here.) This is a pseudo-null Λ-module, and it has characteristic ideal $(1) = \Lambda$. We define $\widetilde{A}^* = \mathrm{Hom}_{\mathbb{Z}_p}(\widetilde{T}_p(E), \mathbb{Q}_p/\mathbb{Z}_p(1))$. By using the Weil pairing $\mathrm{Hom}(T_p(E), \mathbb{Q}_p/\mathbb{Z}_p(1)) \cong E_{p^\infty}$, one finds that $\widetilde{A}^* \cong \widetilde{T}_p(E)' \otimes_\Lambda \widehat{\Lambda}$, where $\widetilde{T}_p(E)' = T_p(E) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa}^{-1})$. As above, one sees that $H^0(\mathbb{Q}, \widetilde{A}^*)$ is pseudo-null. With the reformulation of the Mazur-Wiles theorem at the end of §1 in mind, it perhaps seems reasonable that $\theta_{\tilde{\rho}} = \theta_E$ is in Λ.

We now consider a quite different deformation of $T_p(E)$ constructed by Hida in some cases. The existence of this deformation is a reflection of congruences between cusp forms of level Np^t, $t \geq 0$, and of varying weights, where N denotes the conductor of E. It is impossible for us to formulate a very precise conjecture in this case because of several unresolved points. One should see **[H1, H2]** and **[M-W2]** for the construction and basic properties. We shall base our discussion on the following hypothesis, although it is unclear to us how generally one should expect it to be valid. First some notation. Let $\Sigma = \{p, \infty\} \cup \{\text{prime factors of } N\}$ and let $\mathbb{Q}_\Sigma$ denote the maximal extension of $\mathbb{Q}$ unramified outside Σ. Let $\Lambda_w = \mathbb{Z}_p[[\Gamma_w]]$, where $\Gamma_w = 1 + p\mathbb{Z}_p$. (Here the subscript w refers to "weight".) Note that $\kappa\colon \Gamma \xrightarrow{\sim} \Gamma_w$ induces an isomorphism $\Lambda \xrightarrow{\sim} \Lambda_w$, but we want to distinguish these rings. Let $\varepsilon\colon \Lambda_w \to \mathbb{Z}_p$ be the continuous $\mathbb{Z}_p$-algebra homomorphism induced from the inclusion map $\varepsilon\colon \Gamma_w = 1 + p\mathbb{Z}_p \to \overline{\mathbb{Q}}_p^*$. Also, for $k \in \mathbb{Z}_p$, the homomorphism $\varepsilon^k\colon \Gamma_w \to \overline{\mathbb{Q}}_p^*$ (raising to the kth power) induces a $\mathbb{Z}_p$-algebra homomorphism $\varepsilon^k\colon \Lambda_w \to \overline{\mathbb{Q}}_p$. Let $\mathscr{L}_w$ denote the fraction field of Λ_w.

HYPOTHESIS W. There exist a compact, local integral domain R containing Λ_w and a continuous representation $\tilde{\sigma}\colon \mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_2(R)$ with the properties:

(1) R is finitely generated as a Λ_w-module.
(2) There exists a continuous $\mathbb{Z}_p$-algebra homomorphism $\psi\colon R \to \mathbb{Z}_p$ such that $\psi \circ \tilde{\sigma}\colon \mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Z}_p)$ gives the Galois action on $T_p(E)$.
(3) Let $k \geq 2$. Let $\phi \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$ be such that $\phi|_{\Lambda_w} = \varepsilon^k$. Then $\phi \circ \tilde{\sigma}\colon \mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ is equivalent (over $\overline{\mathbb{Q}}_p$) to the $\overline{\mathbb{Q}}_p$-representation corresponding to a weight k cusp form (an eigenform

for the Hecke operators) for $\Gamma_1(Np)$. We call the form f_ϕ and regard it as having coefficients in $\overline{\mathbb{Q}}_p$. The trace of the *arithmetic* Frobenius for a prime l not dividing Np is the lth Fourier coefficient of f_ϕ.

(4) Let $T_p(\tilde{\sigma})$ denote the free R-module of rank two on which $G_{\mathbb{Q}}$ acts by $\tilde{\sigma}$. There exists an R-submodule $F^+T_p(\tilde{\sigma})$ which is invariant under the action of $G_{\mathbb{Q}_p}$ and such that the inertia group I_p acts trivially on $T_p(\tilde{\sigma})/F^+T_p(\tilde{\sigma})$. Both $F^+T_p(\tilde{\sigma})$ and $T_p(\tilde{\sigma})/F^+T_p(\tilde{\sigma})$ are free R-modules of rank one.

The ring R should be contained in $h_N^\circ \otimes_{\Lambda_w} \mathscr{L}_w$, where h_N° is Hida's universal ordinary Hecke algebra for level N, and chosen (if possible!) to yield a *free* R-module on which $\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q})$ acts. We denote by $\mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$ the set of continuous, nonzero $\mathbb{Z}_p$-algebra homomorphisms. We shall use this notation for any compact $\mathbb{Z}_p$-algebra R. Since R is integral over Λ_w, every element of $\mathrm{Spec}(\Lambda_w, \overline{\mathbb{Q}}_p)$ is the restriction of at least one and at most d elements of $\mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$, where $d = \mathrm{rank}_{\Lambda_w}(R)$. If $\phi \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$, then ϕ induces a continuous homomorphism (also called ϕ): $\mathrm{GL}_2(R) \to \mathrm{GL}_2(\phi(R))$. $\phi(R)$ is a $\mathbb{Z}_p$-subalgebra of the integers in some finite extension of $\mathbb{Q}_p$.

We assume that E has ordinary reduction at p. As a $G_{\mathbb{Q}_p}$-module, $T_p(E)$ has an unramified quotient $T_p(E)/F^+T_p(E)$ of $\mathbb{Z}_p$-rank one. It is clear that $\psi(F^+T_p(\tilde{\sigma})) = F^+T_p(E)$. More generally, Mazur and Wiles show that the underlying $\overline{\mathbb{Q}}_p$-representation space $V_p(f_\phi)$ corresponding to f_ϕ (where $\phi|_{\Lambda_w} = \varepsilon^k$, $k \in \mathbb{Z}$, $k \geq 2$) also has an unramified quotient $V_p(f_\phi)/F^+V_p(f_\phi)$ of dimension one. The R-submodule $F^+T_p(\tilde{\sigma})$ is characterized by the fact that its image under every such ϕ is contained in $F^+V_p(f_\phi)$. Later, we shall discuss the special case where E has complex multiplication. Hypothesis W is then easy to verify. The deformation ring R will simply be Λ_w. Another quite general case is when the representation of $G_{\mathbb{Q}}$ on $E_p = T_p(E)/pT_p(E)$ is absolutely irreducible (or equivalently, in this case, just irreducible). This is assumption (R) of Mazur-Tilouine **[M-T]**, and their assumption (P) can also be verified, as pointed out to us by J. Tilouine. Theorems 7 and 9 of **[M-T]** then imply the validity of Hypothesis W. For the underlying ring R, one can take the quotient of a local component of h_N° by some minimal prime ideal.

To state an analogue of Iwasawa's and Mazur's conjecture for $\tilde{\sigma}$, we must define a Selmer group and a p-adic L-function. The definition of a Selmer group is quite easy, assuming the above hypothesis. Let $\widehat{R} = \mathrm{Hom}(R, \mathbb{Q}_p/\mathbb{Z}_p)$, which we regard as a discrete R-module in the natural way. Let $\widetilde{B} = T_p(\tilde{\sigma}) \otimes_R \widehat{R}$. As an R-module, $\widetilde{B}$ is isomorphic to $\widehat{R} \times \widehat{R}$. $\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q})$ acts on $\widetilde{B}$ through the first factor $T_p(\tilde{\sigma})$. We define $F^+\widetilde{B} =$

$F^+T_p(\tilde{\sigma}) \otimes_R \widehat{R}$. Thus, just as before, we define a Selmer group $S_{\widetilde{B}}(\mathbb{Q}) \subseteq H^1(\mathbb{Q}, \widetilde{B})$. It is a discrete R-module which is easily shown to be cofinitely generated as an R-module. (That is, its Pontryagin dual is finitely generated.) Briefly, one verifies that $S_{\widetilde{B}}(\mathbb{Q}) \subseteq H^1(\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}), \widetilde{B})$ and that this last group is cofinitely generated over Λ_w.

It seems reasonable to conjecture that $S_{\widetilde{B}}(\mathbb{Q})$ is cotorsion as an R-module. The R-modules $H^0(\mathbb{Q}, \widetilde{B})$ and $H^0(\mathbb{Q}, \widetilde{B}^*)$, where $\widetilde{B}^* = \mathrm{Hom}(T_p(\tilde{\sigma}), \mathbb{Q}_p/\mathbb{Z}_p(1))$, will also play a role in the conjecture below. We will prove that these are actually R-cotorsion. For $p \geq 11$, they are in fact trivial. Let P be the kernel of $\psi\colon R \to \mathbb{Z}_p$. Then $\widetilde{B}[P] = \{b \in \widetilde{B} \mid rb = 0$ for all $r \in P\} \cong T_p(E) \otimes (\mathbb{Q}_p/\mathbb{Z}_p) \cong E_{p^\infty}$. Now the maximal ideal of R is $\mathfrak{m} = P + pR$ and we have $\widetilde{B}[\mathfrak{m}] \cong E_p$. Thus $\widetilde{B}^{G_\mathbb{Q}}[\mathfrak{m}] \cong E_p^{G_\mathbb{Q}} = 0$ for $p \geq 11$ by Mazur's theorem on the p-torsion of $E(\mathbb{Q})$. This implies that $H^0(\mathbb{Q}, \widetilde{B}) = \widetilde{B}^{G_\mathbb{Q}} = 0$ by Nakayama's Lemma (in the "dual" version of cofinitely generated modules over the local Noetherian domain R). Also $\widetilde{B}^*[\mathfrak{m}] = \mathrm{Hom}(T_p(\tilde{\sigma})/\mathfrak{m}T_p(\tilde{\sigma}), \mathbb{Z}/p\mathbb{Z}(1)) \cong \mathrm{Hom}(E_p, \mathbb{Z}/p\mathbb{Z}(1)) \cong E_p$ and hence $H^0(\mathbb{Q}, \widetilde{B}^*) = 0$.

In general, we have $H^0(\mathbb{Q}, \widetilde{B}) \subseteq H^0(\mathbb{Q}_p, \widetilde{B}) = \widetilde{B}^{G_{\mathbb{Q}_p}}$. Now $(F^+\widetilde{B})[P] = F^+E_{p^\infty}$ and so $F^+\widetilde{B}[\mathfrak{m}] \cong \mathbb{Z}/p\mathbb{Z}$ has a nontrivial action of $G_{\mathbb{Q}_p}$ (even I_p). Thus $\widetilde{B}^{G_{\mathbb{Q}_p}} \cap F^+\widetilde{B} = 0$. It follows that $\widetilde{B}^{G_{\mathbb{Q}_p}}$ is mapped injectively into $(\widetilde{B}/F^+\widetilde{B})^{G_{\mathbb{Q}_p}}$. There exist homomorphisms $\tilde{\alpha}, \tilde{\beta}\colon G_{\mathbb{Q}_p} \to \mathrm{GL}_1(R) = R^*$ which given the action of $G_{\mathbb{Q}_p}$ on the R-modules $T_p(\tilde{\sigma})/F^+T_p(\tilde{\sigma})$ and $F^+T_p(\tilde{\sigma})$, respectively. We will use the notations $R(\tilde{\alpha})$ (and $\widehat{R}(\tilde{\alpha})$ for the rank one (and corank one) R-modules on which $G_{\mathbb{Q}_p}$ acts by $\tilde{\alpha}$. Similarly for $R(\tilde{\beta})$, etc. We have an exact sequence

$$0 \to \widehat{R}(\tilde{\beta}) \to \widetilde{B} \to \widehat{R}(\tilde{\alpha}) \to 0.$$

Now $\tilde{\alpha}$ is unramified. Let Frob (in $G_{\mathbb{Q}_p}/I_p$) denote the arithmetic Frobenius and let $\tilde{\alpha}(\mathrm{Frob}) = u \in R^*$. We have $\widetilde{B}/F^+\widetilde{B} \cong \widehat{R}(\tilde{\alpha})$ and $(\widetilde{B}/F^+\widetilde{B})^{G_{\mathbb{Q}_p}} \cong \widehat{R}[I]$, where I is the ideal $(u-1)R$. If $u \neq 1$, this is R-cotorsion with Pontryagin dual $R/(u-1)R$. But if $\phi \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$ is such that $\phi|_{\Lambda_w} = \varepsilon^k$ with $k \in \mathbb{Z}$, $k \geq 2$, then $\phi(u)$ is the eigenvalue of Frob acting on $V_p(f_\phi)/F^+V_p(f_\phi)$. This eigenvalue is algebraic and has complex absolute value > 1 for $k > 2$. Thus $u - 1 \neq 0$ and hence $H^0(\mathbb{Q}_p, \widetilde{B})$ and, a fortiori, $H^0(\mathbb{Q}, \widetilde{B})$ are R-cotorsion, annihilated by $u-1$.

The R-module $H^0(\mathbb{Q}_p, \widetilde{B}^*)$ (and, therefore $H^0(\mathbb{Q}, \widetilde{B}^*)$) turns out to be finite and thus pseudo-null. The argument is not difficult but we shall just give a sketch. We have an exact sequence

$$0 \to \widehat{R}(\chi\tilde{\alpha}^{-1}) \to \widetilde{B}^* \to \widehat{R}(\chi\tilde{\beta}^{-1}).$$

Now $\tilde{\alpha}\tilde{\beta} = \tilde{\delta}|_{G_{\mathbb{Q}_p}}$, where $\tilde{\delta} = \det(\tilde{\sigma})\colon G_{\mathbb{Q}} \to R^*$. If $\phi|_{\Lambda_w} = \varepsilon^k$ as before, then $\phi \circ \tilde{\delta}$ gives the determinant on $V_p(f_\phi)$. Using this one can describe $\tilde{\delta}$ quite precisely, up to a character of finite order. As above, $\widetilde{B}^{G_{\mathbb{Q}_p}}$ maps injectively into $\widetilde{R}(\chi\tilde{\beta}^{-1})^{G_{\mathbb{Q}_p}}$. One finds that $\widetilde{B}[p]^{G_{\mathbb{Q}_p}}$ is finite. It follows that if $\widetilde{B}^{G_{\mathbb{Q}_p}}$ is infinite, then for some $\phi \in \operatorname{Spec}(R, \overline{\mathbb{Q}}_p)$, $\widetilde{B}[P_\phi]^{G_{\mathbb{Q}_p}}$ is also infinite. And so $\phi \circ \tilde{\beta} = \chi$. From this one sees that ϕ has "weight 2", i.e., $\phi\varepsilon^{-2}|_{\Gamma_w}$ is of finite order. Then $\widetilde{B}[P_\phi]$ is (up to isogeny) contained in $A[p^\infty]$, where $A = \operatorname{Jac}(X_1(Np^t))$ for some $t \geq 0$. The finiteness of the H^0's follows from the fact that the torsion on $A(\mathbb{Q}_p)$ is finite for any abelian variety A.

We shall describe the p-adic L-function attached to the deformation $\tilde{\sigma}$ in one special case where we can be fairly precise, namely the case where E has complex multiplication by the ring of integers of an imaginary quadratic field K. Since E has ordinary reduction at p, the prime p splits in K. Let ψ_E denote the grossencharacter of K associated to E. Then $L(z, \psi_E)$ is the Hasse-Weil L-function $L(z, E)$. Using the embedding $\sigma_p\colon K \to \mathbb{Q}_p$, we can also regard ψ_E as a continuous homomorphism $\psi_E\colon G_K \to \mathbb{Z}_p^*$. It describes the action of G_K on E_{π^∞}, where π generates the prime ideal of $\mathscr{O}_K$ determined by σ_p, and E_{π^∞} is the π-power torsion on $E(\overline{\mathbb{Q}})$. Let $K_{\pi^\infty} = K(E_{\pi^\infty})$. Then ψ_E induces an isomorphism $\operatorname{Gal}(K_{\pi^\infty}/K) \xrightarrow{\sim} \mathbb{Z}_p^* = \mu_{p-1} \times (1+p\mathbb{Z}_p)$. Let ω_E and κ_E be composition of ψ_E with projection to the first and second factors, respectively. We identify the second factor with Γ_w. Define $\tilde{\psi}\colon \operatorname{Gal}(K_{\pi^\infty}/K) \to \operatorname{GL}_1(\Lambda_w) = \Lambda_w^*$ by $\tilde{\psi}(g) = \omega_E(g)^2\psi_E(g)^{-1}\kappa_E(g)$ for all $g \in \operatorname{Gal}(K_{\pi^\infty}/K)$. Here $\omega_E(g) \in \mu_{p-1}$ and $\psi_E(g) \in \mathbb{Z}_p^*$ are in Λ_w^* as "constants" and $\kappa_E(g) \in \Gamma_w \subseteq \Lambda_w^*$. Then $\tilde{\psi}$ is a continuous homomorphism making the following diagram commutative:

$$\begin{array}{ccc} G_K & \xrightarrow{\ \tilde{\psi}\ } & \Lambda_w^* \\ & \searrow^{\psi_E^{k-1}} & \downarrow{\scriptstyle \varepsilon^k} \\ & & \mathbb{Z}_p^* \end{array}$$

for all $k \equiv 2 \pmod{p-1}$. Notice the close analogy with the definition of the deformation $\tilde{\chi}$ in §1, except for a shift. We define $\tilde{\sigma}\colon \operatorname{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \operatorname{GL}_2(\Lambda_w)$ by $\tilde{\sigma} = \operatorname{Ind}_K^{\mathbb{Q}}(\tilde{\psi})$. We let $R = \Lambda_w$ and $\psi = \varepsilon^2$. Then $\psi \circ \tilde{\sigma} = \operatorname{Ind}_K^{\mathbb{Q}}(\psi_E)$ gives the action of $G_{\mathbb{Q}}$ on $T_p(E)$. Also for $k \equiv 2 \pmod{p-1}$, $k \geq 2$, we have $\varepsilon^k \circ \tilde{\sigma} = \operatorname{Ind}_K^{\mathbb{Q}}(\psi_E^{k-1})$, which gives the action of $G_{\mathbb{Q}}$ on the Deligne p-adic representation attached to the modular form f_k of weight k whose Mellin transform is $L(z, \psi_E^{k-1})$. The kernel of $\tilde{\sigma}$ corresponds to the field $\mathbb{Q}(E_{p^\infty}) \subseteq \mathbb{Q}_\Sigma$.

Finally, to verify Hypothesis W consider $\tilde{\sigma}|_{D_p}$. Now $D_p \subseteq G_K$ and we have $T_p(\tilde{\sigma}) \cong \Lambda_w(\tilde{\psi}) \times \Lambda_w(\tilde{\psi}')$ as a representation over Λ_w of G_K. Here $\tilde{\psi}'$

is defined just as $\tilde{\psi}$, except starting with the homomorphism $\psi'_E: G_K \to \mathbb{Z}_p^*$ giving the action on $E_{(\pi')^\infty}$, where $\pi\pi' = p$. Then $\varepsilon^k \circ \tilde{\psi}' = (\psi'_E)^{k-1}$ for $k \equiv 2 \pmod{p-1}$. We can assume that D_p and I_p are the decomposition and inertia groups for a prime of $\overline{\mathbb{Q}}$ over π. Then $I_\pi \subseteq \ker(\psi'_E)$ and hence if we define $F^+T_p(\tilde{\sigma}) = \Lambda_w(\tilde{\psi})$, we shall have part (4) of Hypothesis W. The Selmer group $S_{\widetilde{B}}(\mathbb{Q})$ is a Λ_w-module. Its structure as a Λ_w-module can be studied by relating it to either the π-Selmer group for E over K_{π^∞} (essentially along the lines of our discussion of $\tilde{\rho}$ earlier) or to a certain Galois group over K_{π^∞} (along the lines of our discussion of $\widetilde{\chi}$ in §1). In both, a suitably defined Selmer group $S_{\widetilde{A}}(K)$, where $\widetilde{A} = \Lambda_w(\tilde{\psi}) \otimes_{\Lambda_w} \widehat{\Lambda}_w$, would intervene. As a consequence, one deduces that $S_{\widetilde{B}}(\mathbb{Q})$ is Λ_w-cotorsion.

Now we describe the p-adic L-function for $\tilde{\sigma}$. Assume that $k \equiv 2 \pmod{p-1}$, $k \geq 2$. Then $\varepsilon^k \circ \tilde{\sigma}$ gives the action of $G_{\mathbb{Q}}$ on the p-adic cohomology of a certain motive M_k. The L-function over $\mathbb{Q}$ for M_k is $L(z, M_k) = L(z+k-1, \overline{f}_k) = L(z, \overline{\psi}_E^{k-1})$. (Actually $\overline{f}_k = f_k$. The form has real coefficients.) Thus $L(0, M_k) = L(k-1, \overline{\psi}_E^{k-1})$ is a critical value. Let Ω_E denote the real period for E. Let $f \in K$ be a generator for the conductor of ψ_E. Let $\Omega_k = (\Omega_E/f)^{k-1}/12(k-2)!$. Let $\mathscr{A}$ denote the ring of integers in the completion of the maximal unramified extension of $\mathbb{Q}_p$. Then there exists a power series $G(T) \in \mathscr{A}[[T]]$ such that for all $k \equiv 2 \pmod{p-1}$, $k \geq 2$,

(10)
$$\begin{aligned} G((1+p)^k - 1) &= \lambda^{1-k}(\sigma_p \circ \sigma_\infty^{-1}((1-\psi_E(\pi)^{k-1}p^{-1})L(k-1, \psi_E^{k-1})/\Omega_k)) \\ &= \lambda^{1-k}(\sigma_p \circ \sigma_\infty^{-1}((1-\beta^{k-1}p^{-1})L(0, M_k)/\Omega_k)). \end{aligned}$$

Here $\lambda \in \mathscr{A}^*$ occurs as the first coefficient in an isomorphism over $\mathscr{A}$ of the formal group for E and the formal multiplicative group. λ is well defined up to a factor in $\mathbb{Z}_p^*$. Also $\beta = \psi_E(\pi)$. The single Euler-like factor is from the other side of the functional equation: $L(1, \psi_E^{k-1})$. The power series is uniquely determined by (10), except for the small ambiguity in the choice of λ. For the existence and further properties, see, for example, **[C-W2]**. One can think of $\mathscr{A}[[T]]$ as the completed group ring over $\mathscr{A}$ for Γ_w, letting T correspond to $\gamma_w = 1+p$, a topological generator for Γ_w. Then ε^k would send T to $(1+p)^k - 1$. The left-hand side of (10) is just $\varepsilon^k(G)$.

It is known that there is a principal ideal I of Λ_w (uniquely determined) such that $I\mathscr{A}[[T]] = G(T)\mathscr{A}[[T]]$. Thus, one can find an element $\theta_{\tilde{\sigma}} \in \Lambda_w$ and an element $u(T) \in \mathscr{A}[[T]]^*$ such that $u(T)G(T) = \theta_{\tilde{\sigma}}$. $\theta_{\tilde{\sigma}}$ is well defined up to a factor in Λ_w^*. It satisfies an interpolation property

(11)
$$\phi(\theta_{\tilde{\sigma}}) = c_\phi(\sigma_p \circ \sigma_\infty^{-1}(e_\phi L(0, M_\phi)/\Omega_\phi))$$

for all $\phi = \varepsilon^k \in \mathrm{Spec}(\Lambda_w, \overline{\mathbb{Q}}_p)$, $k \equiv 2 \pmod{p-1}$, $k \geq 2$. To simplify the notation, we put $e_\phi = (1 - \beta^{k-1}p^{-1})$, $M_\phi = M_k$, $\Omega_\phi = \Omega_k$, and $c_\phi = \phi(u(T))\lambda^{1-k} = u((1+p)^k - 1)\lambda^{1-k}$, which must apparently be in $\mathbb{Z}_p^*$, since $L(0, M_\phi)/\Omega_\phi$ is in $\mathbb{Q}$. It would seem impossible to give a direct description of a choice of the c_ϕ's. One can reformulate an old conjecture of Coates and Wiles (the so-called "one-variable main conjecture" stated in **[C-W2]**, which has been proven by Rubin **[R2]**) in terms of $\tilde{\sigma}$. Still assuming, of course, that E has a complex multiplication, let $\widetilde{B} = T_p(\tilde{\sigma}) \otimes_{\Lambda_w} \widehat{\Lambda}_w$. Then the statement is that the characteristic ideal of the cotorsion Λ_w-module $S_{\widetilde{B}}(\mathbb{Q})$ is generated by $\theta_{\tilde{\sigma}}$. We should point out that $E(\mathbb{Q})$ has no p-torsion in this case and so both $H^0(\mathbb{Q}, \widetilde{B})$ and $H^0(\mathbb{Q}, \widetilde{B}^*)$ vanish.

If E does not have complex multiplication, one can still define a p-adic L-function for $\tilde{\sigma}$: $L_p(\phi, \tilde{\sigma}) = \phi(\theta_{\tilde{\sigma}})$ for $\phi \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$, where $\theta_{\tilde{\sigma}} \in \mathscr{K} = R \otimes_{\Lambda_w} \mathscr{L}_w$ and satisfies an interpolation property of the form (11) for all ϕ such that $\phi|_{\Lambda_w} = \varepsilon^k$, $k \in \mathbb{Z}$, $k \geq 2$. It is not hard to verify that $\bigcap \ker(\phi\colon R \to \overline{\mathbb{Q}}_p) = \{0\}$, where the intersection is over all such ϕ's. Also, M_ϕ is the motive over $\mathbb{Q}$ (with coefficients) whose p-adic cohomology (together with $\sigma_p\colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$) gives a representation equivalent over $\overline{\mathbb{Q}}_p$ to $\phi \circ \tilde{\sigma}$, i.e., to $V_p(f_\phi)$. Regarding f_ϕ as having coefficients in $\mathbb{C}$ (vice $\sigma_\infty \circ \sigma_p^{-1}$), we let $L(z, f_\phi)$ be the corresponding L-function. Then $L(z, M_\phi) = L(z + k - 1, f_\phi)$ and so $L(0, M_\phi) = L(k-1, \overline{f}_\phi)$ which is a critical value. Now $T_p(\tilde{\sigma})^+$, the $(+1)$-eigenspace for complex conjugation, is a free rank-one R-module. A choice of generator for $T_p(\tilde{\sigma})^+$ should give a specific $\theta_{\tilde{\sigma}}$, unique up to a factor in R^*. In (10), e_ϕ would be a certain Euler-like factor which one can specify, but one cannot specify precisely the quantities c_ϕ and Ω_ϕ. We remark though that $\phi(T_p(\tilde{\sigma}))$ will be a lattice in $V_p(f_\phi)$ and this lattice should allow one to specify the Deligne period Ω_ϕ up to a factor in the coefficient field of f_ϕ which is mapped to a unit in $\overline{\mathbb{Q}}_p$ by σ_p. The good normalization for the purpose of the following conjecture is not clear. We should also remark that $L_p(\phi, \tilde{\sigma})$ will be the "improved" p-adic L-function in **[G-S]**.

Let S be any discrete, cofinitely generated, cotorsion R-module. Let X be the R-module $\mathrm{Hom}(S, \mathbb{Q}_p/\mathbb{Z}_p)$. We let $\mathscr{R}$ denote the integral closure of R in its fraction field $\mathscr{K}$. Then $X \otimes_R \mathscr{R}$ is a finitely generated, torsion $\mathscr{R}$-module and hence one can define its divisor as in **[B]**. It will be an element in the free abelian group on the prime ideals of $\mathscr{R}$ of height one. We write $\mathrm{Div}(X)$ for the divisor of $X \otimes_R \mathscr{R}$. We define $\mathrm{Div}(S) = \mathrm{Div}(X)$. (This, of course, is a substitute for the notion of characteristic ideal.) If $f \in \mathscr{R}$, $f \neq 0$, we define the divisor of f (denoted by $\mathrm{div}(f)$) as the divisor of the $\mathscr{R}$-module $\mathscr{R}/f\mathscr{R}$. If $fg^{-1} \in \mathscr{K}$, where $f, g \in \mathscr{R}$, and nonzero, we define $\mathrm{div}(fg^{-1}) = \mathrm{div}(f) - \mathrm{div}(g)$. Modulo the question of the good normalization and definition of $\theta_{\tilde{\sigma}}$, here is a natural conjecture.

CONJECTURE 2.2. *Let* $\widetilde{B} = T_p(\tilde{\sigma}) \otimes_R \widehat{R}$. *Then* $S_{\widetilde{B}}(\mathbb{Q})$ *is cotorsion as an* R-*module and* $\mathrm{Div}(S_{\widetilde{B}}(\mathbb{Q})) - \mathrm{Div}(H^0(\mathbb{Q}, \widetilde{B})) - \mathrm{Div}(H^0(\mathbb{Q}, \widetilde{B}^*))$ *is the principal divisor* $\mathrm{div}(\theta_{\tilde{\sigma}})$.

The conjecture predicts $\theta_{\tilde{\sigma}}$, up to a factor in $\mathscr{R}^*$. If $p \geq 11$, then, as remarked above, the H^0-terms are zero. It would follow that $\theta_{\tilde{\sigma}} \in \mathscr{R}$. Possibly $\theta_{\tilde{\sigma}}$ could be chosen in R.

Both $\tilde{\rho}$ and $\tilde{\sigma}$ are specializations of another deformation $\tilde{\tau}$ of $T_p(E)$. Let $\mathfrak{R} = R[[\Gamma]]$. Let $T_p(\tilde{\tau}) = T_p(\tilde{\sigma}) \otimes_R \mathfrak{R}(\tilde{\kappa})$, where $\mathfrak{R}(\tilde{\kappa})$ is the free, rank one $\mathfrak{R}$-module on which $\gamma \in \Gamma \subseteq \mathfrak{R}$ acts by multiplication. $T_p(\tilde{\tau})$ is a free $\mathfrak{R}$-module of rank two. If we choose a basis for $T_p(E)$ over $\mathbb{Z}_p$, then we get one for $T_p(\tilde{\sigma})$ over R, $T_p(\tilde{\rho})$ over Λ, and $T_p(\tilde{\tau})$ over $\mathfrak{R}$. This gives a continuous homomorphism $\tilde{\tau}: \mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_2(\mathfrak{R})$. Let $\Phi: \mathfrak{R} \to R$ be the homomorphism induced by $\gamma \to 1$ for all $\gamma \in \Gamma$. Let $\Psi: \mathfrak{R} \to \Lambda$ be induced from $\psi: R \to \mathbb{Z}_p$, where ψ is as in Hypothesis W. Let $\phi_0 = \psi \circ \Phi: \mathfrak{R} \to \mathbb{Z}_p$. We then have $\tilde{\rho} = \Psi \circ \tilde{\tau}$, $\tilde{\sigma} = \Phi \circ \tilde{\tau}$, and $\rho_0 = \phi_0 \circ \tilde{\tau}$, which gives the Galois action on $T_p(E)$. Let $A(\tilde{\tau}) = T_p(\tilde{\tau}) \otimes_{\mathfrak{R}} \widehat{\mathfrak{R}}$. We define $F^+T_p(\tilde{\tau}) = F^+T_p(\tilde{\sigma}) \otimes_R \mathfrak{R}$ and $F^+A(\tilde{\tau}) = F^+T_p(\tilde{\tau}) \otimes_{\mathfrak{R}} \widehat{\mathfrak{R}}$. We can then define a Selmer group $S_{A(\tilde{\tau})}(\mathbb{Q})$, which will be a cofinitely generated $\mathfrak{R}$-module. $H^0(\mathbb{Q}, A(\tilde{\tau}))$ and $H^0(\mathbb{Q}, A(\tilde{\tau})^*)$, where $A(\tilde{\tau}^*) = \mathrm{Hom}_{\mathbb{Z}_p}(T_p(\tilde{\tau}), \mathbb{Q}_p/\mathbb{Z}_p(1))$, are also cofinitely generated $\mathfrak{R}$-modules, which one can show are pseudo-null.

$\mathfrak{R}$ contains $\mathfrak{L} = \Lambda_w[[\Gamma]] = \mathbb{Z}_p[[\Gamma_w \times \Gamma]]$ and is finitely generated as an $\mathfrak{L}$-module. Let $k, n \in \mathbb{Z}$. Then $\varepsilon^k\kappa^n: \Gamma_w \times \Gamma \to \mathbb{Z}_p$ induces a continuous $\mathbb{Z}_p$-algebra homomorphism $\varepsilon^k\kappa^n: \mathfrak{L} \to \mathbb{Z}_p$. Assume that $\phi \in \mathrm{Spec}(\mathfrak{R}, \overline{\mathbb{Q}}_p)$ is such that $\phi|_{\mathfrak{L}} = \varepsilon^k\kappa^n$, where $k \geq 2$, $n \equiv 0 \pmod{p-1}$, and $1 \leq n+k-1 \leq k-1$. One can verify that $\bigcap_\phi \ker(\phi: \mathfrak{R} \to \overline{\mathbb{Q}}_p) = 0$, the intersection being taken over such ϕ's. Also $\phi \circ \tilde{\tau}$ gives the Galois action on $V_p(f_{\phi_R})(n)$, where $\phi_R = \phi|_R \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$ is such that $\phi_R|_{\Lambda_w} = \varepsilon^k$. The corresponding L-value is $L(n+k-1, \overline{f}_{\phi_R})$, which is a critical value. A p-adic L-function $L_p(\phi, \tilde{\tau})$ interpolating these L-values (with suitable factors) does exist (**[K, G-S]**) and corresponds to an element $\theta_{\tilde{\tau}}$ in the fraction field of $\mathfrak{R}$, although, as before, it is not clear what the right definition for our purpose should be. One can state a conjecture just as above. In the case where E is an elliptic curve with complex multiplication, one can make a more precise definition of the p-adic L-function $L_p(\phi, \tilde{\tau})$ (**[Kz1, Y]**). The conjecture is then equivalent to a special case of the so-called "two-variable main conjecture" of Coates and Yager **[Y]**, which has recently been proven by Rubin **[R3]**.

If E has supersingular reduction at p, then the situation seems to be extremely different. Assuming again that E is modular, a p-adic L-function for E still exists. (See **[M-T-T]**, for example.) It corresponds to a distribution on Γ but not to a $\mathbb{Z}_p$-valued measure. (It is unbounded.) On the algebraic side, Perrin-Riou **[PeR]** constructs a certain distribution on Γ in terms of the behavior of the Selmer groups for E in the tower $\mathbb{Q}_\infty = \bigcup \mathbb{Q}_n$. It is a kind of "characteristic power series" and is well defined up to a factor in Λ^*. Thus one can formulate the analogue of Mazur's conjecture in this case. One can view this as a conjecture about the deformation $\tilde{\rho}$ of $T_p(E)$ considered before. There are other deformations of $T_p(E)$. The whole question of p-adic L-functions and the analogue of Iwasawa's conjecture in this context is completely mysterious. One should possibly also consider deformations where the underlying local $\mathbb{Z}_p$-algebra R is in some larger category.

3. The cyclotomic deformation of a motive

Let M be a motive over $\mathbb{Q}$. For each prime ℓ, let $H_\ell(M)$ denote the ℓ-adic cohomology of M. The L-function attached to M is defined as usual by an Euler product over all primes q,

$$L(z, M) = \prod_q E_q(q^{-z})^{-1},$$

where

$$E_q(T) = \det(I - F_q T : H_\ell(M)^{I_q}).$$

Here F_q is the geometric Frobenius and ℓ is any prime, $\ell \neq q$. We must, of course, assume that $\{H_\ell(M)\}$ is a compatible system of ℓ-adic representations of $G_\mathbb{Q}$. We will assume that $L(z, M)$ can be analytically continued (with a functional equation of the shape conjectured in **[D]**) and also that $L(0, M)$ is a critical value.

Let p be a prime. We will assume that $H_p(M)$ is Hodge-Tate as a representation of $G_{\mathbb{Q}_p}$. That is, $H_p(M) \otimes_{\mathbb{Q}_p} \mathbb{C}_p \cong \bigotimes_i \mathbb{C}_p(i)^{h_i}$ for certain h_i's, $i \in \mathbb{Z}$. Here $\mathbb{C}_p$ is the topological completion of $\overline{\mathbb{Q}}_p$, $G_{\mathbb{Q}_p}$ acts naturally on $\mathbb{C}_p$, and $\mathbb{C}_p(i)$ is the Tate twist. If $d = \dim_{\mathbb{Q}_p}(H_p(M))$, then $\Sigma_i h_i = d$. Let $d^\pm$ denote the dimension over $\mathbb{Q}_p$ of the (± 1)-eigenspaces for complex conjugation acting on $H_p(M)$, so that $d^+ + d^- = d$. We assume that M is of pure weight. The Gamma factors in the conjectural functional equation for $L(z, M)$ can be described in terms of the H_i's and $d^\pm$ (see **[G2,** §6**]**). The assumption that $L(0, M)$ is a critical value is equivalent to

$$\sum_{i \geq 1} h_i = d^+. \tag{12}$$

We shall make the following restrictive assumption on the $G_{\mathbb{Q}_p}$-representation space $H_p(M)$.

PANCHISHKIN CONDITION. *$H_p(M)$ contains a $\mathbb{Q}_p$-subspace W_p invariant under $G_{\mathbb{Q}_p}$ with the property that $W_p \otimes_{\mathbb{Q}_p} \mathbb{C}_p \cong \bigoplus_{i\geq 1} \mathbb{C}_p(i)^{h_i}$.*

It follows that $(H_p(M)/W_p) \otimes_{\mathbb{Q}_p} \mathbb{C}_p \cong \bigotimes_{i \leq 0} \mathbb{C}_p(i)^{h_i}$. If W_p exists, it is characterized by the above property and we shall denote it by $F^+H_p(M)$. By

(12), we have $\dim_{\mathbb{Q}_p}(F^+H_p(M)) = d^+$. The Panchishkin condition holds for Tate motives and also for the motive M such that $H_p(M) = V_p(E)$, where E is an elliptic curve/$\mathbb{Q}$ with ordinary reduction at p. The Panchishkin condition also holds if $H_p(M)$ is ordinary in the sense defined in **[G2, G3]**.

Here is a simple example (pointed out to us by Coates) where the Panchishkin condition holds, but $H_p(M)$ is not ordinary. Let E be an elliptic curve with supersingular reduction at p. Then $V_p(E)$ is not ordinary but is Hodge-Tate with $h_0 = h_1 = 1$. Let $V_p(\Delta)$ be the Deligne representation for Ramanujan's Δ-function. Assume that $p \nmid \tau(p)$ and, therefore, that $V_p(\Delta)$ is ordinary and, hence, Hodge-Tate with $h_0 = h_{11} = 1$. Let $n \in \mathbb{Z}$, $-10 \leq n \leq -1$. Let M be the motive such that $H_p(M) = V_p(E) \otimes_{\mathbb{Q}_p} V_p(\Delta)(n)$. The Panchishkin condition is satisfied, with $W_p = V_p(E) \otimes_{Q_p} (F^+V_p(\Delta))(n)$. Also, $L(0, M)$ is a critical value.

With the above assumptions on the motive M, we can define a Selmer group, essentially just as in **[G2]**. Choose a $G_{\mathbb{Q}}$-invariant lattice $T_p(M)$ in $H_p(M)$. We let $A = H_p(M)/T_p(M)$ and let F^+A denote the image of $F^+H_p(M)$ in A. For every prime ν of $\mathbb{Q}_\infty$, let D_ν and I_ν be the decomposition and inertia groups in $G_{\mathbb{Q}_\infty}$ for some prime of $\overline{\mathbb{Q}}$ over ν. We define

$$L_\nu = \begin{cases} \ker(H^1(D_\nu, A) \to H^1(I_\nu, A)) & \text{if } \nu \nmid p, \\ \ker(H^1(D_\nu, A) \to H^1(I_\nu, A/F^+A)) & \text{if } \nu | p. \end{cases}$$

Then we define $S_A(\mathbb{Q}_\infty)$ as in (4). (We should point out that, in this case, $L_\nu = 0$ for $\nu \nmid p$ since D_ν/I_ν has profinite order prime to p.) $\Gamma = \mathrm{Gal}(\mathbb{Q}_\infty/\mathbb{Q})$ acts in a natural way on $S_A(\mathbb{Q}_\infty)$. Let $\Sigma = \{p, \infty\} \cup \{\text{ramified primes in } H_p(M)\}$, which we assume is a finite set. Let Σ_∞ denote the primes of $\mathbb{Q}_\infty$ lying above those in Σ. If $\nu \notin \Sigma_\infty$, then I_ν acts trivially on $H_p(M)$. It follows that $S_A(\mathbb{Q}_\infty) \subseteq H^1(\mathbb{Q}_\Sigma/\mathbb{Q}_\infty, A)$. We put

$$H_\Sigma = \prod_{\nu \in \Sigma_\infty} (H^1(D_\nu, A)/L_\nu).$$

Then we have

$$S_A(\mathbb{Q}_\infty) = \ker(H^1(\mathbb{Q}_\Sigma/\mathbb{Q}_\infty, A) \to H_\Sigma). \tag{13}$$

Now Γ acts naturally on $H^1(\mathbb{Q}_\Sigma/\mathbb{Q}_\infty, A)$ and on H_Σ, as well as on $S_A(\mathbb{Q}_\infty)$. We can consider them as discrete Λ-modules. They turn out to be cofinitely generated. The results of **[G2, §§3 and 4]** imply that

$$\mathrm{corank}_\Lambda(H^1(\mathbb{Q}_\Sigma/\mathbb{Q}_\infty, A)) \geq d^- \quad \text{and} \quad \mathrm{corank}_\Lambda(H_\Sigma) = \sum_{i \leq 0} h_i. \tag{14}$$

These are consequences of Tate's calculation of global and local Euler-Poincaré characteristics and the facts that $\mathrm{corank}_{\mathbb{Z}_p}(A^-) = d^-$ and $\mathrm{corank}_{\mathbb{Z}_p}(A/F^+A) = \Sigma_{i\leq 0} h_i$. But we have $d^- = \sum_{i \leq 0} h_i$ by (12). Also, it seems extremely likely that the first part of (14) is an equality. (This was conjectured by Schneider.) It then seems reasonable to conjecture that $S_A(\mathbb{Q}_\infty)$ is Λ-cotorsion. Let $A^* = \mathrm{Hom}(T_p(M), \mathbb{Q}_p/\mathbb{Z}_p(1))$. Then $H^0(\mathbb{Q}_\infty, A)$ and $H^0(\mathbb{Q}_\infty, A^*)$ are also discrete Λ-modules. They are obviously Λ-cotorsion.

Let $\phi \in \widehat{\Gamma}$, which we regard as having values in $\overline{\mathbb{Q}}_p$, $\overline{\mathbb{Q}}$, or $\mathbb{C}$ (via σ_p and σ_∞), depending on the context. Let $M_\phi = M \otimes M(\phi)$, where $M(\phi)$ denotes the Artin motive. For $\phi \in \widehat{\Gamma}$, $L(0, M_\phi)$ is always a critical value, since ϕ is even. The Panchishkin condition stated above is equivalent to a condition relating the Newton polygon for M (and p) and the Hodge polygon for M. Under this condition, Panchishkin **[P]** conjectures the existence of a pseudo-measure (i.e., an element θ_M in $\mathscr{L}$) satisfying an interpolation property of the form:

$$\phi(\theta_M) = \sigma_p \circ \sigma_\infty^{-1}(e_\phi L(0, M_\phi)/\Omega_\phi), \tag{15}$$

where e_ϕ is a factor involving the eigenvalues of F_p (on $H_\ell(M)$, $\ell \neq p$) and Ω_ϕ involves a choice of a Deligne period Ω_M for M, the Gaussian sum for ϕ, and values of the infinite Euler factors for $L(z, M)$. A precise (conjectural) description of the interpolation property is given in **[C-P]** and **[C]**. For our purpose, it would be necessary to specify the choice of Deligne period Ω_M, up to a factor in $\mathbb{Q}^*$ which is a p-unit. It should be possible to describe such a choice in terms of the lattice $T_p(M)$ in $H_p(M)$. We could then make the following conjecture.

CONJECTURE 3.1. *$S_A(\mathbb{Q}_\infty)$ is Λ-cotorsion and $\theta_M = f/gh$, where f, g, and $h \in \Lambda$ are generators for the characteristic ideals of $S_A(\mathbb{Q}_\infty)^\iota$, $H^0(\mathbb{Q}_\infty, A)^\iota$, and $H^0(\mathbb{Q}_\infty, A^*)$, respectively.*

Here, if S is a discrete Λ-module, then the definition of the characteristic ideal of S is as given near the end of §1. Also, ι denotes the involution of $\Lambda = \mathbb{Z}_p[[\Gamma]]$ induced by $\gamma \to \gamma^{-1}$ for all $\gamma \in \Gamma$. S^ι denotes the discrete Λ-module which is S as a set but $\lambda \in \Lambda$ acts as $\iota(\lambda)$.

We must explain the relationship between this conjecture and Conjecture 2 of **[G2]** (which is described in a form closer to the above in **[G3]**). In those articles, we consider the L-function $L_V(z)$ corresponding to the compatible system of ℓ-adic representations $V = \{H_\ell(M)\}$, using the *arithmetic* Frobenius in the definition. The p-adic L-function considered there satisfies an interpolation property involving the values $L_V(1, \phi)$, twisting the Dirichlet series by ϕ in the usual way. Let $V' = \{H_\ell(M)'\}$ and $V^* = \{H_\ell(M)^*\}$, where $H_\ell(M)'$ is the contragradient and $H_\ell(M)^* = \mathrm{Hom}(H_\ell(M), \mathbb{Q}_\ell(1)) = H_\ell(M)'(1)$. The functional equation relates $L_V(1, \phi)$ to $L_{V^*}(1, \phi^{-1}) = L_{V'}(0, \phi^{-1}) = L(0, M, \phi^{-1}) = L(0, M_\phi)$. Thus, the p-adic L-function satisfying (14) will coincide with that in **[G2, G3]**, assuming that the factors e_ϕ, Ω_ϕ above and c_ϕ in **[G3]** are chosen suitably. The Selmer groups also coincide and the ι replaces taking the Pontryagin dual as we have in

[G2, G3]. The above conjecture is more general in that we assume the Panchishkin condition on M instead of assuming that $H_p(M)$ is ordinary.

We shall now give a reformulation of the above conjecture. Let $\rho_0\colon G_{\mathbb{Q}} \to \mathrm{Aut}_{\mathbb{Z}_p}(T_p(M))$ give the Galois action on $T_p(M)$. Let $\tilde{\rho}\colon G_{\mathbb{Q}} \to \mathrm{Aut}_\Lambda(\widetilde{T}_p(M))$, where $\widetilde{T}_p(M) = T_p(M) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa})$, be defined by letting $G_{\mathbb{Q}}$ act on $T_p(M)$ by ρ_0 and by $\tilde{\kappa}\colon G_{\mathbb{Q}} \to \Lambda^*$ on $\Lambda(\tilde{\kappa})$. $\widetilde{T}_p(M)$ is a free Λ-module of rank d. Let $A(\tilde{\rho}) = \widetilde{T}_p(M) \otimes_\Lambda \widehat{\Lambda}$. We shall write more briefly $\widetilde{A} = A(\tilde{\rho})$. Define $F^+T_p(M) = T_p(M) \cap (F^+H_p(M))$. We then put

$$F^+\widetilde{T}_p(M) = (F^+T_p(M)) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa}) \quad \text{and} \quad F^+\widetilde{A} = (F^+\widetilde{T}_p(M)) \otimes_\Lambda \widehat{\Lambda}\,.$$

We have $F^+\widetilde{A} \subseteq \widetilde{A}$ and, as Λ-modules, $\widetilde{A} \cong \widehat{\Lambda}^d$, $F^+\widetilde{A} \cong \widehat{\Lambda}^{d^+}$, and $\widetilde{A}/F^+\widetilde{A} \cong \widehat{\Lambda}^{d^-}$. For every prime ν of $\mathbb{Q}$ we define $L_\nu \subseteq H^1(D_\nu, \widetilde{A})$ exactly as previously. (For example, see (8).) We define $S_{\widetilde{A}}(\mathbb{Q})$ as in (4). It is a discrete Λ-module. $S_A(\mathbb{Q}_\infty)$ is a $\mathbb{Z}_p[[\Gamma]]$-module and hence also a Λ-module, identifying Λ with $\mathbb{Z}_p[[\Gamma]]$ as usual. We shall prove the following result.

PROPOSITION 3.2. *As Λ-modules, $S_{\widetilde{A}}(\mathbb{Q})$ is isomorphic to $S_A(\mathbb{Q}_\infty)^\iota$.*

To prove this, consider the restriction map $H^1(\mathbb{Q}, \widetilde{A}) \to H^1(\mathbb{Q}_\infty, \widetilde{A})^\Gamma$. The kernel is $H^1(\Gamma, \widetilde{A}^{G_{\mathbb{Q}_\infty}})$. Let $\widehat{\Lambda}(\tilde{\kappa})$ denote $\widehat{\Lambda}$ where $g \in G_{\mathbb{Q}}$ acts by multiplication by $\tilde{\kappa}(g) \in \Lambda^*$. Then $\widehat{\Lambda}(\tilde{\kappa}) = \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), \mathbb{Q}_p/\mathbb{Z}_p)$, where we use the traditional action of a group on Hom(,). Now $\widetilde{A} = T_p(M) \otimes_{\mathbb{Z}_p} \widehat{\Lambda}(\tilde{\kappa}) = T_p(M) \otimes_{\mathbb{Z}_p} \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), \mathbb{Q}_p/\mathbb{Z}_p) = \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), A)$ since $A = T_p(M) \otimes_{\mathbb{Z}_p} (\mathbb{Q}_p/\mathbb{Z}_p)$. Let $B = A^{G_{\mathbb{Q}_\infty}}$. We have $\widetilde{A}^{G_{\mathbb{Q}_\infty}} = \mathrm{Hom}_{G_\infty}(\Lambda(\tilde{\kappa}^{-1}), A) = \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), B)$. To show that $H^1(\Gamma, \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), B))$ vanishes for any discrete p-primary Γ-module B, it is not difficult to reduce to the case $B = \mathbb{Z}/p\mathbb{Z}$. In that case, the Pontryagin dual of $\mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), B)$ is $\Lambda(\tilde{\kappa}^{-1})/p\Lambda(\tilde{\kappa}^{-1})$. This is simply $\Lambda/p\Lambda = \mathbb{Z}/p\mathbb{Z}[[\Gamma]]$ and $\gamma_0 - 1$ acts by multiplication by $\gamma_0^{-1} - 1$, which is nonzero. This is injective since $\Lambda/p\Lambda$ is an integral domain. Thus, $\mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), B)/(\gamma_0 - 1)\mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), B)$ is in fact zero. Also, $H^2(\Gamma, \widetilde{A}^{G_{\mathbb{Q}_\infty}}) = 0$ since Γ is a free pro-p group. Therefore, we see that $H^1(\mathbb{Q}, \widetilde{A}) \xrightarrow{\sim} H^1(\mathbb{Q}_\infty, \widetilde{A})^\Gamma$ by the restriction map.

Now

$$H^1(\mathbb{Q}_\infty, \widetilde{A}) = H^1(\mathbb{Q}_\infty, \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), A)) = \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), H^1(\mathbb{Q}_\infty, A))$$

as Λ-modules with an action of Γ. As a $\mathbb{Z}_p[[\Gamma]]$-module, $\Lambda(\tilde{\kappa}^{-1}) \cong \Lambda(\tilde{\kappa})^\iota$ and

$$\begin{aligned}\mathrm{Hom}_\Gamma(\Lambda(\tilde{\kappa}^{-1}), H^1(\mathbb{Q}_\infty, A)) &= \mathrm{Hom}_\Gamma(\Lambda(\tilde{\kappa}), H^1(\mathbb{Q}_\infty, A)^\iota) \\ &= \mathrm{Hom}_\Lambda(\Lambda, H^1(\mathbb{Q}_\infty, A)^\iota),\end{aligned}$$

considering $\Lambda(\tilde{\kappa})$ and $H^1(\mathbb{Q}_\infty, A)^\iota$ as Λ-modules. But this can be identified with $H^1(\mathbb{Q}_\infty, A)^\iota$ as a $\mathbb{Z}_p[[\Gamma]]$-module by restricting an element of $\mathrm{Hom}_\Lambda(\Lambda, \ )$ to $1 \in \Lambda$. We have

$$H^1(\mathbb{Q}, \tilde{A}) \cong H^1(\mathbb{Q}_\infty, \tilde{A})^\Gamma \cong H^1(\mathbb{Q}_\infty, A)^\iota$$

as Λ-modules, identifying Λ with $\mathbb{Z}_p[[\Gamma]]$ as before.

Let ℓ be a prime of $\mathbb{Q}$, $\ell \neq p$. Then $I_\ell \subseteq G_{\mathbb{Q}_\infty}$. For any prime ν of $\mathbb{Q}_\infty$ lying over ℓ, I_ν is conjugate in $G_\mathbb{Q}$ to I_ℓ. Let $\sigma \in H^1(\mathbb{Q}, \tilde{A})$ be such that $\sigma|_{I_\ell}$ is trivial. Then $\sigma|_{I_\nu}$ is trivial for all $\nu|\ell$. Tracing the above isomorphisms, one sees that σ is mapped to $\sigma' \in H^1(\mathbb{Q}_\infty, A)^\iota$ and $\sigma'|_{I_\nu}$ is trivial. Conversely, if $\sigma'|_{I_\nu}$ is trivial, then $\sigma|_{I_\ell}$ is trivial. If $\ell = p$, let π be the unique prime of $\mathbb{Q}_\infty$ over p. Then $I_p/I_\pi \cong \Gamma$ canonically. As above, $\sigma \to \sigma'$ maps $H^1(I_p, \tilde{A}/F^+\tilde{A})$ isomorphically to $H^1(I_\pi, A/F^+A)^\iota$. Therefore, $\sigma|_{D_p} \in L_p$ if and only if $\sigma'|_{D_\pi} \in L_\pi$. The proposition follows from these remarks.

We saw above that $H^0(\mathbb{Q}_\infty, \tilde{A}) = \tilde{A}^{G_{\mathbb{Q}_\infty}} = \mathrm{Hom}(\Lambda(\tilde{\kappa}^{-1}), B)$, where $B = A^{G_{\mathbb{Q}_\infty}}$. Hence, $H^0(\mathbb{Q}, \tilde{A}) = \mathrm{Hom}_\Gamma(\Lambda(\tilde{\kappa}^{-1}), B)$ which is isomorphic to $B^\iota = H^0(\mathbb{Q}_\infty, A)^\iota$ as a Λ-module. Let $\tilde{A}^* = \mathrm{Hom}_{\mathbb{Z}_p}(\tilde{T}_p(M), \mathbb{Q}_p/\mathbb{Z}_p(1)) = \mathrm{Hom}_{\mathbb{Z}_p}(T_p(M) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa}), \mathbb{Q}_p/\mathbb{Z}_p(1)) = \mathrm{Hom}_{\mathbb{Z}_p}(\Lambda(\tilde{\kappa}), A^*)$. By the same argument, we see that $H^0(\mathbb{Q}, \tilde{A}^*)$ is isomorphic to $(A^*)^{G_{\mathbb{Q}_\infty}} = H^0(\mathbb{Q}_\infty, A^*)$. We obtain the following reformulation of Conjecture 3.1.

CONJECTURE 3.3. *$S_{\tilde{A}}(\mathbb{Q})$ is Λ-cotorsion. $\theta_M = f/gh$, where f, g, and $h \in \Lambda$ generate the characteristic ideals of the Λ-modules $S_{\tilde{A}}(\mathbb{Q})$, $H^0(\mathbb{Q}, \tilde{A})$, and $H^0(\mathbb{Q}, \tilde{A}^*)$, respectively.*

We also want to point out that the interpolation property (15) characterizing θ_M can be stated in terms of the deformation $\tilde{\rho}$. If $\phi \in \hat{\Gamma}$, then ϕ induces an element $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$. If $\phi = \phi_0$ (= the trivial character), then $\phi_0 \circ \tilde{\rho} = \rho_0$, which gives the Galois action on $T_p(M)$ and, hence, on $H_p(M) = T_p(M) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$. In general, $\phi \circ \tilde{\rho}\colon G_\mathbb{Q} \to \mathrm{GL}_d(\overline{\mathbb{Q}}_p)$ gives the Galois action on $H_p(M) \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}}_p(\phi)$, where $\overline{\mathbb{Q}}_p(\phi)$ is the one-dimensional space on which $G_\mathbb{Q}$ acts by ϕ. But this is the $\overline{\mathbb{Q}}_p$-representation obtained from $H_p(M_\phi)$ together with the embedding $\sigma_p\colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$.

The reformulation of Mazur's conjecture discussed in §2 is the special case $A = E_{p^\infty}$. We must point out though that $S_A(\mathbb{Q}_\infty) \cong S_A(\mathbb{Q}_\infty)^\iota$ as Λ-modules. This is a consequence of the Weil pairing. (See **[M1]** or **[G2]**.)

There are two important cases where the conjecture has been proven. Let M be of the form $\mathbb{Q}(n) \otimes M(\xi)$, where ξ is an irreducible Artin character over $\mathbb{Q}$, $M(\xi)$ is the corresponding Artin motive, and $n \in \mathbb{Z}$. Then

$L(0, M)$ is a critical value only when ξ factors through either a totally real extension of $\mathbb{Q}$ or a totally complex quadratic extension of such a field and, in either case, for suitable $n \in \mathbb{Z}$. In this case, the conjecture is a consequence of a theorem of Wiles **[W2]**. If M is the motive arising from an elliptic curve $E/\mathbb{Q}$, the conjecture is equivalent to Mazur's conjecture, as explained in §2. If E has complex multiplication, it has been proven by Rubin **[R3]**.

4. A general conjecture

Let M be a motive over $\mathbb{Q}$. Assume that $L(0, M)$ is a critical value. Let $T_p(M)$ be a $G_{\mathbb{Q}}$-invariant lattice in the $\mathbb{Q}_p$-representation space $H_p(M)$. We shall assume that the Panchishkin condition holds for M. Let $F^+T_p(M) = T_p(M) \cap F^+H_p(M)$, as before. Then $F^+T_p(M)$ and $T_p(M)/F^+T_p(M)$ are free $\mathbb{Z}_p$-modules of rank d^+ and d^-, respectively. Choose a basis for $T_p(M)$ over $\mathbb{Z}_p$ and let $\rho_0: G_{\mathbb{Q}} \to \mathrm{GL}_d(\mathbb{Z}_p)$ give the action of $G_{\mathbb{Q}}$ on $T_p(M)$. Let Σ be a finite set of primes of $\mathbb{Q}$ containing p, ∞, and all primes ramified in the field corresponding to $\ker(\rho_0)$. Let $\tilde{\rho}$: $\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_d(R)$ be a deformation of ρ_0. We assume that R is a local, Noetherian, commutative $\mathbb{Z}_p$-algebra with maximal ideal $\mathfrak{m}$ that is compact in its $\mathfrak{m}$-adic topology, and that there is a continuous $\mathbb{Z}_p$-algebra homomorphism $\phi_0: R \to \mathbb{Z}_p$ such that $\phi_0 \circ \tilde{\rho} = \rho_0$. Let $P_0 = \ker(\phi_0)$. If $T_p(\tilde{\rho})$ denotes the free R-module of rank d on which $\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q})$ acts via $\tilde{\rho}$, then the "reduction modulo P_0" $T_p(\tilde{\rho})/P_0T_p(\tilde{\rho})$ is isomorphic to $T_p(M)$.

Let $\phi \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$. Then $\phi(R)$ is a subring of the integers in some finite extension K_ϕ of $\mathbb{Q}_p$. We shall call ϕ a "critical specialization for $\tilde{\rho}$" if there is a motive M_ϕ over $\mathbb{Q}$ (with coefficient field $E_\phi \subseteq \overline{\mathbb{Q}}$) such that $L(0, M_\phi)$ is a critical value, σ_p maps E_ϕ into K_ϕ, and $\phi \circ \tilde{\rho}$: $\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_d(K_\phi)$ gives the Galois action on $H_p^\phi = H_p(M_\phi) \otimes_{E_\phi} K_\phi$. In the tensor product, K_ϕ is regarded as an E_ϕ-module via σ_p. If M_ϕ satisfies the Panchishkin condition, we define $F^+H_p^\phi = F^+H_p(M_\phi) \otimes_{E_\phi} K_\phi$, which is a K_ϕ-vector space of dimension d^+ since $L(0, M_\phi)$ is assumed to be critical. (One can show that the $(+1)$-eigenspace in H_p^ϕ for complex conjugation has dimension d^+ for all ϕ.) The subspace $F^+H_p^\phi$ of H_p^ϕ is $G_{\mathbb{Q}_p}$-invariant and is characterized by the property: $F^+H_p^\phi$ is of positive Hodge-Tate type, $H_p^\phi/F^+H_p^\phi$ is of nonpositive Hodge-Tate type.

Let $\mathrm{Crit\,Spec}_{\tilde{\rho}}(R, \overline{\mathbb{Q}}_p)$ denote the set of critical specializations for $\tilde{\rho}$. A subset C of $\mathrm{Crit\,Spec}_{\tilde{\rho}}(R, \mathbb{Q}_p)$ is said to be "ample" if $\bigcap_{\phi \in C} \ker(\phi) = 0$. In the cases discussed in the earlier sections, ample sets exist, although in general this is not true. We shall need to assume the following restrictive hypothesis on $\tilde{\rho}$.

PANCHISHKIN CONDITION FOR $\tilde{\rho}$. *There exist a $G_{\mathbb{Q}_p}$-invariant R-submodule $F^+T_p(\tilde{\rho})$ of $T_p(\tilde{\rho})$ and a subset C of* Crit Spec$_{\tilde{\rho}}(R, \overline{\mathbb{Q}}_p)$ *with the following properties.*

(1) *$F^+T_p(\tilde{\rho})$ and $T_p(\tilde{\rho})/F^+T_p(\tilde{\rho})$ are free R-modules of ranks d^+ and d^-, respectively.*
(2) *For $\phi \in C$, M_ϕ satisfies the Panchishkin condition and $\phi(F^+T_p(\tilde{\rho})) \subseteq F^+H_p^\phi$.*
(3) *C is ample.*
(4) *$\phi_0 \in C$.*

For the deformations considered in the preceding sections, the Panchishkin condition is satisfied. Also, $F^+T_p(\tilde{\rho})$, if it exists, is uniquely determined by the ample set C. We call $(F^+T_p(\tilde{\rho}), C)$ a Panchishkin type for $\tilde{\rho}$ if (1), (2), and (3) are satisfied.

Given a Panchishkin type for $\tilde{\rho}$, one can easily define a Selmer group. Let $A(\tilde{\rho}) = T_p(\tilde{\rho}) \otimes_R \widehat{R}$, where $\widehat{R}$ is the discrete R-module $\mathrm{Hom}(R, \mathbb{Q}_p/\mathbb{Z}_p)$. Then $A(\tilde{\rho}) \cong \widehat{R}^d$ as an R-module. $G_{\mathbb{Q}}$ acts on $A(\tilde{\rho})$ through its action on $T_p(\tilde{\rho})$. Let $F^+A(\tilde{\rho}) = (F^+T_p(\tilde{\rho})) \otimes_R \widehat{R}$, which is a $G_{\mathbb{Q}_p}$-invariant R-submodule of $A(\tilde{\rho})$. For a prime ν of $\mathbb{Q}$, we let $L_\nu = H^1_{\mathrm{unr}}(D_\nu, A(\tilde{\rho}))$ if $\nu \neq p$ and $L_\nu = \ker(H^1(D_\nu, A(\tilde{\rho})) \to H^1(I_\nu, A(\tilde{\rho})/F^+A(\tilde{\rho}))$ if $\nu = p$. We then define $S_{A(\tilde{\rho})}(\mathbb{Q})$ exactly as in (4) of section 1, where we take $F = \mathbb{Q}$ and $A = A(\tilde{\rho})$. It is an R-module.

PROPOSITION. *$S_{A(\tilde{\rho})}(\mathbb{Q})$ is cofinitely generated as an R-module.*

PROOF. Since I_ν acts trivially on $A(\tilde{\rho})$ if $\nu \notin \Sigma$, one sees that $S_{A(\tilde{\rho})}(\mathbb{Q}) \subseteq H^1(\mathbb{Q}_\Sigma/\mathbb{Q}, A(\tilde{\rho}))$. R-submodules and quotient modules of a cofinitely generated R-module will also be cofinitely generated. Let $\widetilde{A} = A(\tilde{\rho})$. We shall show that $H^1(\mathbb{Q}_\Sigma/\mathbb{Q}, \widetilde{A})$ is R-cofinitely generated, just assuming that $\widetilde{A}$ is. Let $\mathfrak{m}$ denote the maximal ideal of R. Then we have $R/\mathfrak{m} \cong \mathbb{Z}/p\mathbb{Z}$ and $\mathfrak{m} = (r_1, r_2, \dots, r_t)$ say, for some $t \geq 1$ and $r_1, r_2, \dots, r_t \neq 0$. If $r \in \mathfrak{m}$, $r \neq 0$, then we have an exact sequence

$$0 \to \widetilde{A}[r] \to \widetilde{A} \xrightarrow{r} \widetilde{B} \to 0,$$

where $\widetilde{B} = r\widetilde{A}$ and $\widetilde{A}[r]$ is the kernel of multiplication by r. $\widetilde{A}[r]$ and $\widetilde{A}/\widetilde{B}$ are cofinitely generated modules over $R/(r)$. We have the following exact sequences, where we omit the $\mathbb{Q}_\Sigma/\mathbb{Q}$ in the notation:

$$H^1(\widetilde{A}[r]) \to H^1(\widetilde{A}) \to H^1(\widetilde{B}),$$
$$H^0(\widetilde{A}/\widetilde{B}) \to H^1(\widetilde{B}) \to H^1(\widetilde{A}).$$

If $H^1(\widetilde{A}[r])$ and $H^0(\widetilde{A}/\widetilde{B})$ are cofinitely generated over $R/(r)$, then these exact sequences imply that $H^1(\widetilde{A})[r]$ is also. This implies that $H^1(\widetilde{A})[\mathfrak{m}]$ is finite and, as we explain below, that $H^1(\widetilde{A})$ is cofinitely generated over

R. Let $r = r_t$. Then $R/(r)$ has maximal ideal $\mathfrak{m}/(r)$, generated by $t-1$ elements. Thus, we reduce to the case where $\mathfrak{m} = 0$ and $\widetilde{A}$ is a cofinitely generated $\mathbb{Z}/p\mathbb{Z}$-module, i.e., $\widetilde{A}$ is finite. In this case, it is well known that $H^1(\mathbb{Q}_\Sigma/\mathbb{Q}, \widetilde{A})$ is finite.

Thus, $H^1(\mathbb{Q}_\Sigma/\mathbb{Q}, \widetilde{A})[\mathfrak{m}]$ is finite. Let $X = \widehat{H}^1(\mathbb{Q}_\Sigma/\mathbb{Q}, \widetilde{A})$. X is a compact R-module and $X/\mathfrak{m}X$ is finite. Suppose $x_1, x_2, \dots, x_t \in X$ generate $X/\mathfrak{m}X$. Let $Y = Rx_1 + \cdots + Rx_t$, which is a closed R-submodule of X. Let $Z = X/Y$. Then Z is compact and $\mathfrak{m}Z = Z$. A simple compactness argument shows that $Z = 0$. Hence, $H^1(\mathbb{Q}_\Sigma/\mathbb{Q}, \widetilde{A})$ is cofinitely generated over R.

Let $\chi: G_\mathbb{Q} \to \mathbb{Z}_p^* \subset R^*$ be the cyclotomic character. Let $\tilde{\rho}'$ denote the contragradient of $\tilde{\rho}$: $\tilde{\rho}' = \lambda \circ \tilde{\rho}$, where $\lambda(\alpha) = (\text{Transpose}(\alpha))^{-1}$ for $\alpha \in \mathrm{GL}_d(R)$. Let $\tilde{\rho}^* = \chi \otimes \tilde{\rho}'$. Then $\phi_0 \circ \tilde{\rho}^* = \chi \otimes \rho_0'$. This gives the Galois action on $T_p(M)^* = \mathrm{Hom}_{\mathbb{Z}_p}(T_p(M), \mathbb{Z}_p(1))$, which is a lattice in $H_p(\check{M}(1))$, where $\check{M}$ is the dual motive to M. We define $T_p(\tilde{\rho}^*)$ as before and let $A(\tilde{\rho}^*) = T_p(\tilde{\rho}^*) \otimes_R \widehat{R} = \mathrm{Hom}_{\mathbb{Z}_p}(T_p(\tilde{\rho}), \mathbb{Q}_p/\mathbb{Z}_p(1))$. Now $H^0(\mathbb{Q}, H_p(M)) = 0$ since otherwise $H_p(M)$ would contain $\mathbb{Q}_p(0)$. The Gamma factors which occur in the functional equation for $L(z, M)$ would then include that for $L(z, \mathbb{Q}(0))$, i.e., $\Gamma(z/2)$, which has a pole at $z = 0$. (This depends on the conjectural description of the Gamma factors given in terms of the p-adic representation $H_p(M)$ in **[G2]**, section 6.) But this is not possible since $L(0, M)$ is assumed to be critical. Similarly, $H^0(\mathbb{Q}, H_p(\check{M}(1)) = 0$ since $L(0, \check{M}(1)) = L(1, \check{M})$ is the L-value on the "other side" of the functional equation, and hence is also critical. $H^0(\mathbb{Q}, A(\tilde{\rho}))$ and $H^0(\mathbb{Q}, A(\tilde{\rho}^*))$ are clearly cofinitely generated as R-modules. We see that they are also R-cotorsion because otherwise either $H^0(\mathbb{Q}, A(\tilde{\rho}))[P_0] = H^0(\mathbb{Q}, A(\tilde{\rho})[P_0]) = H^0(\mathbb{Q}, T_p(M) \otimes (\mathbb{Q}_p/\mathbb{Z}_p))$ or the corresponding group for $\check{M}(1)$ would be infinite. The above remarks show that this is not the case.

We shall now assume that R is an integral domain, although this may not be absolutely essential for what follows. (For example, one could assume that the total ring of fractions of R is a direct sum of fields, slightly modifying the following discussion.) Let $\mathscr{K}$ be the fraction field for R, and let $\mathscr{R}$ denote the integral closure of R in $\mathscr{K}$. A theorem of Nagata states that $\mathscr{R}$ is finitely generated as an R-module and, hence, is also a compact, Noetherian local $\mathbb{Z}_p$-algebra. (See **[Dd]**.) Assume that we have a Panchishkin type $F^+T_p(\tilde{\rho})$, C. Although we cannot give a precise interpolation property that would characterize a p-adic L-function, we expect that it should exist and satisfy

(1) $L_p(\phi, \tilde{\rho}) = c_\phi(\sigma_p \circ \sigma_\infty^{-1}(e_\phi L(0, M_\phi)/\Omega_\phi))$ for all $\phi \in C$.

(2) There exists an element $\theta_{\tilde{\rho}} = rs^{-1}$ in $\mathscr{K}$ (where $r, s \in R$, $s \neq 0$) such that $L_p(\phi, \tilde{\rho}) = \phi(\theta_{\tilde{\rho}})$ for all $\phi \in \mathrm{Spec}(R, \overline{\mathbb{Q}}_p)$ with $\phi(s) \neq 0$.

Here $c_\phi \in \overline{\mathbb{Q}}_p^*$ (some kind of "p-adic period"), $e_\phi = a$ factor involving the Euler factors at p and the eigenvalues of F_p on $H_\ell(M_\phi)$, $\ell \neq p$, and $\Omega_\phi \in \mathbb{C}^*$ is a Deligne period for M_ϕ.

If X is a finitely generated, torsion R-module, we define $\mathrm{Div}(X)$ to be the divisor of the $\mathscr{R}$-module $X \otimes_R \mathscr{R}$. If S is a cofinitely generated, cotorsion R-module, we define $\mathrm{Div}(S)$ to be $\mathrm{Div}(X)$, where $X = \widehat{S}$. If $\theta \in \mathscr{K}$, $\theta \neq 0$, then we write $\theta = rs^{-1}$, with $r, s \in R$, not zero, and define $\mathrm{div}(\theta) = \mathrm{div}(r) - \mathrm{div}(s)$, where $\mathrm{div}(f) = \mathrm{Div}(R/(f))$ for all nonzero f in R.

We can now state the general conjecture, which we view with some tentativeness because the evidence is limited to just a few cases in which R is a rather well-behaved $\mathbb{Z}_p$-algebra.

Conjecture 4.1. *If $\theta_{\tilde{\rho}} = 0$, then $S_{A(\tilde{\rho})}(\mathbb{Q})$ is not R-cotorsion. If $\theta_{\tilde{\rho}} \neq 0$, then $S_{A(\tilde{\rho})}(\mathbb{Q})$ is R-cotorsion and*

$$\mathrm{div}(\theta_{\tilde{\rho}}) = \mathrm{Div}(S_{A(\tilde{\rho})}(\mathbb{Q})) - \mathrm{Div}(H^0(\mathbb{Q}, A(\tilde{\rho}))) - \mathrm{Div}(H^0(\mathbb{Q}, A(\tilde{\rho}^*))).$$

We intend to discuss various aspects of this conjecture elsewhere, including hopefully a proof that the right-hand element in the divisor group of $\mathscr{R}$ is in fact principal, and also that the conjecture is compatible with the functional equation. We should add that in the above formulation we have chosen one of several possible substitutes for the notion of characteristic ideal. It may not be the best choice.

The equality $\theta_{\tilde{\rho}} = 0$ does occur. For example, consider the deformation $\tilde{\tau}$ described in §2 for the case where E has complex multiplication. Then $\theta_{\tilde{\tau}} \in \mathfrak{R} = \mathbb{Z}_p[[\Gamma_w \times \Gamma]]$, which is isomorphic to a formal power series ring over $\mathbb{Z}_p$ in two variables. Assume that the Hasse-Weil L-function $L(z, E)$ has an odd order zero at $z = 1$. Then $\theta_{\tilde{\tau}}$ is divisible by a certain irreducible element Θ of $\mathfrak{R}$, the so-called "critical divisor" described in **[G1]**. We have $\phi_0(\Theta) = 0$ and so $\Theta \in \ker(\phi_0)$. Let $R = \mathfrak{R}/(\Theta)$ and let $\lambda \colon \mathfrak{R} \to R$ be the canonical homomorphism. Then $\lambda \circ \tilde{\rho}$ is a deformation of ρ_0 satisfying the Panchishkin condition and such that $\theta_{\lambda\circ\tilde{\rho}} = \theta_{\tilde{\rho}}|_{\mathrm{Spec}(R,\overline{\mathbb{Q}}_p)}$ vanishes identically. In **[G1]**, we prove a result equivalent to the assertion that $S_{A(\lambda\circ\tilde{\rho})}$ is not R-cotorsion.

Another interesting example is the following. Let A be an abelian variety of dimension g defined over $\mathbb{Q}$ which has complex multiplication (by a CM field K, say, with $[K : \mathbb{Q}] = 2g$). Assume that A has good, ordinary reduction at p. Let $\overline{A}$ denote the reduction. Let $\rho_0 \colon G_{\mathbb{Q}} \to \mathrm{GL}_{2g}(\mathbb{Z}_p)$ give the action of $G_{\mathbb{Q}}$ on $T_p(A)$. Let M be a motive over $\mathbb{Q}$ such that $H_p(M) = V_p(A) = T_p(A) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$. The Panchishkin condition for M is easily verified: $F^+H_p(M)$ is the kernel of the reduction map $V_p(A) \to V_p(\overline{A})$. Let $\widetilde{K}_\infty$ denote the composite of all $\mathbb{Z}_p$-extensions of K. Then $G_\infty = \mathrm{Gal}(\widetilde{K}_\infty/K_\infty) \cong \mathbb{Z}_p^{g+1}$, assuming that Leopoldt's conjecture is valid for K. Let $R = \mathbb{Z}_p[[G_\infty]]$, which is isomorphic to the formal power series ring over $\mathbb{Z}_p$ in $g+1$ variables. One can describe rather explicitly a deformation $\tilde{\rho}: G_{\mathbb{Q}} \to \mathrm{GL}_{g+1}(R)$ of ρ_0. The Panchishkin condition for $\tilde{\rho}_0$ holds. $F^+T_p(\tilde{\rho})$ can be described in a way that involves the CM type of A.

For C we can take the set of ϕ's in $\mathrm{Crit\,Spec}_{\tilde{\rho}}(R, \overline{\mathbb{Q}}_p)$ satisfying $\phi(F^+T_p(\tilde{\rho})) \subseteq F^+H_p(M_\phi)$. C is ample. One interesting feature in this example is that C will be a proper subset of $\mathrm{Crit\,Spec}_{\tilde{\rho}}(R, \overline{\mathbb{Q}}_p)$, which depends again on the CM type of A. A p-adic L-function satisfying (1) and (2) above (with $\theta_{\tilde{\rho}}$ actually in R) can be obtained from Katz's $(g+1)$-variable p-adic L-function corresponding to K and the CM type of A **[Kz2]**. For $g > 1$, the most significant progress on Conjecture 4.1 is the work of Hida and Tilouine (**[H-T]** and subsequent papers). The p-adic deformation point of view plays an important role in their work.

Suppose now that M is a motive over $\mathbb{Q}$ such that $L(0, M)$ is *not* a critical value. It is sometimes still possible to formulate a conjecture analogous to Conjecture 4.1. We shall assume that $H_p(M)$ contains a $G_{\mathbb{Q}_p}$-invariant subspace W_p satisfying the condition

$$\dim_{\mathbb{Q}_p}(W_p) = \dim_{Q_p}(H_p(M)^+). \tag{16}$$

Choose a $G_{\mathbb{Q}}$-invariant lattice $T_p(M)$ in $H_p(M)$. We simply *define* $F^+T_p(M) = T_p(M) \cap W_p$. One can then consider the cyclotomic deformation $\tilde{\rho}$ described in §3, $\tilde{\rho} = \rho_0 \otimes \tilde{\kappa}$. We have $T_p(\tilde{\rho}) = T_p(M) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa})$, $F^+T_p(\tilde{\rho}) = F^+T_p(M) \otimes_{\mathbb{Z}_p} \Lambda(\tilde{\kappa})$, and part (1) of the Panchishkin condition for $\tilde{\rho}$ is satisfied. One can then define $A(\tilde{\rho})$ and a Selmer group $S_{A(\tilde{\rho})}(\mathbb{Q})$, which the remarks of §2 suggest may conceivably be Λ-cotorsion. (To avoid trivial counterexamples, assume that $H_p(M)$ is irreducible for the action of $G_{\mathbb{Q}}$.) To define a p-adic L-function $L_p(\phi, \tilde{\rho})$ we might proceed as follows. We shall make the following additional restrictive assumption on M and p. Let $\overline{\rho}_0: G_{\mathbb{Q}} \to \mathrm{GL}_d(\mathbb{Z}/p\mathbb{Z})$ be the reduction modulo p of ρ_0, which gives the action of $G_{\mathbb{Q}}$ on $T_p(M)/pT_p(M)$. Obviously $F^+T_p(M)/pF^+T_p(M)$ is a $G_{\mathbb{Q}_p}$-invariant $\mathbb{Z}/p\mathbb{Z}$-subspace. Let $\overline{\alpha}: G_{\mathbb{Q}_p} \to \mathrm{GL}_{d^+}(\mathbb{Z}/p\mathbb{Z})$ give the action on this subspace and $\overline{\beta}$ give the action on the corresponding quotient space. We assume that

(1) $\overline{\rho}_0$ is absolutely irreducible.
(2) The irreducible representations of $G_{\mathbb{Q}_p}$ over $\mathbb{Z}/p\mathbb{Z}$ occurring in $\overline{\alpha}$ are distinct from those occurring in $\overline{\beta}$.

The usefulness of assumption (2) was pointed out to us by Tilouine. Under these assumptions one can construct a "universal deformation" of the pair ρ_0, $F^+T_p(M)$, which might conceivably satisfy the Panchishkin condition.

Let Σ be as before. Let $S = R(\overline{\rho}_0)$ denote the universal deformation ring attached to $\mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q})$ and $\overline{\rho}_0$, whose existence is proved in **[M2]** under assumption (1). Let $\tilde{\sigma}: \mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_d(S)$ be the corresponding representation. Then $\phi_0 \circ \tilde{\sigma} = \rho_0$ for some $\phi_0 \colon S \to \mathbb{Z}_p$. If $P_0 = \ker(\phi_0)$, then $T_p(\tilde{\sigma})/P_0T_p(\tilde{\sigma}) \cong T_p(M)$ contains a $G_{\mathbb{Q}_p}$-invariant $\mathbb{Z}_p$-submodule $F^+T_p(M)$ which by assumption (2) can be characterized as the maximal $G_{\mathbb{Q}_p}$-invariant $\mathbb{Z}_p$-submodule U with the property $\mathrm{Irred}(U/pU) \subseteq \overline{\alpha}$. By this we mean that the irreducible representations of $G_{\mathbb{Q}_p}$ over $\mathbb{Z}/p\mathbb{Z}$ occurring in U/pU also occur in $\overline{\alpha}$ (and hence not in $\overline{\beta}$). Let $\phi \in \mathrm{Spec}(S, \overline{\mathbb{Q}}_p)$ and let $P_\phi = \ker(\phi)$. We say that ϕ is "admissible" if in $T_p(\tilde{\sigma})/P_\phi T_p(\tilde{\sigma})$ (which we denote by T_ϕ) the maximal $G_{\mathbb{Q}_p}$-invariant $S/P_\phi S$-submodule U_ϕ such that $\mathrm{Irred}(U_\phi/pU_\phi) \subset \overline{\alpha}$ has rank d^+ over $S_\phi = S/P_\phi S$. One sees then that U_ϕ and T_ϕ/U_ϕ are free S_ϕ-modules of rank d^+ and d^-, respectively. Let $R = S/I$ with $I = \bigcap_\phi P_\phi$, where the intersection is over all admissible ϕ's in $\mathrm{Spec}(S, \overline{\mathbb{Q}}_p)$. Let $\tilde{\tau}\colon \mathrm{Gal}(\mathbb{Q}_\Sigma/\mathbb{Q}) \to \mathrm{GL}_d(R)$ be the induced representation and $T_p(\tilde{\tau}) = T_p(\tilde{\sigma})/IT_p(\tilde{\sigma})$ be the underlying Galois module. Since ϕ_0 is admissible, it factors through R and we have $\phi_0 \circ \tilde{\tau} = \rho_0$, i.e., $\tilde{\tau}$ is a deformation of ρ_0. Let $\mathfrak{m}$ be the maximal ideal of R. Let U be the maximal R-submodule of $T_p(\tilde{\tau})$ invariant under $G_{\mathbb{Q}_p}$ and such that $\mathrm{Irred}(U/\mathfrak{m}U) \subseteq \overline{\alpha}$. Then one can prove that $\phi(U) = U_\phi$ for every admissible ϕ and that U and $T_p(\tilde{\tau})/U$ are free R-modules of rank d^+ and d^-, respectively. We denote U by $F^+T_p(\tilde{\tau})$. Its image under ϕ_0 is $F^+T_p(M)$. Also, $T_p(\tilde{\tau})$ is a deformation of $T_p(\tilde{\rho})$ corresponding to a homomorphism $R \to \Lambda$. The image of $F^+T_p(\tilde{\tau})$ is $F^+T_p(\tilde{\rho})$. R might not be an integral domain. If not, one could replace it by R/P, where P is a minimal prime ideal of R contained in $\ker(R \to \Lambda)$.

Let C be the set of all $\phi \in \mathrm{Crit\,Spec}_{\tilde{\tau}}(R, \overline{\mathbb{Q}}_p)$ such that M_ϕ satisfies the Panchishkin condition and $\phi(F^+T_p(\tilde{\tau})) \subseteq F^+H_p^\phi$. We should point out that $\tilde{\tau}$, $F^+T_p(\tilde{\tau})$, and C depend on our choice of $G_{\mathbb{Q}_p}$-invariant subspace W_p of $H_p(M)$. The lattice $T_p(M)$ is unique, up to homothety, because of assumption (1). It is tempting to hope that C is ample and hence that the Panchishkin condition for $\tilde{\tau}$ holds, except for property (4) of course. If so, then let $L_p(\phi, \tilde{\rho})$ be the restriction of $L_p(\phi, \tilde{\tau})$ to $\mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$, noticing that all $\phi \in \mathrm{Spec}(\Lambda, \overline{\mathbb{Q}}_p)$ are admissible. This should correspond to an element $\theta_{\tilde{\rho}}$ in $\mathscr{L}$. The Conjecture 4.1 would make sense.

As an example, let $M = M(\xi)$, where ξ is an Artin character over $\mathbb{Q}$. Assume that both d^+ and d^- are positive. Then $\mathrm{Crit\,Spec}_{\tilde{\rho}}(\Lambda, \overline{\mathbb{Q}}_p)$ is empty. We shall assume that $\overline{\rho}_0 = \overline{\xi}$ is absolutely irreducible. But it happens quite often that $\xi|_{G_{\mathbb{Q}_p}}$ is reducible. One possibility might be that $H_p(M)$ is a direct sum of distinct one-dimensional $G_{\mathbb{Q}_p}$-invariant $\mathbb{Q}_p$-subspaces corresponding to characters $\eta_1, \eta_2, \ldots, \eta_d$ of $G_{\mathbb{Q}_p}$. Assuming that p is odd, these characters will also be distinct $\mathrm{mod}\, p\mathbb{Z}_p$. One can choose W_p in $\binom{d}{d^+}$

ways, so that it satisfies (16). Assumptions (1) and (2) will also hold. It then seems rather likely that for each such choice of W_p, one can define as above a corresponding p-adic L-function and hence a "main conjecture" for the cyclotomic deformation of ξ.

Here is one concrete case. Let ξ be the two-dimensional Artin character corresponding to the S_3-extension $L/\mathbb{Q}$, where L is the Hilbert class field of $K = \mathbb{Q}(\sqrt{-23})$. Let $p = 23$. Then $\xi|_{G_{\mathbb{Q}_p}} = \eta_1 + \eta_2$, where η_1 is the trivial character and $\eta_2 = \omega^{11}$. Let $\tilde{\sigma}$ denote the p-adic deformation denoted by $\rho_{p,\Delta}$ in the introduction to **[M-W2]**. Then $\tilde{\sigma}$ specializes in weight one to ξ. The Panchishkin condition for $\tilde{\sigma}$ is valid, except for (4). The image of $F^+T_p(\tilde{\sigma})$ in $H_p(M)$ spans $H_p(M)^{\eta_2}$. However, one could twist $\tilde{\sigma}$ by the quadratic Artin character ψ of conductor 23, obtaining another p-adic deformation $\tilde{\sigma} \otimes \psi$ of $\xi \otimes \psi = \xi$. This time, the image of $F^+T_p(\tilde{\sigma} \otimes \phi)$ in $H_p(M)$ spans $H_p(M)^{\eta_1}$, the other possible choice of W_p. If we let $\tilde{\tau}$ be either $\tilde{\sigma} \otimes \tilde{\kappa}$ or $(\tilde{\sigma} \otimes \psi) \otimes \tilde{\kappa}$, then the set C described above will in fact be ample. It is not clear, however, that $\tilde{\tau}$ is "universal".

References

[B] N. Bourbaki, *Algèbre commutative*, Hermann, Paris, 1983.

[C] J. Coates, *Motivic p-adic L-functions*, LMS Lecture Note Series, vol. 153, Cambridge University Press, London and New York, 1991, pp. 141–172.

[C-P] J. Coates and B. Perrin-Riou, *On p-adic L-functions attached to motives over* $\mathbb{Q}$, Adv. Stud. Pure Math. **17** (1989), 23–54.

[C-W1] J. Coates and A. Wiles, *On the conjecture of Birch and Swinnerton-Dyer*, Invent. Math. **39** (1977), 223–251.

[C-W2] ——, *On p-adic L-functions and elliptic units*, J. Austral. Math. Soc. (Ser. A) **26** (1978), 1–25.

[D] P. Deligne, *Valeurs de fonctions L et périodes d'intégrales*, Proc. Sympos. Pure Math. **33** (1979), 313–346.

[Dd] J. Dieudonné, *Topics on local algebra*, Notre Dame Math. Lectures **10** (1967).

[G1] R. Greenberg, *On the Birch and Swinnerton-Dyer conjecture*, Invent. Math. **72** (1983), 241–265.

[G2] ——, *Iwasawa theory for p-adic representations*, Adv. Stud. Pure Math. **17** (1989), 97–137.

[G3] ——, *Iwasawa theory for motives*, LMS Lecture Note Series, vol. 153, Cambridge University Press, London and New York, 1991, pp. 211–234.

[G-S] R. Greenberg and G. Stevens, *p-adic L-functions and p-adic periods of modular forms*, Invent. Math. **111** (1993), 407–447.

[H1] H. Hida, *Galois representations into* $GL_2(\mathbb{Z}_p[[X]])$ *attached to ordinary cusp forms*, Invent. Math. **85** (1986), 545–613.

[H2] ——, *Iwasawa modules attached to congruences of cusp forms*, Ann. Sci. École Norm. Sup. **19** (1986), 231–273.

[H3] ——, *p-adic L-functions for base change lifts of* GL_2 *to* GL_3, Automorphic Forms, Shimura Varieties and L-functions, Vol. II, Academic Press, New York, 1990, pp. 93–142.

[H4] ——, *On p-adic L-functions of* $GL(2) \times GL(2)$ *over totally real fields*, Ann. Inst. Fourier (Grenoble) **41** (1991), no. 2, 311–391.

[H-T] H. Hida and J. Tilouine, *Katz p-adic L-functions, congruence modules and deformations of Galois representations*, LMS Lecture Note Series, vol. 153, Cambridge University Press, London and New York, 1991, pp. 271–294.

[I1] K. Iwasawa, *Lectures on p-adic L-functions*, Ann. of Math. Stud. no. 74, Princeton University Press, Princeton, NJ, 1972.

[I2] ——, *On $\mathbb{Z}_l$-extensions of algebraic number fields*, Ann. of Math. **98** (1973), 246–326.

[Ka] K. Kato, *Iwasawa theory and p-adic Hodge theory*, preprint.

[Kz1] N. Katz, *p-adic interpolation of real analytic Eisenstein series*, Ann. of Math. **104** (1976), 459–571.

[Kz2] ——, *p-adic L-functions for CM Fields*, Invent. Math. **49** (1978), 199–297.

[K] K. Kitagawa, *On standard p-adic L-functions of families of elliptic cusp forms*, UCLA Ph.D. thesis, 1991.

[Ma] J. Manin, *Periods of parabolic forms and p-adic Hecke series*, Mat. Sb. **92 (134)** (1973), 378–401; English transl. in Math. USSR Sb. **21** (1973).

[M1] B. Mazur, *Rational points of abelian varieties with values in towers of number fields*, Invent. Math. **18** (1972), 183–266.

[M2] ——, *Deforming Galois representations*, Galois Groups over $\mathbb{Q}$, Springer-Verlag, Berlin and New York, 1989, pp. 385–438.

[M-SwD] B. Mazur and P. Swinnerton-Dyer, *Arithmetic of Weil Curves*, Invent. Math. **25** (1974), 1–61.

[M-T] B. Mazur and J. Tilouine, *Représentations galoisiennes, différentielles de Kähler et conjectures principales*, Inst. Hautes Études Sci. Publ. Math., no. 71 (1990), 65–103.

[M-T-T] B. Mazur, J. Tate, and J. Teitelbaum, *On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer*, Invent. Math. **84** (1986), 1–48.

[M-W1] B. Mazur and A. Wiles, *Class fields of abelian extensions of $\mathbb{Q}$*, Invent. Math. **76** (1984), 179–330.

[M-W2] ——, *On p-adic analytic families of Galois representations*, Comp. Math. **59** (1986), 231–264.

[P] A. Panchishkin, *Motives over totally real fields and p-adic L-functions*, Ann. Sci. Inst. Fourier (to appear).

[PeR] B. Perrin-Riou, *Théorie d'Iwasawa p-adique locale et globale*, Invent. Math. **99** (1990), 247–292.

[Ro] D. Rohrlich, *On L-functions of elliptic curves and cyclotomic towers*, Invent. Math. **75** (1984), 409–423.

[R1] K. Rubin, *The main conjecture*. Appendix to: Cyclotomic Fields I–II, 2nd ed. (S. Lang, ed.), Springer-Verlag, Berlin and New York, 1990.

[R2] ——, *The one variable main conjecture for elliptic curves with complex multiplication*, LMS Lecture Note Series, vol. 153, Cambridge University Press, New York, 1991, pp. 353–372.

[R3] ——, *The "main conjectures" of Iwasawa theory for imaginary quadratic fields*, Invent. Math. **103** (1991), 25–68.

[S] P. Schneider, *Motivic Iwasawa theory*, Adv. Stud. Pure Math. **17** (1989), 421–456.

[Se] J.-P. Serre, *Facteurs locaux des fonctions zêta des variétés algébriques* (*définitions et conjectures*), Sém. Delange-Pisot-Poitou 1969/70, exposé 19.

[Wa] L. Washington, *Introduction to Cyclotomic Fields*, Springer-Verlag, Berlin and New York, 1980.

[W1] A. Wiles, *On ordinary λ-adic representations associated to modular forms*, Invent. Math. **94** (1988), 529–573.

[W2] ——, *The Iwasawa conjecture for totally real fields*, Ann. of Math. **131** (1990), 493–540.

[Y] R. Yager, *On two variable p-adic L-functions*, Ann. of Math. **115** (1982), 411–449.

UNIVERSITY OF WASHINGTON, SEATTLE

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

p-adic Points of Motives

PETER SCHNEIDER

Classically one has two tools to study the group of rational points of an abelian variety over a local number field: One is the logarithm map into the tangent space. The other one is the connecting homomorphism arising from the Kummer sequence into the Galois cohomology of the Tate module. In their seminal paper **[BK]** Bloch and Kato define the group of points of a motive over a local number field directly in terms of the Galois cohomology of its p-adic realization. Moreover, using the Faltings-Fontaine-Messing theory of p-adic Galois representations they construct an exponential map from the de Rham realization into the group of points.

In this talk we observe that the conjectural formalism of motivic cohomology is perfectly suited to extend the classical Kummer theoretic approach to the setting of motives. We shall compare this to the Bloch-Kato construction. In this way we shall be naturally led to a conjectural expression of the Bloch-Kato groups of points as the cohomology of certain complexes of sheaves. These latter complexes really exist independently of the formalism of motivic cohomology; the conjecture only concerns the bijectivity of the map between the two sides in the expression. Under additional assumptions we shall prove that the two sides at least have the same dimension. We shall finish with some speculations on a formal deformation theory of motivic cohomology.

I want to thank the Research Institute for Mathematical Sciences (RIMS) at Kyoto for its support and hospitality which I enjoyed very much.

1. Review of some of the local results in [BK]

Throughout the paper $K/\mathbb{Q}_p$ is a fixed finite extension, $G_K := \mathrm{Gal}(\overline{K}/K)$ is the Galois group of an algebraic closure $\overline{K}$ of K, and $K_0 \subseteq K$ is the maximal unramified subextension. For any finite-dimensional $\mathbb{Q}_p$-vector space V

1991 *Mathematics Subject Classification.* Primary 11G25, 14G20, 14F20, 14F30; Secondary 14L05, 19E20, 19F27.

This paper is in final form and no version of it will be submitted for publication elsewhere.

with a continuous G_K-action we define

$$Crys(V) := H^0(K, B_{\mathrm{crys}} \otimes V) \quad \text{and}$$
$$DR(V) := H^0(K, B_{\mathrm{DR}} \otimes V).$$

The former is a K_0-vector space with a Frobenius f, the latter is a K-vector space with a decreasing de Rham filtration $DR(V)^{\cdot}$. For a summary of the properties of the rings B_{crys} and B_{DR} we refer to **[BK]**. One has

$$\dim_{K_0} Crys(V) \leq \dim_K DR(V) \leq \dim_{\mathbb{Q}_p} V$$

together with a natural injection

$$K \bigotimes_{K_0} Crys(V) \hookrightarrow DR(V).$$

The G_K-representation V is called

de Rham if $\dim_K DR(V) = \dim_{\mathbb{Q}_p} V$, resp.

crystalline if $\dim_{K_0} Crys(V) = \dim_{\mathbb{Q}_p} V$.

We put

$$H^1_e(K, V) := \ker(H^1(K, V) \to H^1(K, B^{f=1}_{\mathrm{crys}} \otimes V)) \quad \text{and}$$
$$H^1_f(K, V) := \ker(H^1(K, V) \to H^1(K, B_{\mathrm{crys}} \otimes V));$$

here $H^*(K, \cdot)$ denotes as usual the continuous Galois cohomology of G_K. The fundamental exact diagram in **[BK]** is

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Q}_p & \xrightarrow{\text{diag}} & B^{f=1}_{\mathrm{crys}} \oplus B^+_{\mathrm{DR}} & \xrightarrow{\beta} & B_{\mathrm{DR}} & \longrightarrow & 0 \\ & & \| & & \downarrow \subseteq & & \downarrow (0,\mathrm{id}) & & \\ 0 & \longrightarrow & \mathbb{Q}_p & \xrightarrow{\text{diag}} & B_{\mathrm{crys}} \oplus B^+_{\mathrm{DR}} & \xrightarrow{\gamma} & B_{\mathrm{crys}} \oplus B_{\mathrm{DR}} & \longrightarrow & 0, \end{array}$$

where

$$\beta(x, y) := x - y \quad \text{and} \quad \gamma(x, y) := (x - f(x), x - y).$$

Tensoring with V and passing to cohomology gives, in case V is de Rham (then $H^1(K, B^+_{\mathrm{DR}} \otimes V) \to H^1(K, B_{\mathrm{DR}} \otimes V)$ is injective!), the exact diagram

$$\begin{array}{ccccc} 0 \to H^0(K, V) & \xrightarrow{\text{diag}} & Crys(V)^{f=1} \oplus DR(V)^0 & \xrightarrow{\beta} & \\ \| & & \downarrow \subseteq & & \\ 0 \to H^0(K, V) & \xrightarrow{\text{diag}} & Crys(V) \oplus DR(V)^0 & \xrightarrow{\gamma} & \end{array}$$

$$\begin{array}{ccc} \to \quad DR(V) & \to & H^1_e(K, V) \to 0 \\ \downarrow (0,\mathrm{id}) & & \downarrow \subseteq \\ \to \quad Crys(V) \oplus DR(V) & \to & H^1_f(K, V) \to 0. \end{array}$$

Assuming from now on that V is de Rham, we obtain in particular:

- $\dim_{\mathbb{Q}_p} H^1_f(K, V) = \dim_{\mathbb{Q}_p} DR(V)/DR(V)^0 + \dim_{\mathbb{Q}_p} H^0(K, V)$;
- $H^1_f(K, V)/H^1_e(K, V) \cong Crys(V)/(1-f)Crys(V)$;
- the connecting homomorphism induces a surjective map
$$\exp\colon DR(V)/DR(V)^0 \twoheadrightarrow H^1_e(K, V)$$
with kernel $Crys(V)^{f=1}/H^0(K, V)$.

Defining
$$P(V; u) := \det_{K_0}(1 - f^{[K_0:\mathbb{Q}_p]}u; Crys(V))$$
we then have the following

FACT. *If V is de Rham with $P(V; 1) \neq 0$, then*
$$\exp\colon DR(V)/DR(V)^0 \xrightarrow{\cong} H^1_e(K, V) = H^1_f(K, V)$$
is an isomorphism.

At least philosophically this theory applies to the following geometric situation: Let M be a motive over K that is integral in the sense that we can speak about its $\mathbb{Z}_p$-adic realization T. Faltings' proof of the de Rham conjecture of Fontaine suggests that the $\mathbb{Q}_p$-adic realization $V = T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ of M always should be de Rham. Also we should have $P(V; 1) \neq 0$ if M has weight ≤ -1. In this situation Bloch and Kato propose
$$A(K) := H^1_f(K, T) := \text{inverse image of } H^1_f(K, V) \text{ in } H^1(K, T)$$
as a candidate for the (pro-p-part of the) "group of K-rational points" of M. (Actually they take T to be a Galois lattice in V.) They consider the isomorphism
$$\exp\colon DR(V)/DR(V)^0 \xrightarrow{\cong} A(K) \otimes \mathbb{Q}$$
as an "exponential map". One basic problem in the local arithmetic of M is then to understand this map on an "integral level". Since, for example, Iwasawa theory is concerned with ramified $\mathbb{Z}_p$-extensions, it is absolutely crucial not to restrict the ramification type of the field K when dealing with this problem.

2. The classical case

Let o_K, resp. $o_{\overline{K}}$, denote the ring of integers in K, resp. $\overline{K}$. Consider a connected p-divisible group $\mathcal{G}$ over o_K. Its Tate module is
$$T := T(\mathcal{G}) := \varprojlim \mathcal{G}_{p^\mu}(o_{\overline{K}}).$$
Fontaine has shown ([**Fon**, §6]) that the G_K-representation $V := T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ is de Rham. The assumption that $\mathcal{G}$ is connected implies $P(V; 1) \neq 0$. We therefore have the isomorphism
$$\exp\colon DR(V)/DR(V)^0 \xrightarrow{\cong} H^1_e(K, V).$$

Classically one starts with the exact ([**Tat,** §2.4, Corollary 1]) *Kummer sequence*

$$0 \to T \to \varprojlim \mathcal{G}(o_{\overline{K}}) \to \mathcal{G}(o_{\overline{K}}) \to 0,$$

which induces the injective connecting homomorphism

$$\mathcal{G}(o_K) \overset{\delta}{\hookrightarrow} H^1(K, T).$$

Its relation to the exponential map exp is as follows.

Step 1. For the multiplicative group we have the commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{Z}_p(1) & \longrightarrow & \varprojlim \widehat{\mathbb{G}}_m(o_{\overline{K}}) & \longrightarrow & \widehat{\mathbb{G}}_m(o_{\overline{K}}) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle \log[\ ]} & & \downarrow{\scriptstyle \log} & & \\
0 & \longrightarrow & \mathbb{Q}_p(1) & \longrightarrow & B_{\mathrm{crys}}^{f=p} \cap B_{\mathrm{DR}}^{+} & \longrightarrow & \mathbb{C}_p & \longrightarrow & 0 \\
 & & & & x & & & & \\
 & & \downarrow{\scriptstyle =} & & \downarrow & & \downarrow & & \\
 & & & & xt^{-1} \otimes t & & & & \\
0 & \longrightarrow & \mathbb{Q}_p(1) & \longrightarrow & B_{\mathrm{crys}}^{f=1} \otimes \mathbb{Z}_p(1) & \longrightarrow & (B_{\mathrm{DR}}/B_{\mathrm{DR}}^{+})(1) & \longrightarrow & 0
\end{array}$$

(see [**BK,** p. 360]). Tensoring with $T(-1)$ and passing to cohomology lead to the commutative diagram

$$\begin{array}{ccc}
H^0(K, \widehat{\mathbb{G}}_m(o_{\overline{K}}) \otimes T(-1)) & \overset{\delta}{\longrightarrow} & H^1(K, T) \\
\downarrow & & \Big| \\
H^0(K, \mathbb{C}_p \otimes T(-1)) & & \Big| \\
\downarrow & & \downarrow \\
DR(V)/DR(V)^0 & \overset{\exp}{\longrightarrow} & H^1(K, V).
\end{array}$$

Step 2. If $\mathcal{G}'$ denotes the dual p-divisible group then Cartier duality says that

$$\mathcal{G}'_{p^\mu}(o_{\overline{K}}) = \mathrm{Hom}_{o_{\overline{K}}}(\mathcal{G}_{p^\mu}, \widehat{\mathbb{G}}_m)$$

and therefore that

$$\mathrm{Hom}_{\mathbb{Z}_p}(T, T(\widehat{\mathbb{G}}_m)) = T(\mathcal{G}') = \varprojlim \mathcal{G}'_{p^\mu}(o_{\overline{K}}) = \mathrm{Hom}_{o_{\overline{K}}}(\mathcal{G}, \widehat{\mathbb{G}}_m).$$

This implies the existence of a natural G_K-equivariant map

$$\mathcal{G}(o_{\overline{K}}) \to \widehat{\mathbb{G}}_m(o_{\overline{K}}) \otimes T(-1)$$

which makes

$$\begin{array}{ccccccccc}
0 & \longrightarrow & T & \longrightarrow & \varprojlim \mathcal{G}(o_{\overline{K}}) & \longrightarrow & \mathcal{G}(o_{\overline{K}}) & \longrightarrow & 0 \\
 & & \Vert & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & T & \longrightarrow & \varprojlim \widehat{\mathbb{G}}_m(o_{\overline{K}}) \otimes T(-1) & \longrightarrow & \widehat{\mathbb{G}}_m(o_{\overline{K}}) \otimes T(-1) & \longrightarrow & 0
\end{array}$$

commutative.

Combining the two steps we obtain the commutative diagram

$$\begin{array}{ccc} \mathscr{G}(o_K) & \overset{\delta}{\hookrightarrow} & H^1_e(K, T) \\ \downarrow & & \Big| \\ H^0(K, \mathbb{C}_p \otimes T(-1)) & & \Big| \\ \downarrow \cong & & \downarrow \\ DR(V)/DR(V)^0 & \xrightarrow[\cong]{\exp} & H^1_e(K, V). \end{array}$$

Via Hodge-Tate theory the space $H^0(K, \mathbb{C}_p \otimes T(-1))$ can be naturally identified with the tangent space of $\mathscr{G}$ (**[Tat,** §4, Theorem 3]). The map $\mathscr{G}(o_K) \to H^0(K, \mathbb{C}_p \otimes T(-1))$ then becomes the usual logarithm map for $\mathscr{G}$, which is a local isomorphism (**[Tat,** §2.4]).

RÉSUMÉ. *From the Kummer sequence for* $\mathscr{G}$ *we obtain a local inverse*

$$\log: \mathscr{G}(o_K) \to DR(V)/DR(V)^0$$

of the exponential map.

3. A digression on weight arguments

From now on we fix a smooth projective variety X over K and integers $i \geq 0$ and $n \in \mathbb{Z}$, and we consider the motive $M := H^i(X)(n)$. This means that we consider the G_K-representation $V := H^i(\overline{X}, \mathbb{Q}_p(n))$ together with the Galois lattice

$$T := H^i(\overline{X}, \mathbb{Z}_p(n))/\operatorname{tor} \hookrightarrow V = H^i(\overline{X}, \mathbb{Q}_p(n))$$

in it; here $\overline{X} := X \times_K \overline{K}$ and cohomology, if not indicated otherwise, is always étale cohomology. Faltings (**[Fal]**) has proved that V is de Rham. We have $\text{weight}(M) = i - 2n$.

Since the group G_K has cohomological dimension 2, the groups $H^*(K, V)$ can be nonzero at most for $0 \leq * \leq 2$. Using Poincaré duality together with a polarization we deduce a G_K-isomorphism

$$\operatorname{Hom}_{\mathbb{Q}_p}(V, \mathbb{Q}_p) \cong V(i - 2n).$$

From local Tate duality we therefore obtain that

$$H^2(K, V) \text{ is dual to } H^0(K, V(i + 1 - 2n)).$$

Throughout the paper we assume that X has *good reduction.* We fix a smooth projective model $\mathscr{X}$ of X over $\operatorname{Spec}(o_K)$, and we let $Y := \mathscr{X} \times_{o_K} k$ denote the reduction of $\mathscr{X}$ over the residue class field k of o_K. By Faltings

([**Fal**]) V is crystalline and

$$Crys(V) = \left[H^i_{\mathrm{crys}}(Y/W(k)) \bigotimes_{W(k)} K_0 \right](n)$$

identifies with the (twisted) crystalline cohomology of the reduction Y. The latter implies, by an argument of Katz and Messing ([**KM**]) based on Deligne's proof of the Weil conjectures, that the reciprocal roots of $P(V; u)$ are algebraic integers of complex absolute value $q^{(i-2n)/2} = q^{\mathrm{weight}(M)/2}$. We see that if weight $(M) \neq 0$ then we have $P(V; 1) \neq 0$ as well as $H^0(K, V) = 0$.

LEMMA. (i) *If* $i \neq 2n$ *then* $P(V; 1) \neq 0$ *and* $H^0(K, V) = 0$.
(ii) *If* $i \neq 2n-2$ *then* $H^2(K, V) = 0$.
(iii) *If* $i \neq 2n-1$ *then* $H^1(K, V) = H^{i+1}(X, \mathbb{Q}_p(n))$.

PROOF. (i) This restates the result of our discussion above. (ii) By duality as explained above, this reduces to the assertion (i) for the motive $H^i(X)(i+1-n)$. (iii) This follows from the Hochschild-Serre spectral sequence for G_K.

4. Kummer sequences through motivic cohomology

From now on we assume $n \geq 0$ and let $\mathbb{Z}(n)_X$ or simply $\mathbb{Z}(n)$ denote the nth Beilinson-Lichtenbaum complex on the étale site $X_{\text{ét}}$ of X. One hopes that such objects fulfilling a certain list of nice properties exist and form the ultimate "motivic" cohomology theory. For the background we refer to [**Licl**]. The properties in that list which will be of use for us will be introduced when they are needed. We shall number them by "Mot?" followed by a reference to where this property is stated in the literature, possibly with some explanation. All our assertions that depend on those hypothetical properties will be called "Claims".

To begin with, the most important property for our purposes is:

MOT 1 [**Licl**]. There are compatible distinguished triangles of complexes of sheaves on $X_{\text{ét}}$

$$\begin{array}{ccc} & \mathbb{Z}/p^\mu\mathbb{Z}(n) & \\ {}^{+1}\swarrow & & \nwarrow \\ \mathbb{Z}(n) & \xrightarrow{p^\mu} & \mathbb{Z}(n) \end{array} \qquad \text{for } \mu \geq 1.$$

The associated cohomology sequences are

$$0 \to H^i(\overline{X}, \mathbb{Z}(n))/p^\mu \to H^i(\overline{X}, \mathbb{Z}/p^\mu\mathbb{Z}(n)) \to H^{i+1}(\overline{X}, \mathbb{Z}(n))^{p^\mu=0} \to 0.$$

We define

$$H^*(\overline{X}, \mathbb{Z}(n))^0 := \text{maximal } p\text{-divisible subgroup in } H^*(\overline{X}, \mathbb{Z}(n)).$$

Our first aim is to show that the above exact sequences in the projective limit lead, for $i \neq 2n$, to the exact *Kummer sequence*

$$(*) \qquad 0 \to T \to \varprojlim H^{i+1}(\overline{X}, \mathbb{Z}(n))^0 \to H^{i+1}(\overline{X}, \mathbb{Z}(n))^0 \to 0,$$

where the projective limit in the middle is formed with respect to multiplication by p as transition maps. For trivial reasons we obtain the exact sequence

$$0 \to \varprojlim H^i(\overline{X}, \mathbb{Z}(n))/p^\mu \to H^i(\overline{X}, \mathbb{Z}_p(n)) \to$$
$$\to \varprojlim H^{i+1}(\overline{X}, \mathbb{Z}(n)) \to H^{i+1}(\overline{X}, \mathbb{Z}(n))^0 \to 0.$$

CLAIM 1.

$$\begin{aligned} H^*(\overline{X}, \mathbb{Z}(n))^0 &= \ker(H^*(\overline{X}, \mathbb{Z}(n)) \to \varprojlim H^*(\overline{X}, \mathbb{Z}(n))/p^\mu) \\ &= \ker(H^*(\overline{X}, \mathbb{Z}(n)) \to H^*(\overline{X}, \mathbb{Z}_p(n))). \end{aligned}$$

PROOF. The second equality follows from the injectivity of the map

$$\varprojlim H^*(\overline{X}, \mathbb{Z}(n))/p^\mu \to H^*(\overline{X}, \mathbb{Z}_p(n)).$$

Concerning the first equality, the inclusion "$\subseteq$" is clear since $H^*(\overline{X}, \mathbb{Z}_p(n))$ is a finitely generated $\mathbb{Z}_p$-module. On the other hand, if p^s is the order of the torsion subgroup in $H^*(\overline{X}, \mathbb{Z}_p(n))$ and if x is in the kernel under consideration then, for any $r \geq 0$, we find a $y_r \in H^*(\overline{X}, \mathbb{Z}(n))$ such that $x = p^r y_r$; we get $x = p(p^s y_{s+1})$, where $p^s y_{s+1}$ also lies in the kernel. This shows that the kernel is p-divisible.

Again, since $H^{i+1}(\overline{X}, \mathbb{Z}_p(n))$ is a finitely generated $\mathbb{Z}_p$-module, Claim 1 implies that

$$\varprojlim H^{i+1}(\overline{X}, \mathbb{Z}(n))^0 = \varprojlim H^{i+1}(\overline{X}, \mathbb{Z}(n)).$$

Since this group obviously is p-torsion-free, it remains to show that

$$\varprojlim H^i(\overline{X}, \mathbb{Z}(n))/p^\mu \text{ is finite for } i \neq 2n.$$

This follows from the next claim.

CLAIM 2. *If $i \neq 2n$ then $H^i(\overline{X}, \mathbb{Z}(n))^0$ is of finite index in $H^i(\overline{X}, \mathbb{Z}(n))$.*

PROOF. By Claim 1 it suffices to show that

$$\operatorname{im}(H^i(\overline{X}, \mathbb{Z}(n)) \to H^i(\overline{X}, \mathbb{Z}_p(n))) \text{ is finite.}$$

We use

MOT 2 (implicitly in **[Lic1]**). The formation of $\mathbb{Z}(n)$ commutes with pro-étale base change.

This implies that

$$H^*(\overline{X}, \mathbb{Z}(n)) = \varinjlim_{\substack{K \subseteq L \subseteq \overline{K} \\ L/K \text{ finite}}} H^*(X_{/L}, \mathbb{Z}(n)).$$

Therefore, the above image is contained in that part of $H^i(\overline{X}, \mathbb{Z}_p(n))$ on which G_K acts discretely. But as explained in §3, the discrete part of $H^i(\overline{X}, \mathbb{Q}_p(n))$ is zero if $i \neq 2n$.

COMMENT. The above considerations actually show that

$$H^i(\overline{X}, \mathbb{Z}(n))/H^i(\overline{X}, \mathbb{Z}(n))^0 = \operatorname{Tor} H^i(\overline{X}, \mathbb{Z}_p(n)) \quad \text{for } i \neq 2n.$$

Is the cokernel of

$$H^{2n}(\overline{X}, \mathbb{Z}(n))/H^{2n}(\overline{X}, \mathbb{Z}(n))^0 \hookrightarrow H^{2n}(\overline{X}, \mathbb{Z}_p(n))^{\text{discrete}}$$

finite?

The above Kummer sequence $(*)$ provides us, for $i \neq 2n$, with the connecting homomorphism

$$H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))^0) \xrightarrow{\delta} H^1(K, T).$$

CLAIM 3. *If* $i \neq 2n$ *then* $\ker(\delta)$ *is uniquely p-divisible.*

PROOF. By our weight assumption we have $H^0(K, T) = 0$ and therefore

$$\ker(\delta) = \varprojlim H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))^0),$$

which obviously is uniquely p-divisible.

EXAMPLE. For $i = n = 1$ we have $\operatorname{im}(\delta) = \operatorname{Pic}(X)^0(K)$. This follows from

MOT 3 **[Lic1]**. $\mathbb{Z}(1) = \mathbb{G}_m[-1]$.

Assuming $i \neq 2n$ we would like to have a commutative diagram

$(**)$

$$\begin{array}{ccc} ? \subseteq H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))^0) & \xrightarrow{\delta} & H^1(K, T) \\ \downarrow & & \downarrow \\ DR(V)/DR(V)^0 & \xrightarrow[\cong]{\exp} & H^1_f(K, V) \subseteq H^1(K, V). \end{array}$$

(E.g., for $i = 0$ and $n = 1$ we have $? = o_K^\times \subseteq K^\times$.) This would lead to a local inverse

$$\begin{array}{ccc} \delta(?) & \subseteq & A(K) \\ \swarrow_{\log} & & \downarrow \\ DR(V)/DR(V)^0 & \xrightarrow[\cong]{\exp} & A(K) \otimes \mathbb{Q} \end{array}$$

of the exponential map. Of course, the question remains whether $\delta(?)$ is of finite index in $A(K)$, or even whether $\delta(?)$ is a $\mathbb{Z}_p$-submodule of $A(K)$ (and also whether $\delta(?)$ only depends on M and not on the particular representation $M = H^i(X)(n)$).

5. The comparing diagram

A first simplification of the desired diagram $(**)$ is achieved by using Galois descent. We will exclude the cases $i = 2n, 2n \pm 1$, which are more complicated.

CLAIM 4 (Galois descent). *For $i \neq 2n$, $2n \pm 1$ we have the exact diagram*

$$\begin{array}{ccccc} & & H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))^0) & & \\ & & \downarrow \subseteq & & \\ H^{i+1}(X, \mathbb{Z}(n)) & \longrightarrow & H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))) & \longrightarrow & \textit{finite} \\ & & \downarrow & & \\ & & \textit{finite.} & & \end{array}$$

PROOF. The column follows from Claim 2. For the row we look at the Hochschild-Serre spectral sequence (whose existence follows from Mot 2)

$$H^r(K, H^s(\overline{X}, \mathbb{Z}(n))) \Rightarrow H^{r+s}(X, \mathbb{Z}(n)).$$

Since K has strict cohomological dimension 2, it suffices to prove that

$$H^2(K, H^i(\overline{X}, \mathbb{Z}(n))) \text{ is finite.}$$

Passing in the cohomology sequences associated with the triangles in Mot 1 to the direct limit we obtain the exact sequence

$$0 \to H^{i-1}(\overline{X}, \mathbb{Z}(n)) \otimes \mathbb{Q}_p/\mathbb{Z}_p \to H^{i-1}(\overline{X}, \mathbb{Q}_p/\mathbb{Z}_p(n)) \to H^i(\overline{X}, \mathbb{Z}(n))(p) \to 0.$$

By Claim 2 the first group vanishes. We therefore have

$$H^i(\overline{X}, \mathbb{Z}(n))^0(p) = H^{i-1}(\overline{X}, \mathbb{Q}_p/\mathbb{Z}_p(n))^0,$$

where also on the right-hand side the superscript ^{0} stands for the maximal divisible subgroup. This implies

$$H^2(K, H^i(\overline{X}, \mathbb{Z}(n))^0) = H^2(K, H^{i-1}(\overline{X}, \mathbb{Q}_p/\mathbb{Z}_p(n))^0).$$

By the arguments in §3 the right-hand side is dual to the Galois invariants of a Galois lattice in $H^{i-1}(\overline{X}, \mathbb{Q}_p(i-n))$ and consequently is zero because $i - 1 \neq 2(i - n)$. Again by Claim 2, the left-hand group differs from $H^2(K, H^i(\overline{X}, \mathbb{Z}(n)))$ only by a finite group.

We have seen in §3 that for $i \neq 2n - 1$ the Hochschild-Serre spectral sequence induces the isomorphism

$$H^1(K, V) \cong H^{i+1}(X, \mathbb{Q}_p(n)).$$

In case $i \neq 2n$, $2n \pm 1$ the diagram

$$\begin{array}{ccccc} H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))^0) & \xrightarrow{\delta} & H^1(K, T) & \longrightarrow & H^1(K, V) \\ \downarrow & & & & \\ H^0(K, H^{i+1}(\overline{X}, \mathbb{Z}(n))) & & & & \uparrow \cong \\ \uparrow & & & & \\ H^{i+1}(X, \mathbb{Z}(n)) & \xrightarrow{c_{\text{ét}}} & H^{i+1}(X, \mathbb{Z}_p(n)) & \longrightarrow & H^{i+1}(X, \mathbb{Q}_p(n)) \end{array}$$

then, for a completely formal reason, is commutative after tensoring with $\mathbb{Q}$ (so that by Claim 4 the upper, resp. lower, vertical arrow on the left-hand side becomes an isomorphism, resp. surjective). The natural map $c_{\text{ét}}$ is induced by the distinguished triangles in Mot 1. What we are looking for, therefore, is (in case $i \neq 2n, 2n \pm 1$) a commutative diagram

$(**_K)$

$$\begin{array}{ccccc} ? & \subseteq H^{i+1}(X, \mathbb{Z}(n)) & \xrightarrow{c_{\text{ét}}} & & H^{i+1}(X, \mathbb{Q}_p(n)) \\ \downarrow & & & & \updownarrow\cong \\ DR(V)/DR(V)^0 & \xrightarrow[\cong]{\exp} & H^1_f(K, V) \subseteq & & H^1(K, V). \end{array}$$

REMARKS. (1) By Mot 1 the map $c_{\text{ét}}$ factors through an injective map

$$\varprojlim H^{i+1}(X, \mathbb{Z}(n))/p^\mu \overset{c_{\text{ét}}}{\hookrightarrow} H^{i+1}(X, \mathbb{Z}_p(n)).$$

(2) We have

$$\begin{aligned} \dim_{\mathbb{Q}_p} H^1(K, V) &= \sum_{r \geq 0} (-1)^{r+1} \dim_{\mathbb{Q}_p} H^r(K, V) \\ &= [K : \mathbb{Q}_p] \cdot \dim_{\mathbb{Q}_p} V, \end{aligned}$$

provided $i \neq 2n$ and $2n - 2$; the first equality is a consequence of the lemma in §3, and the second one follows from the theorem about local Euler-Poincaré characteristics. On the other hand, Faltings' result gives a natural identification

$$DR(V) = H^i_{\text{DR}}(X)(n)$$

with the (twisted) algebraic de Rham cohomology of X over K. For $i \neq 2n$ we obtain

$$\begin{aligned} \dim_{\mathbb{Q}_p} H^1_f(K, V) &= [K : \mathbb{Q}_p] \cdot \dim_K H^i_{\text{DR}}(X)/F^n \\ &= [K : \mathbb{Q}_p] \cdot (\dim_{\mathbb{Q}_p} V - \dim_K F^n H^i_{\text{DR}}(X)). \end{aligned}$$

This shows that

$$H^1_f(K, V) = H^1(K, V)$$

if $i < n$ with the exception of the case $i = 0$, $n = 1$. In this situation the existence of a diagram $(**_K)$ of course is trivial.

6. The vanishing cycle spectral sequence

One may view the existence of a diagram $(**_K)$ as a statement about the image of $c_{\text{ét}}$. From this point of view it seems important to find a characterization of the subspace in $H^{i+1}(X, \mathbb{Q}_p(n))$ that corresponds to $H^1_f(K, V)$:

$$\begin{array}{ccc} H^1_f(K, V) \subseteq & H^1(K, V) & \\ \updownarrow\cong & \updownarrow\cong & \qquad (i \neq 2n-1). \\ ? \quad \subseteq & H^{i+1}(X, \mathbb{Q}_p(n)) & \end{array}$$

We shall approach this question by using the p-adic vanishing cycle spectral sequence. Let

$$Y \xrightarrow{\sigma} \mathscr{X} \xleftarrow{\tau} X$$

denote the immersions of the fibers into the model $\mathscr{X}$. It seems that the canonical filtration $t_{\leq \cdot} R\tau_* \mathscr{F}^{\cdot}$ of the total direct image complex $R\tau_* \mathscr{F}^{\cdot}$ behaves rather differently in the two cases $\mathscr{F}^{\cdot} = \mathbb{Z}(n)$ and $\mathscr{F}^{\cdot} = \mathbb{Z}/p^{\mu}\mathbb{Z}(n)$, respectively. This is what we want to explore.

In complete generality the decomposition theorem in étale topology states that for any bounded below complex of sheaves $\mathscr{C}^{\cdot}$ on the small étale site of $\mathscr{X}$ one has a natural distinguished triangle

$$\begin{array}{ccc} & R\tau_*\tau^*\mathscr{C}^{\cdot} & \\ {}^{+1}\swarrow & & \nwarrow \\ \sigma_* R\sigma^! \mathscr{C}^{\cdot} & \rightarrow & \mathscr{C}^{\cdot}. \end{array}$$

If $\mathscr{C}^{\cdot}$ is acyclic in degrees $> n$, then

$$\begin{array}{ccc} & t_{\leq n} R\tau_*\tau^*\mathscr{C}^{\cdot} & \\ {}^{+1}\swarrow & & \nwarrow \\ \sigma_* t_{\leq n+1} R\sigma^! \mathscr{C}^{\cdot} & \rightarrow & \mathscr{C}^{\cdot} \end{array}$$

is a distinguished triangle, too, and

$$R^{n+j}\tau_*\tau^*\mathscr{C}^{\cdot} = \sigma_* R^{n+j+1}\sigma^! \mathscr{C}^{\cdot} \quad \text{for } j \geq 1.$$

Because of

MOT 4 **[Lic1]**. $\mathbb{Z}(0) = \mathbb{Z}$, and $\mathbb{Z}(n)$, for $n \geq 1$, is acyclic outside $[1, n]$ we may apply this to the complex $\mathscr{C}^{\cdot} = \mathbb{Z}(n)_{\mathscr{X}}$. Using

MOT 5 (Purity, **[Lic2]** and **[Mil2]**). $t_{\leq n+1} R\sigma^! \mathbb{Z}(n)_{\mathscr{X}} = \mathbb{Z}(n-1)_Y[-2]$ we obtain the distinguished triangle

$$\begin{array}{ccc} & t_{\leq n} R\tau_* \mathbb{Z}(n)_X & \\ {}^{+1}\swarrow & & \nwarrow \\ \sigma_* \mathbb{Z}(n-1)_Y[-2] & \rightarrow & \mathbb{Z}(n)_{\mathscr{X}} \end{array}$$

and the isomorphisms

$$R^{n+j}\tau_* \mathbb{Z}(n)_X = \sigma_* R^{n+j+1}\sigma^! \mathbb{Z}(n)_{\mathscr{X}} \quad \text{for } j \geq 1.$$

The associated cohomology sequence reads

$$\cdots \to H^{i-1}(Y, \mathbb{Z}(n-1)) \to H^{i+1}(\mathscr{X}, \mathbb{Z}(n)) \to H^{i+1}(\mathscr{X}, t_{\leq n} R\tau_* \mathbb{Z}(n)_X)$$
$$\to H^i(Y, \mathbb{Z}(n-1)) \to \cdots .$$

We now invoke

MOT 6 **[Lic1]**. $H^*(Y, \mathbb{Z}(n-1))$ is finite for $* \neq 2n-2, 2n$ and is finitely generated for $* = 2n-2$.

Excluding the cases $i = 2n$, $2n \pm 1$, or $2n - 2$ (which involve the codimension $n - 1$ cycles on Y through the term $H^{2n-2}(Y, \mathbb{Z}(n-1))$) we see that the map

$$H^{i+1}(\mathscr{X}, \mathbb{Z}(n)) \to H^{i+1}(\mathscr{X}, t_{\leq n} R\tau_* \mathbb{Z}(n)_X)$$

should have finite kernel and cokernel. Next we have to look at the map

$$H^{i+1}(\mathscr{X}, t_{\leq n} R\tau_* \mathbb{Z}(n)_X) \to H^{i+1}(\mathscr{X}, R\tau_* \mathbb{Z}(n)_X) = H^{i+1}(X, \mathbb{Z}(n)).$$

We hope that its kernel and cokernel are torsion and therefore could be neglected in looking for our logarithm map since the latter will take values in a $\mathbb{Q}_p$-vector space. For trivial reasons, of course, the map is an isomorphism if $i < n$. What are the reasons for our hope?

A) (Compare **[Mil2**, p. 74, last paragraph].) It might be possible that the purity axiom Mot 5 can be complemented by the requirement that the sheaves

$$R^j \sigma^! \mathbb{Z}(n)_{\mathscr{X}} \text{ for } j \geq n + 2 \geq 3 \text{ are } p\text{-primary torsion.}$$

Then the sheaves

$$R^j \tau_* \mathbb{Z}(n)_X \text{ for } j > n \geq 1 \text{ are } p\text{-primary torsion, too,}$$

which would imply that the kernels and cokernels in question are p-primary torsion if $n \geq 1$.

CLAIM 5. *In case* $X = \operatorname{Spec}(K)$ *we have* $R^j \tau_* \mathbb{Z}(n)_X = 0$ *for* $j > n \geq 1$.

PROOF. If I denotes the inertia subgroup in G_K we have

$$R^j \tau_* \mathbb{Z}(n)_X = H^j(I, \mathbb{Z}(n)_X).$$

Since I has strict cohomological dimension two and $\mathbb{Z}(n)$ is acyclic in degrees $> n$ by Mot 4 it is clear that $H^j(I, \mathbb{Z}(n)_X) = 0$ for $j > n + 2$. The same argument shows that for $j = n + 2$ it suffices to prove the vanishing of $H^2(I, h^n(\mathbb{Z}(n)_X))$. In case $n = 1$ this is, by Mot 3, the well-known vanishing of $H^2(I, \overline{K}^\times)$. For $n \geq 2$ it follows from Mot 1 that $h^n(\mathbb{Z}(n)_X)$ is uniquely divisible and therefore cohomologically trivial. The vanishing of $H^{n+1}(I, \mathbb{Z}(n)_X)$ is a consequence of

MOT 7 (Hilbert 90, **[Lic1]**). $R^{n+1} \alpha_* \mathbb{Z}(n) = 0$, where $\alpha: X_{\text{ét}} \to X_{\text{Zar}}$ is the natural morphism of sites.

B) In the presence of other expected properties of the complexes $\mathbb{Z}(n)$ the purity axiom Mot 5 actually can be sharpened:

CLAIM 6. $R^{n+1} \tau_* \mathbb{Z}(n)_X = \sigma_* R^{n+2} \sigma^! \mathbb{Z}(n)_{\mathscr{X}} = 0$.

PROOF. Let $\overline{x}$ be a geometric point of $\mathscr{X}$. We have to show that

$$(R^{n+1} \tau_* \mathbb{Z}(n)_X)_{\overline{x}} = H^{n+1}(\mathscr{O}_{\mathscr{X}, \overline{x}}[\tfrac{1}{p}], \mathbb{Z}(n))$$

(use Mot 2) vanishes. The axiom about the K-theoretic Beilinson complexes

MOT 8 **[Lic1]**. $\mathbb{Z}_B(n) = t_{\leq n} R\alpha_* \mathbb{Z}(n)$

together with Mot 7 implies that

$$H^{n+1}(\mathscr{O}_{\mathscr{X},\bar{x}}[\tfrac{1}{p}], \mathbb{Z}(n)) = H^{n+1}_{\mathrm{Zar}}(\mathscr{O}_{\mathscr{X},\bar{x}}[\tfrac{1}{p}], \mathbb{Z}_B(n)).$$

But in the K-theoretic localization sequence

$$\cdots \to H^{n+1}_{\mathrm{Zar}}(\mathscr{O}_{\mathscr{X},\bar{x}}, \mathbb{Z}_B(n)) \to H^{n+1}_{\mathrm{Zar}}(\mathscr{O}_{\mathscr{X},\bar{x}}[\tfrac{1}{p}], \mathbb{Z}_B(n)) \to$$
$$\to H^{n}_{\mathrm{Zar}}(\mathscr{O}_{\mathscr{X},\bar{x}}/p, \mathbb{Z}_B(n-1)) \to \cdots$$

the outer terms vanish being the Zariski cohomology over a local ring in a degree in which the coefficient complex is acyclic.

As a consequence the map in question would be an isomorphism even for $i \leq n$.

C) A similar argument as in the proof of Claim 6 shows that

$$R\tau_*^{\mathrm{Zar}}\mathbb{Z}_B(n)_X \text{ is acyclic in degrees} > n.$$

Because of

$$t_{\leq n}R\tau_*^{\mathrm{Zar}}\mathbb{Z}_B(n)_X = t_{\leq n}R\tau_*^{\mathrm{Zar}}(t_{\leq n}R\alpha_*\mathbb{Z}(n)_X)$$
$$= t_{\leq n}R\alpha_*(t_{\leq n}R\tau_*\mathbb{Z}(n)_X),$$

this leads to the factorization

$$\begin{array}{ccc} H^*_{\mathrm{Zar}}(X, \mathbb{Z}_B(n)) & \to & H^*(X, \mathbb{Z}(n)) \\ \searrow & & \nearrow \\ & H^*(\mathscr{X}, t_{\leq n}R\tau_*\mathbb{Z}(n)_X) & \end{array}$$

showing that we lose nothing coming from K-theory. Also it seems reasonable to believe that the horizontal arrow always has torsion kernel and cokernel. For $\mathbb{Z}(1) = \mathbb{G}_m[-1]$ this is shown to be the case in **[Gro,** Proposition 1.4].

We have the obvious commutative diagram

$$\begin{array}{ccc} H^{i+1}(\mathscr{X}, \mathbb{Z}(n)) & \longrightarrow & H^{i+1}(X, \mathbb{Z}(n)) \\ \downarrow & & \downarrow c_{\text{ét}} \\ H^{i+1}\left(\mathscr{X}, \mathbb{Z}(n)_{\mathscr{X}} \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z}\right) & \longrightarrow & H^{i+1}\left(X, \mathbb{Z}(n)_X \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z}\right) \\ & & \| \\ & & H^{i+1}(X, \mathbb{Z}/p^\mu\mathbb{Z}(n)), \end{array}$$

where the identity on the right-hand side comes from Mot 1. So far we have discussed the upper horizontal arrow. We now turn to the computation of the complexes

$$S_\mu(n) := \mathbb{Z}(n)_{\mathscr{X}} \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z}.$$

With $\mathbb{Z}(n)_{\mathscr{X}}$ (by Mot 4) also $S_\mu(n)$ is acyclic in degrees $> n$. The decomposition theorem then gives us the distinguished triangle

$$\begin{array}{ccc} & t_{\leq n} R\tau_* \tau^* S_\mu(n) & \\ {}^{+1}\swarrow & & \nwarrow \\ \sigma_* t_{\leq n+1} R\sigma^! S_\mu(n) & \rightarrow & S_\mu(n). \end{array}$$

For the further arguments observe first that $\cdot \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z}$ commutes with any other triangulated functor.

- From Mot 1 we deduce

$$\tau^* S_\mu(n) = \tau^* \mathbb{Z}(n)_{\mathscr{X}} \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z} = \mathbb{Z}(n)_X \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z} = \mathbb{Z}/p^\mu\mathbb{Z}(n).$$

- From Claim 6 we deduce that the distinguished triangle

$$\begin{array}{ccc} & R\sigma^! S_\mu(n) & \\ {}^{+1}\swarrow & & \nwarrow \\ R\sigma^! \mathbb{Z}(n)_{\mathscr{X}} & \xrightarrow{p^\mu} & R\sigma^! \mathbb{Z}(n)_{\mathscr{X}} \end{array}$$

remains a distinguished triangle after $t_{\leq n+1}$-truncation. Using in addition the purity axiom Mot 5 we therefore obtain

$$\begin{aligned} t_{\leq n+1} R\sigma^! S_\mu(n) &= (t_{\leq n+1} R\sigma^! \mathbb{Z}(n)_{\mathscr{X}}) \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z} \\ &= (\mathbb{Z}(n-1)_Y \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z})[-2]. \end{aligned}$$

About the p-torsion in characteristic p we have

Mot 9 **[Mil2]**. $\mathbb{Z}(n)_Y \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z} = \nu_\mu(n)[-n]$.

We recall the definition of the sheaves $\nu_\mu(n)$ on the étale site $Y_{\text{ét}}$: For $n = 0$ put

$$\nu_\mu(0) := \mathbb{Z}/p^\mu\mathbb{Z} \subseteq W_\mu \mathscr{O}_Y.$$

Using the homomorphism

$$\begin{aligned} d\log\colon \mathscr{O}_Y^\times &\to W_\mu \Omega_Y^1, \\ f &\mapsto \frac{d\underline{f}}{\underline{f}}, \qquad \underline{f} := \text{Teichmüller representative of } f, \end{aligned}$$

we define for $n \geq 1$

$$\begin{aligned} \nu_\mu(n) := {} & \text{additive subsheaf in } W_\mu\Omega_Y^n \text{ generated étale-} \\ & \text{locally by the sections } d\log f_1 \wedge \cdots \wedge d\log f_n. \end{aligned}$$

Equivalently we have

$$\begin{aligned} \nu_\mu(n) = \text{image } (&\underline{\underline{K}}^{\text{Milnor}}_{n, Y_{\text{ét}}} \xrightarrow{\text{symbol map}} W_\mu\Omega_Y^n), \\ &\{f_1, \dots, f_n\} \mapsto d\log f_1 \wedge \cdots \wedge d\log f_n. \end{aligned}$$

Putting everything together we have the distinguished triangle

$$\begin{array}{ccc} & t_{\leq n} R\tau_* \mathbb{Z}/p^\mu \mathbb{Z}(n) & \\ {}^{+1}\swarrow & & \nwarrow \\ \sigma_* \nu_\mu(n-1)[-n-1] & \rightarrow & S_\mu(n), \end{array}$$

resp., by translation

$$S_\mu(n) = \operatorname{cone}(t_{\leq n} R\tau_* \mathbb{Z}/p^\mu \mathbb{Z}(n) \rightarrow \sigma_* \nu_\mu(n-1)[-n])[-1].$$

This cone is determined by a sheaf homomorphism

$$s\colon R^n \tau_* \mathbb{Z}/p^\mu \mathbb{Z}(n) \rightarrow \sigma_* \nu_\mu(n-1)$$

at which we shall take a closer look.

As a consequence of Claim 6 the natural homomorphism

$$R^n \tau_* \mathbb{Z}(n)_X \twoheadrightarrow R^n \tau_* \mathbb{Z}/p^\mu \mathbb{Z}(n)$$

is surjective so that s is determined by the composite homomorphism

$$R^n \tau_* \mathbb{Z}(n)_X \rightarrow \sigma_* \nu_\mu(n-1),$$

or, equivalently by adjointness, by a homomorphism

$$\sigma^* R^n \tau_* \mathbb{Z}(n)_X \rightarrow \nu_\mu(n-1).$$

The decomposition theorem together with Mot 4 gives us the exact sequence

$$h^n(\mathbb{Z}(n)_{\mathscr{X}}) \rightarrow R^n \tau_* \mathbb{Z}(n)_X \rightarrow \sigma_* R^{n+1} \sigma^! \mathbb{Z}(n)_{\mathscr{X}} \rightarrow 0,$$

resp.

$$\sigma^* h^n(\mathbb{Z}(n)_{\mathscr{X}}) \rightarrow \sigma^* R^n \tau_* \mathbb{Z}(n)_X \rightarrow R^{n+1} \sigma^! \mathbb{Z}(n)_{\mathscr{X}} \rightarrow 0.$$

By the purity axiom Mot 5 the right-hand side is equal to $h^{n-1}(\mathbb{Z}(n-1)_Y)$. We now use

Mot 10 (**[Lic2]** and **[Mil2]**). $R^n \alpha_* \mathbb{Z}(n) = \underline{\underline{K}}^{\text{Milnor}}_{n,\text{Zar}}$.

It implies

$$h^n(\mathbb{Z}(n)) = \underline{\underline{K}}^{\text{Milnor}}_{n,\text{ét}},$$

so that the above exact sequence finally becomes

$$\sigma^* \underline{\underline{K}}^{\text{Milnor}}_{n,\mathscr{X}_{\text{ét}}} \rightarrow \sigma^* R^n \tau_* \mathbb{Z}(n)_X \rightarrow \underline{\underline{K}}^{\text{Milnor}}_{n-1,Y_{\text{ét}}} \rightarrow 0.$$

The expected compatibility of the various axioms with multiplicative structures (see **[Mil2]**) then shows that the exact diagram

$$\begin{array}{ccc}
\sigma^* \underline{\underline{K}}^{\text{Milnor}}_{n,\mathscr{X}_{\text{ét}}}/p^\mu & & \\
\downarrow & & \\
\sigma^* R^n\tau_*\mathbb{Z}(n)_X/p^\mu & \xrightarrow{\cong} & \sigma^* R^n\tau_*\mathbb{Z}/p^\mu\mathbb{Z}(n) \\
\downarrow & & \downarrow{\scriptstyle \sigma^* s} \\
\underline{\underline{K}}^{\text{Milnor}}_{n-1,Y_{\text{ét}}}/p^\mu & \xrightarrow[\text{symbol map}]{\cong} & \nu_\mu(n-1) \\
\downarrow & & \\
0 & &
\end{array}$$

is commutative (the lower isomorphism comes from Mot 9 and Mot 10). Since by Mot 10 and Mot 4 we have

$$\underline{\underline{K}}^{\text{Milnor}}_{n,\mathscr{X}_{\text{ét}}}/p^\mu = h^n(\mathbb{Z}(n)_{\mathscr{X}})/p^\mu = h^n(S_\mu(n)),$$

we arrive at an exact sequence

$$0 \to \underline{\underline{K}}^{\text{Milnor}}_{n,\mathscr{X}_{\text{ét}}}/p^\mu \to R^n\tau_*\mathbb{Z}/p^\mu\mathbb{Z}(n) \xrightarrow{s} \sigma_*\nu_\mu(n-1) \to 0.$$

7. A conjecture

Bloch and Kato show in [**BKØ**, (6.1.1)] that the composed homomorphism

$$(\sigma^*\tau_*\mathbb{G}_m)^{\otimes n} \to (\sigma^* R^1\tau_*\mathbb{Z}/p^\mu\mathbb{Z}(1))^{\otimes n} \xrightarrow{\text{cupproduct}} \sigma^* R^n\tau_*\mathbb{Z}/p^\mu\mathbb{Z}(n)$$

is surjective. This enables them, by using the above symbol map, to actually define the homomorphism $\sigma^* s$ (and then also s) independently of any conjecture about motivic cohomology [loc. cit., (6.6)]. We therefore may take

$$S_\mu(n) = \text{cone}(t_{\leq n}R\tau_*\mathbb{Z}/p^\mu\mathbb{Z}(n) \to \sigma_*\nu_\mu(n-1)[-n])[-1]$$

as the definition of the complexes $S_\mu(n)$. Put

$$H^*(\mathscr{X}, S_{\mathbb{Z}_p}(n)) := \varprojlim H^*(\mathscr{X}, S_\mu(n)) \quad \text{and}$$
$$H^*(\mathscr{X}, S_{\mathbb{Q}_p}(n)) := H^*(\mathscr{X}, S_{\mathbb{Z}_p}(n)) \otimes \mathbb{Q}_p.$$

CONJECTURE. *For $i \neq 2n, 2n-1$ the natural map $H^{i+1}(\mathscr{X}, S_{\mathbb{Q}_p}(n)) \to H^{i+1}(X, \mathbb{Q}_p(n))$ is injective and its image corresponds to $H^1_f(K, V)$ under the isomorphism $H^{i+1}(X, \mathbb{Q}_p(n)) \cong H^1(K, V)$, i.e.,*

$$\begin{array}{ccc}
H^{i+1}(\mathscr{X}, S_{\mathbb{Q}_p}(n)) & \hookrightarrow & H^{i+1}(X, \mathbb{Q}_p(n)) \\
\updownarrow\cong & & \updownarrow\cong \\
H^1_f(K, V) & \subseteq & H^1(K, V).
\end{array}$$

As explained above this conjecture makes sense independently of the existence of motivic cohomology.

Concerning the evidence we begin by looking at the case $i < n$ and not $i = 0$, $n = 1$; then $H^1_f(K, V) = H^1(K, V)$. It is known that

$$\varprojlim H^*(Y, \nu_\mu(n)) \otimes \mathbb{Q}_p = 0 \quad \text{for } * \neq n, n+1$$

(see **[Mil2, p. 80, first paragraph]**). It follows that the map

$$H^{i+1}(\mathscr{X}, S_{\mathbb{Q}_p}(n)) \to [\varprojlim H^{i+1}(\mathscr{X}, t_{\leq n} R\tau_* \mathbb{Z}/p^\mu\mathbb{Z}(n))] \otimes \mathbb{Q}_p$$

is injective for $i \neq 2n, 2n-1$ and is bijective for $i \neq 2n, 2n-1, 2n-2$. This shows that

$$H^{i+1}(\mathscr{X}, S_{\mathbb{Q}_p}(n)) \to H^{i+1}(X, \mathbb{Q}_p(n))$$

is injective for $i < n$ and is an isomorphism for $i < n$ and not $i = 0$, $n = 1$, so that the conjecture holds in the latter case.

If $X = \mathrm{Spec}(K)$ and $i = 0$, only the case $n = 1$ is not covered by that consideration. But we know that

$$\begin{array}{ccc} o_K^\times \otimes \mathbb{Q} & \subseteq & (\varprojlim K^\times / K^{\times p^\mu}) \otimes \mathbb{Q} \\ \updownarrow \cong & & \updownarrow \cong \\ H^1_f(K, V) & \subseteq & H^1(K, V); \end{array}$$

on the other hand, we have

$$\begin{aligned} H^1(\mathrm{Spec}(o_K), S_{\mathbb{Q}_p}(1)) &= \ker(H^1(K, \mathbb{Q}_p(1)) \xrightarrow{s} H^0(k, \mathbb{Z}_p)) \\ &= \ker((\varprojlim K^\times / K^{\times p^\mu}) \otimes \mathbb{Q}_p \xrightarrow{\text{valuation}} \mathbb{Q}_p) \\ &= o_K^\times \otimes \mathbb{Q}. \end{aligned}$$

Let us now assume that $K/\mathbb{Q}_p$ is unramified and that $n < p$. Then more evidence can be obtained from considering the syntomic sheaves of Fontaine-Messing and Kato. On the syntomic site $\mathscr{X}_{\mathrm{syn}}$ we have the sheaves

$$\begin{aligned} \mathscr{O}_\mu(V) &:= H^0(V_\mu, \mathscr{O}), \\ \mathscr{O}_\mu^{\mathrm{crys}}(V) &:= H^0((V_1 \to \mathrm{Spec}(o_K/p^\mu))_{\mathrm{crys}}, \mathscr{O}^{\mathrm{crys}}), \\ J_\mu &:= \ker(\mathscr{O}_\mu^{\mathrm{crys}} \to \mathscr{O}_\mu), \quad \text{and} \\ J_\mu^{[n]} &:= n\text{th divided power of } J_\mu; \end{aligned}$$

here we have put $V_\mu := V \times_{o_K} (o_K/p^\mu)$ for any syntomic scheme V over o_K. It is shown in **[FM, II.2.3 and III.1.1]** that (for $n < p$) the homomorphism

$$J_\mu^{[n]} \xrightarrow{1 - p^{-n} f} \mathscr{O}_\mu^{\mathrm{crys}}$$

is well defined and surjective. Denoting its kernel by $s_\mu(n)$ we therefore have the exact sequence

$$0 \to s_\mu(n) \to J_\mu^{[n]} \xrightarrow{1-p^{-n}f} \mathcal{O}_\mu^{\text{crys}} \to 0.$$

Put

$$H^*_{\text{syn}}(\mathscr{X}, s_{\mathbb{Q}_p}(n)) := [\varprojlim H^*_{\text{syn}}(\mathscr{X}, s_\mu(n))] \otimes \mathbb{Q}_p.$$

Passing to cohomology then gives, again by **[FM]**, the exact sequence

$$\begin{aligned} 0 \to H^i_{\text{DR}}(X)/(1-p^{-n}f)F^n \to H^{i+1}_{\text{syn}}(\mathscr{X}, s_{\mathbb{Q}_p}(n)) \\ \to \{v \in F^n H^{i+1}_{\text{DR}}(X) \colon fv = p^n v\} \to 0. \end{aligned}$$

From the crystalline Weyl conjecture **[KM]** we know that $1 - p^{-n}f$ is an automorphism of $H^*_{\text{DR}}(X) = H^*_{\text{crys}}(Y/W(k)) \otimes \mathbb{Q}_p$ if $* \neq 2n$. Hence we obtain natural $\mathbb{Q}_p$-linear isomorphisms

$$H^i_{\text{DR}}(X)/F^n \xrightarrow[1-p^{-n}f]{\cong} H^i_{\text{DR}}(X)/(1-p^{-n}f)F^n \xrightarrow{\cong} H^{i+1}_{\text{syn}}(\mathscr{X}, s_{\mathbb{Q}_p}(n))$$

if $i \neq 2n, 2n-1$. Moreover, Faltings' result gives an isomorphism

$$DR(V)/DR(V)^0 \cong H^i_{\text{DR}}(X)/F^n.$$

We see that

$$\dim_{\mathbb{Q}_p} H^{i+1}_{\text{syn}}(\mathscr{X}, s_{\mathbb{Q}_p}(n)) = \dim_{\mathbb{Q}_p} DR(V)/DR(V)^0$$

holds for $n < p$ and $i \neq 2n, 2n-1$. On the other hand, if $\beta \colon \mathscr{X}_{\text{syn}} \to \mathscr{X}_{\text{ét}}$ denotes the natural morphism of sites, then Kurihara has established in **[Kur]** the existence of distinguished triangles

$$\begin{array}{ccc} & \nu_\mu(n-1)[-n] & \\ {}^{+1}\swarrow & & \nwarrow \\ \sigma^* R\beta_* s_\mu(n) & \rightarrow & \sigma^* t_{\leq n} R\tau_* \mathbb{Z}/p^\mu\mathbb{Z}(n) \end{array}$$

provided $n < p - 1$. This means that we have (noncanonical) isomorphisms in the derived category

$$\sigma^* R\beta_* s_\mu(n) \sim \sigma^* S_\mu(n)$$

and therefore by the proper base change theorem isomorphisms on cohomology groups

$$H^*_{\text{syn}}(\mathscr{X}, s_\mu(n)) \cong H^*(\mathscr{X}, S_\mu(n))$$

provided $n < p - 1$. In this way we get the following result in the direction of our conjecture.

PROPOSITION. *If* $K/\mathbb{Q}_p$ *is unramified and if* $n < p-1$, *then*

$$\dim_{\mathbb{Q}_p} H^{i+1}(\mathscr{X}, S_{\mathbb{Q}_p}(n)) = \dim_{\mathbb{Q}_p} H^1_f(K, V) \quad \textit{for } i \neq 2n, 2n-1.$$

Of course, one may hope that the various isomorphisms in the above considerations can be made compatible in such a way that one has the commutative diagram

$$\begin{array}{ccccc} H^{i+1}_{\text{syn}}(\mathscr{X}, s_{\mathbb{Q}_p}(n)) & \xrightarrow{\cong} & H^{i+1}(\mathscr{X}, S_{\mathbb{Q}_p}(n)) & \hookrightarrow & H^{i+1}(X, \mathbb{Q}_p(n)) \\ \updownarrow \cong & & & & \\ H^i_{\text{DR}}(X)/(1-p^{-n}f)F^n & & & & \updownarrow \cong \\ \updownarrow \cong & & & & \\ DR(V)/DR(V)^0 & \underset{\exp}{\xrightarrow{\cong}} & H^1_f(K, V) & \xrightarrow{\subseteq} & H^1(K, V), \end{array}$$

thereby fully proving the conjecture in this case.

8. Motivic Tate modules

The next step would be to understand the map

$$H^{i+1}(\mathscr{X}, \mathbb{Z}(n)) \to H^{i+1}(\mathscr{X}, S_{\mathbb{Z}_p}(n)).$$

This seems not to be possible at present. Nevertheless we want to propagate the point of view that

$$\mathscr{G}^{i+1}_n(L) := \ker(H^{i+1}(\mathscr{X} \times o_L, \mathbb{Z}(n)) \to H^{i+1}(\mathscr{X} \times k_L, \mathbb{Z}(n)))$$

for L varying through the finite extensions of K is a kind of generalization of the formal group of an abelian variety; here o_L, resp. k_L, denotes the ring of integers in L, resp. the residue class field of o_L. (Possibly $H^{i+1}(\mathscr{X} \times_{o_L}, \mathbb{Z}(n))$ contains a large uniquely divisible subgroup which one may want to divide out.) In this light our discussion so far has dealt with the "tangent space"

$$L \mapsto H^i_{\text{DR}}(X_{/L})/F^n$$

and, if $i \neq 2n, 2n-1$, the conjectural "logarithm map"

$$\mathscr{G}^{i+1}_n(L) \to H^{i+1}(\mathscr{X} \times o_L, S_{\mathbb{Q}_p}(n)) \cong H^1_f(L, V) \cong H^i_{\text{DR}}(X_{/L})/F^n$$

into the tangent space. A third natural object to consider is the Tate module

$$\mathrm{Tate}(\mathscr{G}_n^{i+1}) := \left[\varprojlim_{\mu} \varinjlim_{L} \mathscr{G}_n^{i+1}(L)_{p^\mu}\right] \otimes \mathbb{Q}_p .$$

We can say something about it under the assumption that the reduction Y is Hodge-Witt (i.e, the slope spectral sequence for Y degenerates at E_1). Put $\overline{\mathscr{X}} := \mathscr{X} \times o_{\overline{K}}$ and let $\overline{\tau}\colon \overline{X} \hookrightarrow \overline{\mathscr{X}}$ denote the corresponding open immersion.

LEMMA.

$$\varinjlim_{L} H^*(\mathscr{X} \times o_L, S_\mu(n)) = H^*(\overline{\mathscr{X}}, t_{\leq n} R\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n)).$$

PROOF. For any finite extension L/K let $\tau_L\colon X_{/L} \hookrightarrow \mathscr{X} \times o_L$ be the natural open immersion. Now fix finite extensions $K \subseteq L \subseteq L'$ and let e denote the ramification index of the extension L'/L. It follows from the explicit description in [**BKØ**, (6.6.1)] of the homomorphism s which defines the cone $S_\mu(n)$ that the following diagram is commutative (and exact):

$$\begin{array}{ccc}
\cdots & & \cdots \\
\downarrow & & \downarrow \\
H^{*-n-1}(Y \times k_L, \nu_\mu(n-1)) & \xrightarrow{e\cdot\mathrm{can}} & H^{*-n-1}(Y \times k_{L'}, \nu_\mu(n-1)) \\
\downarrow & & \downarrow \\
H^*(\mathscr{X} \times o_L, S_\mu(n)) & \xrightarrow{\mathrm{can}} & H^*(\mathscr{X} \times o_{L'}, S_\mu(n)) \\
\downarrow & & \downarrow \\
H^*(\mathscr{X} \times o_L, t_{\leq n} R\tau_{L*}\mathbb{Z}/p^\mu\mathbb{Z}(n)) & \xrightarrow{\mathrm{can}} & H^*(\mathscr{X} \times o_{L'}, t_{\leq n} R\tau_{L'*}\mathbb{Z}/p^\mu\mathbb{Z}(n)) \\
\downarrow & & \downarrow \\
H^{*-n}(Y \times k_L, \nu_\mu(n-1)) & \xrightarrow{e\cdot\mathrm{can}} & H^{*-n}(Y \times k_{L'}, \nu_\mu(n-1)) . \\
\downarrow & & \downarrow \\
\cdots & & \cdots
\end{array}$$

This implies that in the limit we have

$$\varinjlim_{L} H^*(\mathscr{X} \times o_L, S_\mu(n)) = \varinjlim_{L} H^*(\mathscr{X} \times o_L, t_{\leq n} R\tau_{L*}\mathbb{Z}/p^\mu\mathbb{Z}(n)).$$

For formal reasons the right-hand side is equal to $H^*(\overline{\mathscr{X}}, t_{\leq n} R\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n))$.

Consider now the exact sequences

$$\begin{array}{c} 0 \\ \downarrow \\ \varinjlim_L H^i(\mathscr{X} \times o_L, \mathbb{Z}(n))/p^\mu \\ \downarrow \\ \varinjlim_L H^i(\mathscr{X} \times o_L, S_\mu(n)) = H^i(\overline{\mathscr{X}}, t_{\leq n} R\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n)) \\ \downarrow \\ \varinjlim_L H^{i+1}(\mathscr{X} \times o_L, \mathbb{Z}(n))_{p^\mu} \\ \downarrow \\ 0. \end{array}$$

Viewed as a short exact sequence of projective systems with respect to μ, we see that the upper term has surjective transition maps. In the projective limit, therefore, we still have a surjection

$$\varprojlim_\mu H^i(\overline{\mathscr{X}}, t_{\leq n} R\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n)) \to \varprojlim_\mu \varinjlim_L H^{i+1}(\mathscr{X} \times o_L, \mathbb{Z}(n))_{p^\mu} \to 0.$$

To be able to say more, we now assume the following conjecture of Bloch in [**Blo**, §3] to hold:

If Y is Hodge-Witt, then the spectral sequences

(B1) $$E_2^{a,b} = H^a(\overline{\mathscr{X}}, R^b\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n)) \Rightarrow H^{a+b}(\overline{X}, \mathbb{Z}/p^\mu\mathbb{Z}(n))$$

degenerate at least up to torsion which is bounded independently of μ.

This conjecture was proved by Kato [**Kat**] under the additional assumptions that $K/\mathbb{Q}_p$ is unramified and that $\dim X < p-1$. Let $F^{\cdot}H^*(\overline{X}, \mathbb{Z}/p^\mu\mathbb{Z}(n))$ denote the filtration induced by the above spectral sequence, i.e.,

$$F^{*-b}H^*(\cdots) = \operatorname{im}(H^*(\overline{\mathscr{X}}, t_{\leq b} R\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n)) \to H^*(\overline{X}, \mathbb{Z}/p^\mu\mathbb{Z}(n)))$$
$$\text{and} \quad F^{\cdot}H^*(\cdots)/F^{\cdot+1}H^*(\cdots) = E_\infty^{\cdot, *-\cdot}.$$

It induces corresponding filtrations on $H^*(\overline{X}, \mathbb{Z}_p(n))$ and $H^*(\overline{X}, \mathbb{Q}_p(n))$. Assuming from now on that Y is Hodge-Witt, (B1) implies that

$$\varprojlim_\mu H^i(\overline{\mathscr{X}}, t_{\leq n} R\overline{\tau}_* \mathbb{Z}/p^\mu\mathbb{Z}(n)) \otimes \mathbb{Q}_p = F^{i-n}H^i(\overline{X}, \mathbb{Q}_p(n)) = F^{i-n}V.$$

Consider the commutative diagram

$$\begin{array}{ccc}
 & & 0 \\
 & & \downarrow \\
 & & H^i(\overline{X}, \mathbb{Z}(n))/p^\mu \\
 & & \downarrow \\
\varinjlim_L H^i(\mathscr{X} \times o_L, S_\mu(n)) & \longrightarrow & H^i(\overline{X}, \mathbb{Z}/p^\mu\mathbb{Z}(n)) \\
\downarrow & & \downarrow \\
\varinjlim_L H^{i+1}(\mathscr{X} \times o_L, \mathbb{Z}(n))_{p^\mu} & \longrightarrow & H^{i+1}(\overline{X}, \mathbb{Z}(n))_{p^\mu} \\
\downarrow & & \downarrow \\
0 & & 0.
\end{array}$$

By (B1) and the Lemma the kernels of the horizontal maps are bounded independently of μ. The same is true in case $i \neq 2n$ for the first term in the right column by Claim 2. Therefore, passing to the projective limit with respect to μ and tensoring with $\mathbb{Q}_p$ lead, for $i \neq 2n$, to the commutative diagram

$$\begin{array}{ccc}
\left[\varprojlim_\mu \varinjlim_L H^i(\mathscr{X} \times o_L, S_\mu(n))\right] \otimes \mathbb{Q}_p & \xrightarrow{\cong} & \left[\varprojlim_\mu \varinjlim_L H^{i+1}(\mathscr{X} \times o_L, \mathbb{Z}(n))_{p^\mu}\right] \otimes \mathbb{Q}_p \\
\downarrow \cong & & \\
F^{i-n} H^i(\overline{X}, \mathbb{Q}_p(n)) & & \downarrow \\
\downarrow \subseteq & & \\
H^i(\overline{X}, \mathbb{Q}_p(n)) & \xrightarrow{\cong} & \left[\varprojlim_\mu H^{i+1}(\overline{X}, \mathbb{Z}(n))_{p^\mu}\right] \otimes \mathbb{Q}_p.
\end{array}$$

On the other hand, from Mot 9 we have the exact sequences

$$\begin{aligned}0 \to H^i(\mathscr{X} \times k_L, \mathbb{Z}(n))/p^\mu \to H^{i-n}(\mathscr{X} \times k_L, \nu_\mu(n)) \\ \to H^{i+1}(\mathscr{X} \times k_L, \mathbb{Z}(n))_{p^\mu} \to 0.\end{aligned}$$

In the limit with respect to L they give (by Mot 2) the exact sequences

$$0 \to H^i(\overline{Y}, \mathbb{Z}(n))/p^\mu \to H^{i-n}(\overline{Y}, \nu_\mu(n)) \to H^{i+1}(\overline{Y}, \mathbb{Z}(n))_{p^\mu} \to 0,$$

where $\overline{Y} := \mathscr{X} \times \overline{k} = Y \times \overline{k}$ with $\overline{k}$ the residue class field of $o_{\overline{K}}$.

CLAIM 7. *For $i \neq 2n$ the group $H^i(\overline{Y}, \mathbb{Z}(n))$ contains a p-divisible subgroup such that the quotient has finite exponent.*

PROOF. This follows by the same arguments as in the proofs of Claims 1 and 2 from the subsequent facts which can be found in **[Mil1,** Proposition 3.1 and Corollary 6.4].

- The torsion subgroup in $H^{i-n}(\overline{Y}, \nu_{\mathbb{Z}_p}(n)) := \varprojlim_{\mu} H^{i-n}(\overline{Y}, \nu_\mu(n))$ has a finite exponent.
- That part of $H^{i-n}(\overline{Y}, \nu_{\mathbb{Z}_p}(n))$ on which $\mathrm{Gal}(\overline{k}/k)$ acts discretely is torsion provided $i \neq 2n$.

We therefore obtain, for $i \neq 2n$,

$$\left[\varprojlim_{\mu} H^{i+1}(\overline{Y}, \mathbb{Z}(n))_{p^\mu}\right] \otimes \mathbb{Q}_p = \varprojlim_{\mu} H^{i-n}(\overline{Y}, \nu_\mu(n)) \otimes \mathbb{Q}_p =: H^{i-n}(\overline{Y}, \nu_{\mathbb{Q}_p}(n)).$$

Putting these results together, we arrive at the following computation of the Tate module.

CLAIM 8. *If Y is Hodge-Witt and* (B1) *holds, then we have, for $i \neq 2n$,*

$$\mathrm{Tate}(\mathcal{G}_n^{i+1}) = \ker(F^{i-n}H^i(\overline{X}, \mathbb{Q}_p(n)) \to H^{i-n}(\overline{Y}, \nu_{\mathbb{Q}_p}(n))).$$

Of course, the map

$$F^{i-n}H^i(\overline{X}, \mathbb{Q}_p(n)) \to H^{i-n}(\overline{Y}, \nu_{\mathbb{Q}_p}(n))$$

is induced by the natural sheaf homomorphisms

$$\begin{array}{ccc} \mathbb{Z}(n)_{\mathscr{X}\times o_L} \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z} & \longrightarrow & \sigma_{L*}\mathbb{Z}(n)_{Y\times k_L} \overset{\mathbb{L}}{\otimes} \mathbb{Z}/p^\mu\mathbb{Z} \\ \| & & \| \\ S_\mu(n)_{\mathscr{X}\times o_L} & \longrightarrow & \sigma_{L*}\nu_\mu(n)_{Y\times k_L}[-n], \end{array}$$

where $\sigma_L : Y \times k_L \hookrightarrow \mathscr{X} \times o_L$ denotes the obvious closed immersion (the vertical identification on the right-hand side comes from Mot 9). A similar analysis as in §7 based again on **[BKØ,** (6.6)] shows that the lower homomorphism can be defined independently of the existence of motivic cohomology. Following **[Blo]** we set

$$U^{i+1-n}H^i(\overline{X}, \mathbb{Q}_p(n)) := \ker(F^{i-n}H^i(\overline{X}, \mathbb{Q}_p(n)) \to H^{i-n}(\overline{Y}, \nu_{\mathbb{Q}_p}(n))).$$

Always assuming (B1) to hold it is easy to see that

$$F^{i+1-n}H^i(\overline{X}, \mathbb{Q}_p(n)) \subseteq U^{i+1-n}H^i(\overline{X}, \mathbb{Q}_p(n)).$$

Since over $\overline{K}$ the Tate twist commutes with forming cohomology, this in particular means that

$$[U^{i+1-(n-1)}H^i(\overline{X}, \mathbb{Q}_p(n-1))](1) \subseteq U^{i+1-n}H^i(\overline{X}, \mathbb{Q}_p(n)).$$

If $i \neq 2n, 2n-2$, the latter can be reformulated by Claim 8 as

$$\mathrm{Tate}(\mathscr{G}_{n-1}^{i+1})(1) \subseteq \mathrm{Tate}(\mathscr{G}_n^{i+1}).$$

Bloch in **[Blo,** §3] also conjectures the following:

If Y is Hodge-Witt, then the functor

$$L \mapsto \ker(H_{\mathrm{Zar}}^{i+1-n}(\mathscr{X} \times o_L, \underline{\underline{K}}_n^{\mathrm{Milnor}}) \to H_{\mathrm{Zar}}^{i+1-n}(Y \times k_L, \underline{\underline{K}}_n^{\mathrm{Milnor}})) \quad \text{(B2)}$$

has a "piece" $L \mapsto \mathscr{B}_n^{i+1}(L)$ which is a p-divisible formal group over o_K with Tate module

$$\mathrm{Tate}(\mathscr{B}_n^{i+1}) = U^{i+1-n} H^i(\overline{X}, \mathbb{Q}_p(n)) / U^{i+1-(n-1)} H^i(\overline{X}, \mathbb{Q}_p(n-1))(1).$$

In case $i \neq 2n$ we then would have the surjection

$$\mathrm{Tate}(\mathscr{G}_n^{i+1}) \twoheadrightarrow \mathrm{Tate}(\mathscr{B}_n^{i+1});$$

does it come from a homomorphism $\mathscr{G}_n^{i+1} \to \mathscr{B}_n^{i+1}$?

From Mot 10 we obtain

$$H_{\mathrm{Zar}}^{i+1-n}(\mathscr{X} \times o_L, \underline{\underline{K}}_n^{\mathrm{Milnor}}) = H_{\mathrm{Zar}}^{i+1-n}(\mathscr{X} \times o_L, R^n \alpha_* \mathbb{Z}(n)).$$

Since the right-hand side, by the Leray spectral sequence for α_*, is closely related to $H^{i+1}(\mathscr{X} \times o_L, \mathbb{Z}(n))$, this conjecture (B2) seems to fit well into our general point of view.

If $\mathscr{G}_n^{i+1}$ is some kind of generalized formal group then it should have a "Dieudonné module". This means understanding the functor

$$A \mapsto \mathscr{G}_n^{i+1}(A) := \ker(H^{i+1}(\mathscr{X} \times A, \mathbb{Z}(n)) \to H^{i+1}(\mathscr{X} \times A_{\mathrm{red}}, \mathbb{Z}(n)))$$

on Artinian o_K-algebras A. Writing down this functor requires the existence of the complexes $\mathbb{Z}(n)_U$ at least for any scheme U such that the associated scheme U_{red} is Noetherian and regular. But so far all proposals for the complexes $\mathbb{Z}(n)$ have been insensitive to nilpotent elements.

References

[Blo] S. Bloch, *p-adic étale cohomology*, Arithmetic and Geometry, vol. I, Birkhäuser, Boston, Basel, and Stuttgart, 1983, pp. 13–26.

[BKØ] S. Bloch and K. Kato, *p-adic étale cohomology*, Inst. Hautes Études Sci. Publ. Math., no. 63 (1986), 107–152.

[BK] ——, *L-functions and Tamagawa numbers of motives*, The Grothendieck Festschrift, vol. I, Birkhäuser, Boston, Basel, and Berlin, 1990, pp. 333–400.

[Fal] G. Faltings, *Crystalline cohomology and p-adic étale cohomology*, Algebraic Analysis, Geometry and Number Theory, Johns Hopkins Univ. Press, 1989, pp. 25–80.

[Fon] J.-M. Fontaine, *Sur certains types de représentations p-adiques du groupe de Galois d'un corps local; construction d'un anneau de Barsotti-Tate*, Ann. of Math. **115** (1982), 529–577.

[FM] J.-M. Fontaine and W. Messing, *p-adic periods and p-adic étale cohomology*, Current Trends in Arithmetical Algebraic Geometry, Contemp. Math., vol. 67, Amer. Math. Soc., Providence, RI, 1987, pp. 179–207.

[Gro] A. Grothendieck, *Le groupe de Brauer*, II. Dix Exposés sur la Cohomologie des Schémas, North-Holland, Amsterdam, 1968, pp. 67–87.

[Kat] K. Kato, *On p-adic vanishing cycles* (*Application of ideas of Fontaine-Messing*), Algebraic Geometry, Sendai, 1985, Adv. Stud. Pure Math., vol. 10, North-Holland, Amsterdam and New York, 1987, pp. 207–251.

[KM] N. M. Katz and W. Messing, *Some consequences of the Riemann hypothesis for varieties over finite fields*, Invent. Math. **23** (1974), 73–77.

[Kur] M. Kurihara, *A note on p-adic étale cohomology*, Proc. Japan Acad. Ser. A Math. Sci. **63** (1987), 275–278.

[Lic1] S. Lichtenbaum, *Values of zeta-functions at non-negative integers*, Number Theory (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer-Verlag, Berlin, Heidelberg, and New York, 1984, pp. 127–138.

[Lic2] ——, *New results on weight-two motivic cohomology*, The Grothendieck Festschrift, vol. III, Birkhäuser, Boston, Basel, and Berlin, 1990, pp. 35–55.

[Mil1] J. S. Milne, *Values of zeta functions of varieties over finite fields*, Amer. J. Math. **108** (1986), 297–360.

[Mil2] ——, *Motivic cohomology and values of zeta functions*, Compositio Math. **68** (1988), 59–102.

[Tat] J. Tate, *p-divisible groups*, Proc. Conf. Local Fields (Driebergen, 1966), Springer-Verlag, Berlin, Heidelberg, and New York, 1967, pp. 158–183.

Mathematisches Institut der Universität zu Köln, Germany

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Admissible Non-Archimedean Standard Zeta Functions Associated with Siegel Modular Forms

A. A. PANCHISHKIN

0. Introduction

Let p be a prime number. In this paper we describe p-adic properties of the special values of the standard zeta function $\mathscr{D}(s, f, \chi)$, where f is a Siegel cusp form of an even degree m and of weight $k > 2m+2$ and χ is a varying Dirichlet character. In our previous work **[Pa14]** we constructed the non-Archimedean interpolation of these special values under the important assumption that the form f is p-ordinary, i.e., $|i_p(\alpha_0(p))|_p = 1$, where $i_p(\alpha_0(p))$ denotes the image in the Tate field $\mathbf{C}_p$ under a fixed embedding $i_p : \overline{\mathbf{Q}} \to \mathbf{C}_p$ of one of the Satake p-parameters $\alpha_0(p)$ of the cusp eigenform f. Let S be a finite set of primes containing p. It follows from the main theorem in **[Pa14]** that the normalized values $\mathscr{D}(s, f, \chi)$ can be explicitly rewritten in terms of some $\mathbf{C}_p$-valued integrals of bounded p-adic measures (over the group $\mathbf{Z}_S^\times$ of S-adic units) provided the character χ has an S-complete conductor, i.e., χ is nontrivial for $S = \{p\}$.

This paper has some further purposes: first we treat the delicate case when the above p-ordinarity condition is not satisfied, and we give an interpretation of the quantity $\operatorname{ord}(\alpha_0(p)^2)$ in terms of the Newton polygon and the Hodge polygon, attached to a conjectural motive associated with the zeta function $\mathscr{D}(s, f, \chi)$; on the other hand we use this motivic interpretation in order to precisely evaluate the corresponding p-adic integrals in the general case (e.g., for the trivial character χ) using a version of the conjecture on p-adic L-functions attached to motives over $\mathbf{Q}$ **[Co-PeRi]**. The interesting fact is that the predicted p-adic identity can be independently proved using the properties of the Hecke algebra acting on Siegel modular forms. Also, it would be interesting to extend these results to Siegel cusp eigenforms of odd degree, using the method of Böcherer **[Bö1]**.

1991 *Mathematics Subject Classification*. Primary 11F03, 11F46, 11F85, 11S40.
This paper is in final form and no version of it will be submitted for publication elsewhere.

Content of the paper. In §1 we recall some properties and definitions of motives over $\mathbf{Q}$ and their L–functions. Then in §2 we describe the definition of the standard zeta functions $\mathscr{D}(s, f, \chi)$, where f is a Siegel cusp form of degree m and of weight k, and give an interpretation of this zeta-function in terms of a motive $M = M_f$ over $\mathbf{Q}$ provided that the cusp form f does not come from a smaller algebraic subgroup, e.g., does not belong to the Maass subspace (in the case of degree two). The generalized Maass subspace should be defined as the space of modular forms, which are liftings from the "smaller" groups; then there is a version of the generalized Ramanujan–Petersson conjecture for all cusp forms f that do not belong to the generalized Maass subspace. Then we recall in §3 the definition by Coates of the modified L–function of a motive M over F and we state there a modified period conjecture which gives a description of the critical special values of arbitrary twists $M(\chi)$ with Dirichlet characters χ in terms of Deligne's periods of the original motive M. Also we show that Deligne's conjecture perfectly matches with the known results on the algebraic properties of the above special values [**St2, Ha1, Bö1**]. In §4 we recall the notion of an h-admissible measure over a Galois group and properties of its Mellin transform. Then in §5 we recall the definitions of the Newton polygon and the Hodge polygon of a motive over $\mathbf{Q}$, and then in §6 we formulate a general conjecture on p-adic L-functions of a critical motive M over $\mathbf{Q}$ without assuming that M is p-ordinary. This conjecture is stated in terms of the existence of some h-admissible measures which provide the p-adic L-functions using the p-adic Mellin transform. Then in §7 we specialize the conjecture to the case of the conjectural motive M_f, and we formulate a theorem on the existence of h-admissible measures attached to the standard zeta function $\mathscr{D}(s, f, \chi)$ of a Siegel cusp eigenform f. In §8 we outline the method of the proof of this result on the existence of the above h-admissible measures using congruences for the Fourier coefficients of Siegel modular forms.

The Appendix was included following a suggestion of U. Jannsen, and it is devoted to p-adic analytic families of Galois representations associated with motives. These families are bigger that those obtained from a fixed motive by cyclotomic twist, and they go back to those constructed by Hida in the case of elliptic modular forms. We formulate two general conjectures on the existence of such families, and on p-adic analytic dependence of the corresponding p-adic L-functions on a parameter in such a family.

Throughout the paper we fix embeddings

$$i_\infty : \overline{\mathbf{Q}} \to \mathbf{C}, \quad i_p : \overline{\mathbf{Q}} \to \mathbf{C}_p,$$

and we shall often regard algebraic numbers (via these embeddings) as both complex and p-adic numbers, where $\mathbf{C}_p = \widehat{\overline{\mathbf{Q}}}_p$ is the Tate field (the completion of a fixed algebraic closure $\overline{\mathbf{Q}}_p$ of $\mathbf{Q}_p$), which is endowed with a unique norm $|\cdot|_p$ such that $|p|_p = p^{-1}$.

1. Motives over Q and their L-functions

By a motive M over $\mathbf{Q}$ with coefficients in a number field T we shall mean a collection of objects M_B, M_{DR}, M_λ, I_∞, I_λ, where

M_B is the Betti realization of M which is a vector space over T of dimension d endowed with a T-rational involution ρ;

M_{DR} is the de Rham realization of M, a vector space over T of dimension d, endowed with a decreasing filtration $\{F^i_{\mathrm{DR}}(M) \subset M_{\mathrm{DR}} \mid i \in \mathbf{Z}\}$ of T-vector spaces;

M_λ is the λ-adic realization of M at a finite place λ of the coefficient field T (a T_λ-vector space of degree d over T_λ, a completion of T at λ), which is a Galois module over $G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ so that we have a compatible system of λ-adic representations denoted by

$$r_{M,\lambda} = r_\lambda : G_{\mathbf{Q}} \to \mathrm{GL}(M_\lambda).$$

Also,

$$I_\infty : M \otimes_{\mathbf{Q}} \mathbf{C} \to M_{\mathrm{DR}} \otimes_{\mathbf{Q}} \mathbf{C}$$

is the complex comparison isomorphism of $(T \otimes \mathbf{C})$-modules, and

$$I_\lambda : M \otimes_T T_\lambda \to M_\lambda$$

is the λ-adic comparison isomorphism of T_λ-vector spaces. It is assumed in the notation that the complex vector space $M \otimes_{\mathbf{Q}} \mathbf{C}$ is decomposed in the Hodge bigraduation of complex subspaces

$$M \otimes_{\mathbf{Q}} \mathbf{C} = \bigoplus_{i,j} M^{i,j}$$

in which $\rho(M^{i,j}) \subset M^{j,i}$ and $h(i,j) = h(i,j,M) = \dim_{\mathbf{C}} M^{i,j}$ are the Hodge numbers. Moreover,

$$I_\infty\left(\bigoplus_{i' \geq i} M^{i',j}\right) = F^i_{\mathrm{DR}}(M) \otimes_{\mathbf{Q}} \mathbf{C}.$$

Also, I_λ takes ρ to the r_λ-image of the Galois automorphism, which is denoted by the same symbol $\rho \in G_{\mathbf{Q}}$ and corresponds to the complex conjugation of $\mathbf{C}$ under the fixed embedding i_∞ of $\overline{\mathbf{Q}}$ to $\mathbf{C}$. We assume that M is pure of weight w (i.e., $i + j = w$).

The L-function $L(M, s)$ of M is defined as the following Euler product (which takes values in $T \otimes \mathbf{C}$):

$$L(M, s) = \prod_p L_p(M, p^{-s}),$$

extended over all primes p and where

$$\begin{aligned} L_p(M, X)^{-1} &= \det(1 - X \cdot r_\lambda(\mathrm{Fr}_p^{-1}) \mid M_\lambda^{I_p}) \\ &= (1 - \alpha^{(1)}(p)X) \cdot (1 - \alpha^{(2)}(p)X) \cdot \dots \cdot (1 - \alpha^{(d)}(p)X) \\ &= 1 + A_1(p)X + \dots + A_d(p)X^d ; \end{aligned}$$

here $\mathrm{Fr}_p \in G_{\mathbf{Q}}$ is the Frobenius element at p, defined modulo conjugation and modulo the inertia subgroup $I_p \subset G_p \subset G_{\mathbf{Q}}$ of the decomposition group G_p (with respect to an arbitrary extension of the place p to $\overline{\mathbf{Q}}$). We make the standard hypothesis that the coefficients of $L_p(M, X)^{-1}$ belong to T, do not depend on λ not dividing p, and we regard this polynomial over the ring $T \otimes \mathbf{C}$ so that

$$L_p(M, X)^{-1} = (L_p^{(\tau)}(M, X)^{-1})_{\tau \in J_T} = (1 + A_1(p)^\tau X + \dots + A_d(p)^\tau X^d)_\tau ,$$

where J_T denotes the set of all complex embeddings of T.

We shall need the following linear algebra operations on motives. These operations are defined by means of their realizations and can be obviously described in terms of the corresponding L-functions:

$M^\vee$ (dual motive): its λ-adic representation is contragredient to that of M;

$M_1 \oplus M_2$ (direct sum of motives M_1, M_2): its λ-adic representation is the direct sum of those for M_1, M_2, and the corresponding L-function is the product of those of M_1 and M_2;

$M_1 \otimes_{\mathbf{Q}} M_2$ (tensor product of motives over $\mathbf{Q}$), its λ-adic representation is the tensor product of those for M_1, M_2, and the corresponding L-function is a kind of a "multiplicative convolution" of the L-functions of M_1 and M_2.

The important examples of motives are: the cyclotomic (Tate) motive $\mathbf{Q}(1)$ and the motive $[\chi]$ associated with a Dirichlet character $\chi : \mathbf{A}_{\mathbf{Q}}^\times/\mathbf{Q}^\times \to \mathbf{C}^\times$. The λ-adic representation of $\mathbf{Q}(1)$ is defined by the action of $G_{\mathbf{Q}}$ on the ℓ-power roots of unity (where ℓ is the characteristic of λ) so that Fr_p acts as a scalar p and $L(\mathbf{Q}(1), s) = \zeta(s+1)$, $\zeta(s)$ being the Riemann zeta function. Also $\mathbf{Q}(s)$ will denote the sth tensor power of $\mathbf{Q}(1)$ if $s \geq 0$ and the $-s$th tensor power of $\mathbf{Q}(-1) = \mathbf{Q}(1)^\vee$ if $s < 0$.

The λ-adic representation of $[\chi]$ is given by class field theory so that $L([\chi], s)$ coincides with the Dirichlet L-function $L(s, \chi)$ of χ.

The twist operation: for an arbitrary motive M over $\mathbf{Q}$ with coefficients in T an integer s and a Dirichlet character χ one can define the twist $N = M(s)(\chi)$ which is again a motive over $\mathbf{Q}$ with the coefficient field $T(\chi)$ of the same rank d and weight w so that we have

$$L(N, s) = \prod_p L_p(M, \chi(p)p^{-s-n}).$$

Conjecturally the function $L(M, s)$ can be analytically continued onto the entire complex plane and it satisfies the functional equation of the type

$$\Lambda(M, s) = \varepsilon(M, s)\Lambda(M^{\vee}(1), -s)$$

where $\Lambda(M, s) = L_{\infty}(M, s)L(M, s)$, $L_{\infty}(M, s)$ is the Γ-factor, which is completely determined by the Hodge structure of M, and $\varepsilon(M, s)$ is a certain ε-factor, which can be decomposed into a product of local factors $\varepsilon_v(M, s)$ (v runs over places of $\mathbf{Q}$). Moreover, $\Lambda(M, s)$ is entire, unless the weight w is even and $\mathbf{Q}(-w/2)$ is a direct summand of M.

We shall use the explicit expression for $\Lambda_{\infty}(M, s)$ and $\varepsilon_{\infty}(M, s)$ in terms of the Hodge structure of M (see **[De3**, p. 329]) and of the functions $\Gamma_{\mathbf{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$, $\Gamma_{\mathbf{R}}(s) = \pi^{-s/2}\Gamma(s/2)$. Conversely, if we have a zeta function $Z(s)$ satisfying a functional equation of the above type, and if we know that $Z(s) = L(M, s)$ for a motive M then we can use this functional equation in order to determine the type of the Hodge structure of this motive. As an example we consider in the next section the standard zeta functions of Siegel cusp eigenforms.

2. The standard zeta function of a Siegel cusp eigenform and its motivic interpretation

Let

$$f = \sum_{\xi>0} a(\xi)e_m(\xi z)$$

be a Siegel cusp form of even degree m and of weight k on the congruence subgroup

$$\Gamma_0(C) = \left\{ \begin{pmatrix} ab \\ cd \end{pmatrix} \in \mathrm{Sp}_m(\mathbf{Z}) \mid c \equiv 0 \bmod C \right\}$$

with a Dirichlet character $\psi \bmod C$. Here z belongs to the Siegel upper half-plane

$$\mathscr{H}_m = \left\{ z \in \mathrm{GL}_m(\mathbf{C}) \mid z = {}^t z, \, \mathrm{Im}(z) > 0 \right\}$$

and we adopt the standard notation $e_m(z) = \exp(2\pi i \,\mathrm{Tr}(z))$.

Suppose that f is an eigenfunction of the global Hecke algebra

$$L^{(m)}(C) = \bigotimes_{q \nmid C} L_{(q)}^{(m)}(C)$$

with the eigenvalue given by a homomorphism $\Lambda : L^{(m)}(C) \to \mathbf{C}$, i.e.,

$$f \mid X = \Lambda(X)f \quad \text{for all } X \in L_{(q)}^{(m)}(C).$$

Let $\alpha_0(q), \alpha_1(q), \dots, \alpha_m(q)$ be the $(m+1)$-tuple of Satake q-parameters of Λ, which uniquely determine Λ. It is known **[An3]** that the relation

$$\alpha_0^2(q)\alpha_1(q)\cdots\alpha_m(q) = \psi(q)^m q^{m(k-(m+1)/2)}$$

holds. Recall that the standard zeta function of f with a Dirichlet character χ mod N is defined as the Euler product

$$D(s, f, \chi) = \prod_{q \nmid C} D^{(q)}(s, f, \chi),$$

with

$$D^{(q)}(s, f, \chi)^{-1} = \left(1 - \frac{\chi(q)\psi(q)}{q^s}\right) \prod_{i=1}^{m} \left(1 - \frac{\chi(q)\psi(q)\alpha_i(q)}{q^s}\right) \left(1 - \frac{\chi(q)\psi(q)\alpha_i^{-1}(q)}{q^s}\right),$$

the product being absolutely convergent for $\mathrm{Re}(s) > 1 + m$.

Together with $D(s, f, \chi)$ let us consider the normalized zeta function

$$D^*(s, f, \chi) = \Gamma_{\mathbf{R}}(s + \delta) \prod_{j=1}^{m} \Gamma_{\mathbf{C}}(s + k - j) D(s, f, \chi)$$

where $\delta = 0$ or 1 according as $\psi\chi(-1) = 1$ or -1. The theorem on analytic properties of the standard zeta functions, proved by Andrianov, Kalinin [**An-K**], and Böcherer [**Bö1**] (see also [**Pa14,** Chapter 2]) states that for a Dirichlet character χ modulo a positive integer N (not necessarily primitive) and a Siegel cusp form f of weight $k > m + \nu$ (with $\nu = 0, 1$ and $\chi(-1) = (-1)^\nu$) the function $D^*(s, f, \chi)$ admits an analytic continuation that is holomorphic for all $s \in \mathbf{C}$ with the possible exclusion of a simple pole at the point $s = 1$ in the case when the character $\chi^2\psi^2$ is trivial. In general this result is established using a detailed study of poles and residues of the Siegel-Eisenstein series as functions of the variable s (see [**Fe, Shi8**]) under a technical condition that $C \det(2\xi_0)$ divides N for some matrix ξ_0 such that $a(\xi_0) \neq 0$. If we compare this result with the general conjectural description of the analytic properties of the L-function $L(M, s)$ of a motive M we may expect that the function $D(s, f, \chi)$ coincides with a function of the type $L(M_f(\chi), s)$ such that $D^*(s, f, \chi) = \Lambda(M_f(\chi), s)$, i.e.,

$$L_\infty(M_f(\chi), s) = \Gamma_{\mathbf{R}}(s + \delta) \prod_{j=1}^{m} \Gamma_{\mathbf{C}}(s + k - j),$$

where $\Gamma_{\mathbf{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$, $\Gamma_{\mathbf{R}}(s) = \pi^{-s/2}\Gamma(s/2)$. This means that M_f may have rank $2m + 1$, weight 0, and Hodge structure of type

$$(-k + 1, 1 - k) + (-k + 2, k - 2) + \cdots + (-k + m, k - m) + (0, 0) + (k - m, -k + m) + \cdots + (k - 1, -k + 1).$$

Moreover, in this case all of the Satake p-parameters $\alpha_1(p), \dots, \alpha_m(p)$ for p not dividing the level C of the form f should have the absolute value 1 because this property holds for the $(2m + 1)$ characteristic numbers

$$\psi(p)\alpha_1(p), \dots, \psi(p)\alpha_m(p), \ \psi(p), \ \psi(p)\alpha_m^{-1}(p), \dots, \psi(p)\alpha_1^{-1}(p)$$

of the corresponding pure motive M_f of weight 0. However, the last condition is not satisfied in general even if we restrict ourselves to cusp forms: in the spaces of Siegel cusp forms there are forms lifted from "smaller" algebraic subgroups, for example, cusp forms in the so-called Maass subspace in the case $m = 2$. For the lifted cusp forms this condition is not satisfied for some obvious reason. An example of another type of cusp form, which does not satisfy the generalized Ramanujan–Petersson conjecture, and is not a lifting of a Siegel cusp form of a lower degree, is given by Freitag: this is a certain cusp form of degree 24 and weight 13 constructed from a Leech lattice. This cusp form is not a lifting of a Siegel cusp form of degree < 24, but it is a lifting from an automorphic form on the orthogonal group $\mathrm{O}(T)$, where T is a definite quadratic form in 24 variables (this example was pointed out by the referee).

Nevertheless, there is a generalization of the Ramanujan conjecture which states that the condition on the absolute values of the characteristic roots should be satisfied by the cusp eigenforms orthogonal to the generalized Maass subspace (defined in the above sense). Combining this with the previous observation we may formulate the following

CONJECTURE ON A MOTIVIC INTERPRETATION OF THE STANDARD ZETA FUNCTION. *Let f be a Siegel cusp eigenform of weight $k > 2m + 2$ with a Dirichlet character ψ* mod N *such that f is not a lifting from an automorphic form on a "smaller" algebraic group (i.e., f belongs to the orthogonal complement to the generalized Maass subspace). Then there exists a motive M_f of rank $2m + 1$, weight* 0, *which has Hodge structure of the type*

$$(-k+1, 1-k) + (-k+2, k-2) + \cdots + (-k+m, k-m) + (0, 0) \\ + (k-m, -k+m) + \cdots + (k-1, -k+1),$$

such that the complex conjugation acts on the $(0, 0)$-subspace via the sign $\psi(-1)$ and the function $D(s, f, \chi)$ coincides with the function $L(M_f(\chi), s)$ and $D^*(s, f, \chi) = \Lambda(M_f(\chi), s)$. Moreover, the motive M_f should have coefficients in the field $T = K_0 = \mathbf{Q}(f, \Lambda, \psi)$, generated by Fourier coefficients of f and the values of Λ and ψ.

There are some other confirmations of this conjecture: the first one is based on the already known description of the critical values of the function $D(s, f, \chi)$ given in the next section, and the other one is based on the shape of the p-adic L-function of $D(s, f, \chi)$ constructed in **[Pa14]**: these facts perfectly match with the general conjecture on critical values and on p-adic L-functions associated with motives over $\mathbf{Q}$.

3. Modified L-function of a motive over Q and the modified standard zeta functions

Now we recall some facts about periods and critical values of a motive M over $\mathbf{Q}$. Assume that ρ acts on $M^{w/2, w/2}$ as a scalar $(-1)^{\varepsilon}$. This will be automatically implied by our assumption made later that M is critical at 0. It follows that one can choose appropriate terms $F^{+}_{\mathrm{DR}}M$ and $F^{-}_{\mathrm{DR}}M$ of the de Rham filtration such that

$$\dim(M_{\mathrm{DR}}/F^{+}_{\mathrm{DR}}M) = d^{+}(M), \quad \dim(M_{\mathrm{DR}}/F^{-}_{\mathrm{DR}}M) = d^{-}(M)$$

and that the comparison map I_∞ induces isomorphisms

$$I^{\pm}_{\infty} : M^{\pm} \otimes_{\mathbf{Q}} \mathbf{C} \to M^{\pm}_{\mathrm{DR}} \otimes_{\mathbf{Q}} \mathbf{C},$$

where $M^{\pm}_{\mathrm{DR}} = M_{\mathrm{DR}}/F^{\pm}_{\mathrm{DR}}M$. In order to define periods we put

$$c^{\pm}(M) = (c^{\pm}(M)^{(\tau)}) = \det(I^{\pm}_{\infty}) \in (T \otimes \mathbf{C})^{\times}$$

where the right-hand side denotes the determinants of matrices representing the maps relative to some T-rational bases of the source and the target. Note that the quantities $c^{\pm}(M)$ are defined modulo the multiplicative subgroup $T^{\times}$ of $(T\otimes\mathbf{C})^{\times}$ so that the quantity $c^{\pm}(M)^{(\tau)}$ is a complex constant which is defined modulo $\tau(T)^{\times}$.

We can easily compute periods of the twisted motive $M(s)$, for $\nu = \mathrm{sgn}(-1)^s$, $s \in \mathbf{Z}$,

$$c^{\varepsilon}(M(s)) = (2\pi i)^{d^{\varepsilon\nu}s} c^{\varepsilon\nu}(M) \mod (T)^{\times}.$$

Note that for an integer s the periods of the twisted motive $M(s)$ are given by

$$c^{\varepsilon}(M(s)) = (2\pi i)^{d^{\varepsilon\nu}s} c^{\varepsilon\nu}(M) \bmod T^{\times}$$

where $(-1)^s = \nu 1$ $(\nu = \pm)$.

Periods of the twist with a Dirichlet character $\chi : \mathbf{A}^{\times}_{\mathbf{Q}}/\mathbf{Q}^{\times} \to \mathbf{C}^{\times}$. We shall denote by the same letter χ the corresponding character of the Galois group $\mathrm{Gal}(\mathbf{Q}^{\mathrm{ab}}/\mathbf{Q})$ such that $\chi(\mathrm{Fr}_p^{-1}) = \chi(p)$ for all prime ideals p that do not divide the conductor $c(\chi)$. Let $\mathrm{sgn}(\chi)$ denote the sign of χ; then we have that $d^{\varepsilon}(M(\chi)) = d^{\varepsilon\varepsilon(\chi)}(M)$ and

$$c^{\varepsilon}(M(\chi)) = (G(\chi)^{-1})^{d^{\varepsilon\cdot\varepsilon(\chi)}(M)} c^{\varepsilon\cdot\varepsilon(\chi)}(M) \in (T\otimes\mathbf{C})^{\times} \mod T(\chi)^{\times}$$

where $\varepsilon = \pm$, $\varepsilon(\chi)$ is any of $\varepsilon_\sigma(\chi)$, and $G(\chi)$ is the Gauss sum (see **[Shi6, De3]**):

$$G(\chi) = \sum_{x\in\mathbf{Z}/c(\chi)\mathbf{Z}} \chi(x) \otimes_{\mathbf{Q}} e(x/c(\chi)) \in \mathbf{Q}(\chi)\otimes\mathbf{C},$$

where $e(x) = \exp(2\pi i x)$. Combining the above equalities we obtain the following general formula for the periods of the twist $M(s)(\chi)$:

$$c^{\varepsilon}(M(\chi)(s)) = (G(\chi)^{-1}(1\otimes(2\pi i)^s))^{d^{\varepsilon\nu\varepsilon(\chi)}(M)} c^{\varepsilon\varepsilon(\chi)\nu}(M) \in (T\otimes\mathbf{C})^{\times} \mod T(\chi)^{\times}.$$

Following Coates we shall formulate this modified period conjecture in a form appropriate for further use in the p-adic construction. First we multiply $L(M, s)$ by an appropriate factor at infinity and define

$$\Lambda_{(\infty)}(M, s)^{(\tau)} = E_\infty(M, s)L(M, s)^{(\tau)}$$

as $\Lambda_{(\infty)}(\tau, M, \rho, s)$ in the notation of Coates **[Co]** with $\rho = i$, where τ runs over the set J_T of all complex embeddings of T. We have that $E_\infty(M, s) = E_\infty(\tau, M, \rho, s)$ is the modified Γ-factor at infinity, which actually does not depend on τ. Also we put

$$\Lambda_{(\infty)}(M, s) = (\Lambda_{(\infty)}(M, s)^{(\tau)})_{\tau \in J_T}$$

for the modified L-function with values at $T \otimes \mathbf{C}$, and we put

$$\Omega(\nu, M) = (\Omega(\nu, M)^{(\tau)}) = c^\nu(M)(2\pi i)^{r(M)} \in (T \otimes \mathbf{C})^\times,$$

where $\nu = (-1)^s$, $r(M) = \sum_{j<0} jh(i, j, M)$, and $c^\nu(M)$ is the period of M. Then Deligne's period conjecture can be stated in the following convenient form: if $s = 0$ is critical for M, then for any $s \in \mathbf{Z}$ such that $M(s)$ is critical at 0 we have that there exists an element α of T such that

$$\frac{\Lambda_\infty(M(s), 0)^{(\tau)}}{\Omega(\nu, M)^{(\tau)}} = \tau(\alpha),$$

i.e., that $\Lambda_{(\infty)}(M(s), 0)\Omega(\nu, M)^{-1} \in T \subset T \otimes \mathbf{C}$, where $\nu = \operatorname{sgn}((-1)^s) = \pm$. This statement is deduced from the original conjecture on critical values in Coates's work **[Co]**, where it was shown that

$$E_\infty(M, 0) \sim (2\pi i)^{r(M)} \mod \mathbf{Q}^\times,$$

and it follows that

$$E_\infty(M(s), 0) \sim (2\pi i)^{r(M) - sd^\varepsilon(M)} = (2\pi i)^{(r(M) - sd^\varepsilon(M))} \mod \mathbf{Q}^\times,$$

where $\varepsilon = +$ if $j < 0$ and $\varepsilon = -$ if $j \geq 0$ for $j = w/2$. If we combine this fact with the equivalence

$$c^+(M(s)) \sim (2\pi i)^{d^\nu ns} c^\nu(M) \mod T^\times$$

we deduce from the above form of the conjecture that

$$\Lambda_{(\infty)}(M(s), 0) \sim (2\pi i)^{(r(M) - sd^\varepsilon(M) + sd^\nu(M))} c^\nu(M).$$

Note that in our situation we have that $d^\varepsilon(M) = d^\nu(M)$ because both M and $M(s)$ are critical at 0: we have that $\nu = +$ only for $j - s < 0$ because $M(s)$ is critical, but according to Lemma 3 in **[Co]** the condition $j < 0$ is equivalent in this situation to $j - s < 0$. Taking into account the above formulae for periods of the twisted motive $M(\chi)(s)$, we now state the following

Modified conjecture on the critical values. *Assume that M is critical at 0. If we put for a given sign $\varepsilon_0 = \pm$, $\Omega(\varepsilon_0, M) = (1 \otimes (2\pi i))^{r(M)} c^{\varepsilon_0}(M)$ with $r(M) = \sum_{j<0} jh(j, k, M)$, then for any integer s and Dirichlet character χ such that $M(\chi)(s)$ is critical at 0 and $\varepsilon(\chi)\nu = \varepsilon_0$ we have that*

$$\Lambda_{(\infty)}(M(\chi)(s), 0)((G(\chi)^{-1})^{d^{\varepsilon_0}(M)}\Omega(\varepsilon_0, M))^{-1} \in T(\chi)$$

where $\nu = sgn((-1)^s) = \pm$. We recall that by definition

$$E_\infty(M, s) = E_\infty(\tau, M, \rho, s) = E_\infty(U, \rho, s),$$

where U runs over direct summands of the Hodge decomposition, $\rho = i$ and $E_\infty(U, \rho, s)$ is given by:

(a) If $U = M^{j,k} \oplus M^{k,j}$ *with* $j < k$, *then* $E_\infty(U, \rho, s) = \Gamma_{\mathbf{C},\rho}(s - j)^{h(j,k)}$;

(b) If $U = M^{k,k}$ *with* $k \geq 0$, *then* $E_\infty(U, \rho, s) = 1$;

(c) If $U = M^{k,k}$ *with* $k < 0$, *then* $E_\infty(U, \rho, s) = R_\infty(U, \rho, s)$. *Here*

$$\rho^{-s} = \exp(-\rho\pi s/2), \qquad \Gamma_{\mathbf{C},\rho}(s) = \rho^{-s}\Gamma_{\mathbf{C}}(s),$$
$$\Gamma_{\mathbf{C}}(s) = 2(2\pi)^{-s}\Gamma(s), \qquad \Gamma_{\mathbf{R}}(s) = \pi^{-s/2}\Gamma(s/2),$$

$$R_\infty(U, \rho, s) = L_\infty(\tau, U, s)/(\varepsilon_\infty(\tau, U, \rho, s)L_\infty(\tau, U^\vee(1), -s)$$

with L- *and* ε-*factors described in* **[De3,** p. 329], *so that we have in case* (c)

$$R_\infty(U, \rho, s) = \frac{\Gamma_{\mathbf{R}}(s-k+\delta)}{i^\delta\Gamma_{\mathbf{R}}(1-s+k-\delta)} = \frac{2\Gamma(s-k+\delta)\cos(\pi(s-k+\delta)/2)}{i^\delta(2\pi)^{s-k+\delta}}$$

with $\delta = 0, 1$ *chosen according to the sign of the scalar action of* ρ_σ *on* $U = M_\sigma^{k,k}$ *so that* ρ_σ *acts as* $(-1)^{k+\delta}$.

Now we restrict ourselves to the case of the conjectural motives associated with the standard zeta functions and recall some known algebraic properities of their special values, which were established independently of their motivic interpretation. In order to precisely formulate the result we first introduce the following normalized standard zeta functions:

$$\mathscr{D}^+(s, f, \chi) = \frac{2\Gamma(s+\delta)\cos(\pi(s+\delta)/2)}{i^\delta(2\pi)^{s+\delta}}\prod_{i=1}^m \Gamma_{\mathbf{C}}(s+k-j)D(s, f, \chi),$$

$$\mathscr{D}^-(s, f, \chi) = \prod_{i=1}^m \Gamma_{\mathbf{C}}(s+k-j)D(s, f, \chi).$$

THEOREM ON ALGEBRAIC PROPERTIES OF THE SPECIAL VALUES OF THE STANDARD ZETA FUNCTIONS (see **[Ha1, Bö2,** and **Pa14,** Theorem 3.4]). (a) *For all integers* s *with* $1 \leq s \leq k - \delta - m$ *and the Dirichlet characters* χ *such that for* $s = 1$ *the character* $\chi^2\psi^2$ *is nontrivial and we have that*

$$G(\chi\psi)^{-m-1}\langle f, f\rangle^{-1}\mathscr{D}^+(s, f, \chi) \in K = \mathbf{Q}(f, \Lambda_f, \psi, \chi)$$

where $K = \mathbf{Q}(f, \Lambda_f, \psi, \chi)$ *denotes the field generated by the Fourier coefficients* f, *by the eigenvalues* $\Lambda_f(X)$ *of Hecke operators on* f, *and by the values of the characters* χ *and* ψ.

(b) *For all integers* s *with* $1 - k + \delta + m \leq s \leq 0$ *and all Dirichlet characters* χ *we have that*

$$G(\chi\psi)^{-m}\langle f, f\rangle^{-1}\mathscr{D}^-(s, f, \chi) \in K = \mathbf{Q}(f, \Lambda_f, \psi, \chi).$$

Note that in view of the definitions of $\mathscr{D}^+(s, f, \chi), \mathscr{D}^-(s, f, \chi)$ *and of the above theorem on analytic properies of* $\mathscr{D}(s, f, \chi)$ *we have that*

$$\mathscr{D}^+(s, f, \chi) = 0 \quad \textit{for } s \in \mathbf{N}, s \not\equiv \delta \pmod 2,$$
$$\mathscr{D}^-(s, f, \chi) = 0 \quad \textit{for } s \in \mathbf{Z}, s \leq 0, s \equiv \delta \pmod 2.$$

Also, one can subtract from the proof of the theorem on the algebraic properties some more explicit information about the action of Galois automorphisms $\sigma \in \operatorname{Aut}\mathbf{C}$ on the above special values, namely, for some nonzero constant $\mu(\Lambda, k, \psi) \in \mathbf{C}^\times$ depending only on $k > 2m+2$, the character ψ and the homomorphism Λ we have that

$$\left(\frac{G(\psi\chi)^{-m-1}D^+(s, f, \chi)}{\mu(\Lambda, k, \psi)}\right)^\sigma = \frac{G(\psi^\sigma\chi^\sigma)^{-m-1}D^+(s, f^\sigma, \chi^\sigma)}{\mu(\Lambda^\sigma, k, \psi^\sigma)},$$

$$\left(\frac{G(\psi\chi)^{-m}D^-(s, f, \chi)}{\mu(\Lambda, k, \psi)}\right)^\sigma = \frac{G(\psi^\sigma\chi^\sigma)^{-m}D^-(s, f^\sigma, \chi^\sigma)}{\mu(\Lambda^\sigma, k, \psi^\sigma)},$$

where we adopt the standard notation

$$f^\sigma = \sum_{\xi > 0} a(\xi)^\sigma e_m(\xi z).$$

Also, the following holds:

$$\mu(\Lambda, k, \psi)^{-1}\langle f, f\rangle_C \in \mathbf{Q}(f, \Lambda, \psi)$$

(i.e., the Petersson scalar product $\langle f, f\rangle_C$ differs from the constant $\mu(\Lambda, k, \psi)$ only by an algebraic multiple from the field $K_0 = \mathbf{Q}(f, \Lambda, \psi)$, generated by Fourier coefficients of f and the values of Λ and ψ).

REMARK. We wish to point out a misprint in the corresponding statement of Theorem 3.4 in **[Pa14]**: the exponent m of the Gaussian sum $G(\chi)$ in (1.10) and (1.11) of **[Pa14]** must be replaced by $-m$.

Let us compare these proven results on algebraicity with the above motivic interpretation. If $M = M_f$ is the (conjectural) motive over $\mathbf{Q}$ with coefficients in T associated with f then we should have that $T = K_0 = \mathbf{Q}(f, \Lambda, \psi)$, and the critical values s for $L(M(\chi), s) = \mathscr{D}(s, f, \chi)$ are given by

(a) $s \in \mathbf{Z}$ with $1 \leq s \leq k - \delta - m, s \equiv \delta \pmod 2$ and the Dirichlet characters χ such that for $s = 1$ the character $\chi^2\psi^2$ is nontrivial,

(b) $s \in \mathbf{Z}$ with $1 - k + \delta + m \leq s \leq 0$, $s \not\equiv \delta \pmod 2$, so that in case (a) we have $\varepsilon_0 = 1$ and in case (b) $\varepsilon_0 = -1$.

Also, for an integer s and a Dirichlet character χ we have that

$$\Lambda_{(\infty)}(M(\chi)(s), 0) = i^{m(2s+2k-m-1)/2}\mathscr{D}^+(s, f, \chi)$$

in case (a) and

$$\Lambda_{(\infty)}(M(\chi)(s), 0) = i^{m(2s+2k-m-1)/2}\mathcal{D}^-(s, f, \chi)$$

in case (b), where the power of i comes from the factor $\Gamma_{\mathbf{C},i}(s) = i^{-s}\Gamma_{\mathbf{C}}(s)$ in the definition of the modified Euler factor $E_\infty(M(\chi), s)$ of $\Lambda_{(\infty)}(M(\chi), s)$ at infinity, taking into account that in this situation

$$\sum_{j=-k+1}^{-k+m}(s-j)h(j, -j) = ms - \frac{m(-2k+m+1)}{2} = \frac{m(2s+2k-m-1)}{2}$$

and

$$r(M(\chi)) = \sum_{j<0} jh(i, j, M(\chi)) = \sum_{j=-k+1}^{-k+m} j = -\frac{m(-2k+m+1)}{2}.$$

Note also that for $\varepsilon_0 = \pm$, $d^{\varepsilon_0}(M) = m+1$ or m according as $\varepsilon_0 = +$ or $-$ and

$$\Omega(\varepsilon_0, M) = (1 \otimes (2\pi i))^{r(M)} c^{\varepsilon_0}(M) = (2\pi i)^{-m(-2k+m+1)/2} c^{\varepsilon_0}(M)$$

so that in case (a) the modified period conjecture reads as

$$\begin{aligned}\Lambda_{(\infty)}(M(\chi)(s), 0)((G(\chi)^{-1})^{d^{\varepsilon_0}(M)}\Omega(\varepsilon_0, M))^{-1} \\ = i^{ms}\mathcal{D}^+(s, f, \chi)G(\chi)^{-m-1}c^+(M)^{-1} \in T(\chi),\end{aligned}$$

and in case (b) as

$$\begin{aligned}\Lambda_{(\infty)}(M(\chi)(s), 0)((G(\chi)^{-1})^{d^{\varepsilon_0}(M)}\Omega(\varepsilon_0, M))^{-1} \\ = i^{ms}\mathcal{D}^-(s, f, \chi)G(\chi)^{-m}c^-(M)^{-1} \in T(\chi).\end{aligned}$$

Comparing with the above theorem we see that (m even)

$$c^+(M) \sim c^-(M) \sim \langle f, f\rangle_C \mod \mathbf{Q}(f, \Lambda, \psi)^\times.$$

4. Non-Archimedean integration and admissible measures

Let S be a finite set of primes containing p. The set on which our non-Archimedean zeta functions are defined is the $\mathbf{C}_p$-adic analytic Lie group

$$\mathscr{X}_S = \mathrm{Hom}_{\mathrm{contin}}(\mathrm{Gal}_S, \mathbf{C}_p^\times),$$

where Gal_S is the Galois group of the maximal abelian extension of $\mathbf{Q}$ unramified outside S and infinity, and $\mathbf{C}_p = \widehat{\overline{\mathbf{Q}}}_p$ is the Tate field (completion of an algebraic closure of the p-adic field $\mathbf{Q}_p$). Now we recall the notion of the h-admissible measures on Gal_S and properties of their Mellin transform.

These Mellin transforms are certain p-adic analytic functions on the $\mathbf{C}_p$-analytic Lie group $\mathscr{X}_S$. Recall that by class field theory the group Gal_S is described as the projective limit

$$\mathrm{Gal}_S = \varprojlim_Q (\mathbf{Z}/Q\mathbf{Z})^\times ,$$

where Q runs over integers with support in the set of primes S . The canonical $\mathbf{C}_p$-analytic structure on $\mathscr{X}_S$ is obtained by shifts from the obvious $\mathbf{C}_p$-analytic structure on the subgroup $\mathrm{Hom}_{\mathrm{contin}}(\mathbf{Z}_p^\times , \mathbf{C}_p^\times) \subset \mathscr{X}_S$. We regard the elements of finite order $\chi \in \mathscr{X}_S^{\mathrm{tors}}$ as Dirichlet characters whose conductor $c(\chi)$ may contain only primes in S , by means of the decomposition,

$$\chi : \mathbf{A}_\mathbf{Q}^\times/\mathbf{Q}^\times \xrightarrow{\text{class field theory}} \mathrm{Gal}_S \to \overline{\mathbf{Q}}^\times \xrightarrow{i_\infty} \mathbf{C}^\times ,$$

where i_∞ is the fixed embedding. The characters $\chi \in \mathscr{X}_S^{\mathrm{tors}}$ form a discrete subgroup $\mathscr{X}_S^{\mathrm{tors}} \subset \mathscr{X}_S$. We shall need also the natural homomorphism

$$x_p : \mathrm{Gal}_S \to \mathbf{Z}_p^\times \to \mathbf{C}_p^\times , \qquad x_p \in \mathscr{X}_S.$$

so that all integers $k \in \mathbf{Z}$ can be regarded as characters of the type $x_p^k : y \mapsto y^k$.

Recall that a p-adic measure on Gal_S may be regarded as a bounded $\mathbf{C}_p$-linear form μ on the space $\mathscr{C}(\mathrm{Gal}_S)$ of all continuous $\mathbf{C}_p$-valued functions

$$\varphi \to \mu(\varphi) = \int_{\mathrm{Gal}_S} d\mu \in \mathbf{C}_p , \qquad \varphi \in \mathscr{C}(\mathrm{Gal}_S) ,$$

which is uniquely determined by its restriction to the subspace $\mathscr{C}^1(\mathrm{Gal}_S)$ of locally constant functions. We denote by $\mu(a + (Q))$ the value of μ on the characteristic function of the set

$$a + (Q) = \{x \in \mathrm{Gal}_S \mid x \equiv a \bmod Q\} \subset \mathrm{Gal}_S .$$

The Mellin transform L_μ of μ is a bounded analytic function

$$L_\mu : \mathscr{X}_S \to \mathbf{C}_p , \qquad L_\mu(\chi) = \int_{\mathrm{Gal}_S} d\mu \in \mathbf{C}_p , \quad \chi \in \mathscr{X}_S ,$$

on $\mathscr{X}_S$, which is uniquely determined by its values $L_\mu(\chi)$ for the characters $\chi \in \mathscr{X}_S^{\mathrm{tors}}$.

A more delicate notion of an h-admissible measure was introduced by Amice and Velu and Višik (see **[Am-V, V1]**). Let $\mathscr{C}^h(\mathrm{Gal}_S)$ denote the space of $\mathbf{C}_p$-valued functions that can be locally represented by polynomials of degree less than a natural number h of the variable $x_p \in \mathscr{X}_S$ introduced above.

DEFINITION. *A* $\mathbf{C}_p$*-linear form* $\mu : \mathscr{C}^h(Gal_S) \to \mathbf{C}_p$ *is called h-admissible measure if for all* $a \in \mathrm{Gal}_S$ *and for all* $r = 0, 1, \dots , h - 1$ *the following* *growth condition is satisfied*:

$$\left|\sup_{a\in \mathrm{Gal}_S}\int_{a+(Q)}(x_p-a_p)^r\,d\mu\right| = o(|\,Q\,|_p^{r-h}).$$

Note that the notion of a bounded measure is covered by the case $h=1$, but the set of 1-admissible measures is bigger: it consists of so-called measures of bounded growth **[Man4, V1]**, which are characterized by the property that they grow on open compact sets $a+(Q)$ slower as $o(|Q|_p^{-1})$. We know (essentially due to Amice and Velu and Višik) that each h-admissible measure can be uniquely extended to a linear form on the $\mathbf{C}_p$-space of all locally analytic functions so that one can associate to its Mellin transform

$$L_\mu:\mathscr{X}_S\to\mathbf{C}_p,\qquad L_\mu(\chi)=\int_{\mathrm{Gal}_S}\chi\,d\mu\in\mathbf{C}_p,\quad \chi\in\mathscr{X}_S,$$

which is a $\mathbf{C}_p$-analytic function on $\mathscr{X}_S$ of the type $o(\log x_p^h)$. Moreover, the measure μ is uniquely determined by the special values of the type $L_\mu(\chi x_p^r)$ ($\chi\in\mathscr{X}_S^{\mathrm{tors}}$, $r=0,1,\dots,h-1$).

5. The Newton polygon and the Hodge polygon of a motive; p-ordinary and p-admissible motives

We shall formulate in the next section a general conjecture on p-adic L-functions of motives in terms of the existence of certain h-admissible measures, where the quantity is defined in terms of the Newton polygons and the Hodge polygons of a motive. Properties of these polygons are closely related to the notions of a p-ordinary and a p-admissible motive; such motives will correspond to the case $h=1$.

From now on we fix an embedding $\tau\colon T\to\mathbf{C}_p$ in order to deal with p-adic L-functions. It is often convenient to omit the symbol (τ) from the notation $L(M,s)^{(\tau)}$, $\Lambda_{(\infty)}(M,s)^{(\tau)}$, $c^\varepsilon(M)^{(\tau)}$, $\Omega(\varepsilon_0,M)^{(\tau)}$, viewing these quantities as complex numbers. Then for a motive M over $\mathbf{Q}$ with coefficients in T under the assumptions of the period conjecture of §3 the algebraic number

$$\Lambda_{(\infty)}(M(\chi)(s),0)G(\chi)^{-d^{\varepsilon_0}(M)}\Omega(\varepsilon_0,M)^{-1}\in T(\chi)$$

can be regarded via the fixed embeddings $i_\infty\colon\overline{\mathbf{Q}}\to\mathbf{C}$ and $i_p\colon\overline{\mathbf{Q}}\to\mathbf{C}_p$ as an element of both $\mathbf{C}$ and $\mathbf{C}_p$. We are now going to define the Newton polygon $P_{\mathrm{Newton}}(u)=P_{\mathrm{Newton}}(u,M)$ and the Hodge polygon $P_{\mathrm{Hodge}}(u)=P_{\mathrm{Hodge}}(u,M)$ attached to M and to the fixed embedding $\tau\in J_T$. First for a prime number p we consider (using i_∞) the local p-polynomial

$$\begin{aligned}L_p(M,X)^{-1}&=1+A_1(p)X+\cdots+A_d(p)X^d\\&=(1-\alpha^{(1)}(p)X)\cdot(1-\alpha^{(2)}(p)X)\cdot\cdots\cdot(1-\alpha^{(d)}(p)X),\end{aligned}$$

and we assume that its inverse roots are indexed in such a way that

$$\mathrm{ord}_p\alpha^{(1)}(p)\le\mathrm{ord}_p\alpha^{(2)}(p)\le\cdots\le\mathrm{ord}_p\alpha^{(d)}(p).$$

The Newton polygon $P_{\text{Newton}}(u)$ $(0 \le u \le d)$ of M at p is the convex hull of the points $(i, \operatorname{ord}_p A_i(p))$ $(i = 0, 1, \dots, d)$. The important property of the Newton polygon is that the length of the horizontal segment of slope i is equal to the number of the inverse roots $\alpha^{(j)}(p)$ such that $\operatorname{ord}_p \alpha^{(j)}(p) = i$ (note that this number may not necessarily be an integer but this will be the case for the p-ordinary motives below).

The Hodge polygon $P_{\text{Hodge}}(u)$ $(0 \le u \le d)$ of M is defined using the Hodge decomposition of the d-dimensional $\mathbf{C}$-vector space

$$M^{(\tau)} = M \otimes_{T,\tau} \mathbf{C} = \bigoplus_{i,j} M^{(\tau)i,j}$$

where we keep τ fixed and regard $M^{(\tau)i,j} = M^{i,j} \otimes_{T,\tau} \mathbf{C}$ as the $\mathbf{C}$-subspace of $M^{i,j}$ on which T acts via $\tau \in J_T$. Note that the dimension $h^{(\tau)}(i, j) = \dim_{\mathbf{C}} M^{(\tau)i,j}$ does not depend on τ; see Deligne **[De3]**.

The Hodge polygon $P_{\text{Hodge}}(u)$ by definition passes through the points

$$(0, 0), \dots, \left(\sum_{i' \le i} h^{(\tau)}(i', j), \sum_{i' \le i} i' h^{(\tau)}(i', j)\right), \dots,$$

so that the length of the horizontal segment of slope i is equal to the dimension $h^{(\tau)}(i, j)$.

Now we recall the definition of a p-ordinary motive in the simplest case $T = \mathbf{Q}$ (see **[Co, Co-PeRi]**). We assume that M is pure of weight w and rank d. Let $G_p = \operatorname{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ be the decomposition group (of a place in $\overline{\mathbf{Q}}$ over p) and $\psi_p : G_p \to \mathbf{Z}_p^\times$ be the cyclotomic character of D_p. Then M is called p-ordinary at p if the following conditions are satisfied:

(i) The inertia group $I_p \subset G_p$ acts trivially on each of the ℓ-adic realizations M_ℓ for $\ell \ne p$.

(ii) There exists a decreasing filtration $F_p^i V$ on $V = M_p = M_B \otimes \mathbf{Q}_p$ of $\mathbf{Q}_p$-subspaces that are stable under the action of G_p such that for all $i \in \mathbf{Z}$ the group D_p acts on $F_p^i V / F_p^{i+1} V$ via some power of the cyclotomic character, say $\psi_p^{-e_i}$. Then $e_1(M) \ge \cdots \ge e_t(M)$ and the following properties take place:

(a)

$$\dim_{\mathbf{Q}_p} F_p^i V / F_p^{i+1} V = h(e_i, w - e_i).$$

(b) The Hodge polygon and the Newton polygon of M coincide:

$$P_{\text{Newton}}(u) = P_{\text{Hodge}}(u).$$

If, furthermore, M is critical at 0, then it is easy to verify that the number d_p of the inverse roots $\alpha^{(j)}(p)$ with $\operatorname{ord}_p \alpha^{(j)}(p) < 0$ is equal to $d^+ = d^+(M)$ of M_σ^+.

In the general case (of a motive M with coefficients in T) the notion of a p-ordinary motive can be defined using the restriction of the coefficient field T to $\mathbf{Q}$ (the last operation corresponds to forgetting of the T-module structure on the realizations of M). In this way we get a motive M' with coefficients in $\mathbf{Q}$ of the same weight w and rank $\mathrm{rk}(M') = [T:\mathbf{Q}]\cdot d$.

For p-ordinary motives over $\mathbf{Q}$ and their Dirichlet twists, Coates and Perrin-Riou **[Co–PeRi]** have formulated a general conjecture on the existence of bounded p-adic measures attached to such motives. However, it turns out that such bounded measures can exist even for certain non-p-ordinary motives, which can be characterized by the following simple condition:

DEFINITION. A motive M over $\mathbf{Q}$ with coefficients in T is called *admissible at* p if $P_{\mathrm{Newton}}(d^+) = P_{\mathrm{Hodge}}(d^+)$. Here $d^+ = d(M)$ is the dimension of M^+.

On the other hand, in a number of cases when M is not p-ordinary and even when M is not admissible at p one can prove the existence of the corresponding (growing) h-admissible measures. It was recently found by Andrzej Dabrowski (Moscow University) that all these cases admit a unified description if we use the following positive integer h which is defined in terms of the difference between the Newton polygon and the Hodge polygon of M:

$$h = \max[P_{\mathrm{Newton}}(d^{\pm}) - P_{\mathrm{Hodge}}(d^{\pm}))] + 1.$$

Note the following important properties of the quantity h:

(i) $h = h(M)$ does not change if we replace M by its Tate twist.

(ii) $h = h(M)$ does not change if we replace M by its twist $M = M(\chi)$ with the Dirichlet character χ.

In the next section we state in terms of this quantity a general conjecture on p-adic L-functions.

6. A conjecture on p-adic L-functions of motives over $\mathbf{Q}$

In order to formulate precisely a general conjecture on p-adic L-functions of a motive M over $\mathbf{Q}$ with coefficients in T we suppose that M is pure of weight w, M has rank d, and there exists an integer s such that the motive $M(s)$ is critical at 0. Then we put

$$s_* = \min\left\{s \mid \exists\chi \in \mathscr{X}_S^{\mathrm{tors}} \text{ such that } N = M(\chi)(s) \text{ is critical at } 0\right\},$$
$$s^* = \max\left\{s \mid \exists\chi \in \mathscr{X}_S^{\mathrm{tors}} \text{ such that } N = M(\chi)(s) \text{ is critical at } 0\right\}.$$

Then the number $s^* - s_* + 2$ coincides with the width of the critical strip of our motive M. The integers s_* and s^* can be characterized in terms of the Hodge decomposition of M:

$$s_* = \max\{j \mid \exists j, k, j < k \text{ such that } h(j, k, M) \neq 0\} + 1,$$
$$s^* = \min\{j \mid \exists j, k, j > k \text{ such that } h(j, k, M) \neq 0\}.$$

Furthermore, for an integer s and a Dirichlet character χ, we put $N = M(\chi)(s)$ and define for $q \in S$ the q-factors:

$$A_q(M(\chi), s) = \begin{cases} \prod_{i=d^+ +1}^{d} (1 - \chi(q)\alpha^{(i)}(q)q^{-s}) \prod_{i=1}^{d^+} (1 - \chi^{-1}(q)\alpha^{(i)}(q)^{-1}q^{s-1}) & \text{for } q \nmid c(\chi), \\ \prod_{i=1}^{d^+} \left(\frac{q^s}{\alpha^{(i)}(q)} \right)^{\operatorname{ord}_q c(\chi)}, & \text{otherwise.} \end{cases}$$

Let us fix a sign $\varepsilon_0 = \pm 1$. Assuming that the modified period conjecture of §3 is true, if we put for ε_0

$$\Omega(\varepsilon_0, M) = (1 \otimes (2\pi i))^{r(M)} c^{\varepsilon_0}(M)$$

with $r(M) = \sum_{j<0} jh(j, k, M)$, then for any integer s and Dirichlet character χ such that $M(\chi)(s)$ is critical at 0 and $\varepsilon(\chi)\nu = \varepsilon_0$, we have

$$\Lambda_{(\infty)}(M(\chi)(s), 0)(G(\chi)^{-d^{\varepsilon_0}(M)} \Omega(\varepsilon_0, M))^{-1} \in T(\chi)$$

where $\nu = \operatorname{sgn}((-1)^s) = \pm$.

CONJECTURE. *For each sign ε_0 there exists a $\mathbf{C}_p$-analytic function $L_{(p)}^{(\varepsilon_0)}$ on $\mathscr{X}_S$ with the properties*:

(i) *for all but a finite number of pairs (s, χ) such that the motive $N = M(\chi)(s)$ is critical at 0 and $\varepsilon_0 = \varepsilon(\chi)\nu$ we have that*

$$L_{(p)}^{(\varepsilon_0)}(\chi x_p^s) = G(\chi)^{-d^{\varepsilon_0}(M)} \prod_{q \in S} A_q(N) \cdot \frac{\Lambda_{(\infty)}(M(\chi)(s), 0)}{\Omega(\varepsilon_0, M)};$$

(ii) *the function $L_{(p)}^{(\varepsilon_0)}$ is holomorphic on $\mathscr{X}_S$ if $M^{k,k} = 0$; otherwise there exists a finite set $\Xi \subset \mathscr{X}_S$ of p-adic characters and positive integers $n(\xi)$ (for $\xi \in \Xi$) such that for any $g_0 \in Gal_S$ we have that the function*

$$\prod_{\xi \in \Xi} (x(g_0) - \xi(g_0))^{n(\xi)} L_{(p)}^{(\varepsilon_0)}(x)$$

is holomorphic on $\mathscr{X}_S$;

(iii) *the holomorphic function in* (ii) *is bounded if*

$$P_{\text{Newton}}(d^+) = P_{\text{Hodge}}(d^+) \quad \text{for } d^+ = d^+(N);$$

(iv) *in the general case the holomorphic function in* (ii) *belongs to the type $o(\log x_p^h)$ and it can be represented as the Mellin transform of an h-admissible measure with h defined at the end of the previous section*;

(v) *if $h \le s^* - s_* + 1$ then the function $L_{(p)}^{(\varepsilon_0)}$ is uniquely determined by the conditions* (i)–(ii).

Note that the last statement follows from the properties of h-admissible measures (see §4).

REMARK. It would be interesting to rewrite the right-hand side of the equality in (i) in a more invariant form using the ε-factors of certain complex representations of the Weil–Deligne groups W'_p. This was done by Coates for p-ordinary motives over $\mathbf{Q}$ [Co]. These representations can probably occur in the complexification $Y = M_\lambda \otimes_{T_\lambda, \tau} \mathbf{C}$ of the λ-adic realization M_λ of M, where for $\tau \in J_T$ the same symbol τ denotes an embedding $T_\lambda \to \mathbf{C}$ extending $\tau : T \to \mathbf{C}$.

7. The p-adic L-function of the conjectural motive M_f

7.1. Now we specialize the conjecture of the previous section to the case of the motive M_f of rank $2m+1$, weight 0, which has the Hodge structure of the type

$$(-k+1, 1-k) + (-k+2, k-2) + \cdots + (-k+m, k-m) + (0, 0) \\ + (k-m, -k+m) + \cdots + (k-1, -k+1),$$

such that the complex conjugation acts on the $(0, 0)$-subspace via the sign $\psi(-1)$ and the function $D(s, f, \chi)$ coincides with the function $L(M_f(\chi), s)$ and $D^*(s, f, \chi) = \Lambda(M_f(\chi), s)$. We know that $T = K_0 = \mathbf{Q}(f, \Lambda, \psi)$, and the critical values s for $L(M(\chi), s) = \mathscr{D}(s, f, \chi)$ are given by

(a) $s \in \mathbf{Z}$ with $1 \le s \le k - \delta - m$, $s \equiv \delta \pmod 2$, and the Dirichlet characters χ such that for $s = 1$ the character $\chi^2\psi^2$ is nontrivial,

(b) $s \in \mathbf{Z}$ with $1 - k + \delta + m \le s \le 0$, $s \not\equiv \delta \pmod 2$. Also, for an integer s and a Dirichlet character χ we have that

$$\Lambda_{(\infty)}(M(\chi)(s), 0) = i^{m(2s+2k-m-1)/2}\mathscr{D}^+(s, f, \chi)$$

in case (a) and

$$\Lambda_{(\infty)}(M(\chi)(s), 0) = i^{m(2s+2k-m-1)/2}\mathscr{D}^-(s, f, \chi)$$

in case (b),

$$r(M(\chi)) = \sum_{j<0} jh(i, j, M(\chi)) = -\frac{m(-2k+m+1)}{2},$$

for $\varepsilon_0 = \pm$, $d^{\varepsilon_0}(M) = m+1$ or m according as $\varepsilon_0 = +$ or $-$,

$$\Omega(\varepsilon_0, M) = (1 \otimes (2\pi i))^{r(M)} c^{\varepsilon_0}(M) = (2\pi i)^{-m(-2k+m+1)/2} c^{\varepsilon_0}(M).$$

According to the modified period conjecture we have in case (a) that $\varepsilon_0 = 1$ and

$$\Lambda_{(\infty)}(M(\chi)(s), 0)((G(\chi)^{-1})^{d^{\varepsilon_0}(M)}\Omega(\varepsilon_0, M))^{-1} \\ = i^{ms}\mathscr{D}^+(s, f, \chi)G(\chi)^{-m-1}c^+(M)^{-1} \in T(\chi)$$

and in case (b) that $\varepsilon_0 = -1$ and

$$\Lambda_{(\infty)}(M(\chi)(s), 0)((G(\chi)^{-1})^{d^{\varepsilon_0}(M)}\Omega(\varepsilon_0, M))^{-1} \\ = i^{ms}\mathscr{D}^-(s, f, \chi)G(\chi)^{-m}c^-(M)^{-1} \in T(\chi)$$

where

$$c^+(M) \sim c^-(M) \sim \langle f, f\rangle_C \mod \mathbf{Q}(f, \Lambda, \psi)^\times .$$

In order to explicitly formulate the result for the motive $M = M_f$ and for its $(2m+1)$ characteristic numbers $\alpha^{(1)}(q) = \psi(q)\alpha_m^{-1}(q), \dots, \alpha^{(m)}(q) = \psi(q)\alpha_1^{-1}(q), \alpha^{(m+1)}(q) = \psi(q), \alpha^{(m+2)}(q) = \psi(q)\alpha_1(q), \dots, \alpha^{(2m+1)}(q) = \psi(q)\alpha_m(q)$, we consider separately the cases $\varepsilon_0 = 1$ and $\varepsilon_0 = -1$ and put

$$A_q^+(M(\chi)(s)) = \begin{cases} (1-(\chi\psi)^{-1}(q)q^{s-1}) \\ \times\prod_{i=1}^m (1-(\chi\psi)(q)\alpha^{(i)}(q)q^{-s})(1-(\chi\psi)(q)^{-1}\alpha^{(i)}(q)^{-1}q^{s-1}), \\ \qquad\qquad \text{in case (a) if } q \nmid c(\chi), \\ \left(\dfrac{q^{s-m(k-(m+1)/2)}}{\psi(q)\alpha_0^2(q)}\right)^{\operatorname{ord}_q c(\chi)}, \quad \text{in case (a) if } q \text{ divides } c(\chi). \end{cases}$$

$$A_q^-(M(\chi)(s)) = \begin{cases} (1-(\chi\psi)(q)q^{-s}) \\ \times\prod_{i=1}^m (1-\chi\psi(q)\alpha^{(i)}(q)q^{-s})(1-(\chi\psi)(q)^{-1}\alpha^{(i)}(q)^{-1}q^{s-1}), \\ \qquad\qquad \text{in case (a) if } q \nmid c(\chi), \\ \left(\dfrac{q^{s-m(k-(m+1)/2}}{\psi(q)\alpha_0^2(q)}\right)^{\operatorname{ord}_q c(\chi)}, \quad \text{in case (a) if } q \text{ divides } c(\chi). \end{cases}$$

This definition is obtained from the general formula for the factor $A_q(M(\chi)(s))$, taking into account the relation

$$\alpha_0^2(q)\alpha_1(q)\cdots\alpha_m(q) = \psi(q)^m q^{m(k-(m+1)/2)}.$$

Note that

$$s_* = \max\{j \mid \exists j, k, j<k \text{ such that } h(j,k,M) \neq 0\} + 1 = m-k+1,$$
$$s^* = \min\{j \mid \exists j, k, j>k \text{ such that } h(j,k,M) \neq 0\} = k-m.$$

Note also that

$$P_{\text{Newton}}(d^+) = \operatorname{ord}_p(\alpha_1(p)^{-1}\alpha_2(p)^{-1}\cdots\alpha_m(p)^{-1}) = 2\operatorname{ord}_p\alpha(p) - m(k-(m+1)/2)$$

because of the relation $\alpha_0^2(q)\alpha_1(q)\cdots\alpha_m(q) = \psi(q)^m q^{m(k-(m+1)/2)}$, and

$$P_{\text{Hodge}}(d^+) = \sum_{j<0} jh(i, j, M(\chi)) = -m\left(k - \frac{m+1}{2}\right),$$

so that the difference

$$P_{\text{Newton}}(d^+) - P_{\text{Hodge}}(d^+) \\ = \text{ord}_p(\alpha_1(p)^{-1}\alpha_2(p)^{-1}\cdots\alpha_m(p)^{-1}) - \sum_{j<0} jh(i, j, M(\chi))$$

is equal to $2\,\text{ord}_p\,\alpha_0(p)$. Therefore the condition

$$P_{\text{Newton}}(d^+) = P_{\text{Hodge}}(d^+)$$

for the existence of the bounded standard p-adic L-function transforms to $\text{ord}_p\,\alpha_0(p) = 0$, or $|\; i_p(\alpha_0(p))\;|_p = 1$, and in this case the conjecture of §7 becomes essentially the main theorem from Chapter 3 of **[Pa14]** which has been (unconditionaly) proven for the standard zeta function of a Siegel cusp eigenform f of an even degree m and of weight $k > 2m+2$, although the existence of the motive $M = M_f$ is not known yet.

Moreover, we proved recently that for the standard zeta functions the conjecture of §6 holds in the general case under a technical assumption.

7.2. We set

$$h = [2\,\text{ord}_p\,\alpha_0(p)] + 1,$$

$\Omega(M) = (2\pi i)^{-m(-2k+m+1)/2}\langle f, f\rangle_C$, and for $\varepsilon_0 = \pm$, $d^{\varepsilon_0}(M) = m+1$ or m according as $\varepsilon_0 = +$ or $-$.

Assume that $a(\xi_0) \neq 0$ for some $\xi_0 > 0$ with $\det(2\xi_0) = 1$.

Moreover, we make the essential assumption that for $a_0(\xi_0) \neq 0$ the Fourier coefficient $a_0(\xi_0)$ of the S-auxiliary form $f_0(z) = \sum_{\xi>0} a_0(\xi)e_m(\xi z)$ defined in §1, (1.47), Chapter 2 of **[Pa14]**. According to its definition, this form satisfies the following multiplicativity property $a_0(Q\xi) = \alpha_0(Q)a_0(\xi)$ for $Q \in \mathbf{N}$ divisible only by primes in S. The importance of the above nonvanishing condition was pointed out to the author by S. Böcherer. It turns out that it actually may not be satisfied by certain Eisenstein series (note that we deal here with cusp forms). However, it is not known in general whether there exist cusp forms violating the nonvanishing condition.

7.3. THEOREM. *Let f be a Siegel cusp eigenform of even degree m and of weight $k > 2m+2$ satisfying the above nonvanishing assumption. For $\varepsilon_0 = \pm 1$ there exists a $\mathbf{C}_p$-analytic function $L_{(p)}^{(\varepsilon_0)}$ on $\mathscr{X}_S$ with the properties:*

(i)a) *for all pairs (s, χ) such that $s \in \mathbf{Z}$ with $1 \leq s \leq k - \delta - m$, $s \equiv \delta \pmod 2$ and for $s = 1$ the character $\chi^2\psi^2$ is nontrivial,*

$$L_{(p)}^{(+)}(\chi x_p^s) = G(\chi)^{-m-1}\prod_{q\in S} A_q^+(N)\cdot i^{ms}\mathscr{D}^+(s, f, \overline{\chi})\langle f, f\rangle_C^{-1},$$

b) *for all pairs (s, χ) such that*

$$s \in \mathbf{Z} \textit{ with } 1-k+\delta+m \leq s \leq 0, \qquad s \not\equiv \delta \pmod 2,$$

$$L_{(p)}^{(-)}(\chi x_p^s) = G(\chi)^{-m}\prod_{q\in S} A_q^-(N)\cdot i^{ms}\mathscr{D}^-(s, f, \overline{\chi})\langle f, f\rangle_C^{-1};$$

(ii) *for each* $c \in \mathbf{Z}_S^\times$ *the regularized function* $(1 - x(c)^2 c^{-2}) L_{(p)}^{(\varepsilon_0)}(x)$ *is holomorphic on* $\mathscr{X}_S$;

(iii) *if* $\operatorname{ord}_p \alpha_0(p) = 0$ *then the holomorphic function in* (ii) *is a bounded* $\mathbf{C}_p$*-analytic function*;

(iv) *in the general case the holomorphic function in* (ii) *belongs to the type* $o(\log x_p^h)$ *with* $h = [2 \operatorname{ord}_p \alpha_0(p)] + 1$ *and it can be represented as the Mellin transform of an* h*-admissible measure*;

(v) *if* $h \le s^* - s_* + 1 = 2k - 2m$, *then the function* $L_{(p)}^{(\varepsilon_0)}$ *is uniquely determined by conditions* (i)–(ii).

8. The construction of admissible measures attached to standard zeta functions

8.1. We now outline the method of the proof of Theorem 7.3 and show the existence of the above h-admissible measures attached to the standard zeta functions $\mathscr{D}(s, f, \chi)$. This proof is based on congruences for the Fourier coefficients of Siegel modular forms. This is very similar to the original proof in the p-ordinary case $\operatorname{ord}_p \alpha_0(p) = 0$ and is based on the Rankin method. We use the relation of the function $D(s, f, \chi)$ with a convolution of Rankin type given by the general Andrianov's identity (see **[An-K]**):

$$a(\xi) \det \xi_0^{-(s+k-1+\nu)/2} D(s, f, \chi)$$
$$= L\left(s + \frac{m}{2}, \psi \chi_{\xi_0} \chi\right) \prod_{i=0}^{m/2-1} L(2s+2i, \psi^2 \chi^2) L\left(\frac{s+k-1+\nu}{2}, f, \theta_{2\xi_0}^{(\nu)}(\chi)\right),$$

where χ is a Dirichlet character, $\nu = 0$ or 1 according as $\chi\psi(-1) = (-1)^\nu$, ξ_0 being a fixed half-integral symmetric positive-definite matrix. In this identity the right-hand side is the convolution of Rankin type of f and a theta function $\theta_{2\xi_0}^{(\nu)}(\chi)$. This convolution is completely determined by the Fourier coefficients of f. The non-Archimedean part of the construction is based on the theory of distributions and h-admissible measures.

Using a general criterion of finite additivity we then construct in §8.2 complex-valued distributions associated with $D(s, f, \chi)$ by means of their values at Dirichlet characters. Then we give in §8.3 an (Archimedean) integral representation for these values, which makes it possible to express the distributions in terms of the Fourier coefficients of Siegel–Eisenstein series of **[Pa14,** Chapter 2], by applying the holomorphic projection operator. Then we give in §8.4 the explicit formulae for certain Fourier coefficients in order to obtain p-adic integral representations for them in terms of Mazur's measure. After a regularization these distributions become h-admissible measures which take algebraic values at compact open subsets of Gal_S. The proof of Theorem 7.3 is completed in §8.5 by the use of the non-Archimedean Mellin transform.

8.2. Complex-valued distributions associated with standard zeta functions of Siegel modular forms (cf. [**Pa14**, p. 88]). Let as in [**Pa14**, p. 49]

$$f(z) = \sum_{\xi \in B_m} c(\xi) e_m(\xi z) \in \mathscr{M}_k^m(N, \psi)$$

be an eigenfunction of the algebra $\mathscr{L} = \mathscr{L}_p^m(N)$ whose eigenvalue is a homomorphism $\lambda_f : \mathscr{L} \to \mathbf{C}$, $f|X = \lambda_f(X) f$ for all $X \in \mathscr{L}$ which is given by the $(m+1)$-tuple $\alpha_0 = \alpha_0(p)$, $\alpha_1 = \alpha_1(p)$, $\dots$, $\alpha_m = \alpha_m(p)$ of the Satake p-parameters. Put

$$f_0 = \sum_{i=1}^{\tilde{m}-1} \alpha_0(p)^{-i} f | V_i^+(p) \qquad (\tilde{m} = 2^m). \tag{8.1}$$

Here $V_i^+(p)$ are operators defined by the following Andrianov's factorization formulae for the Hecke polynomials:

$$Q(z) = \left(\sum_{i=1}^{2^m - 1} V_i^+ z^i \right) (1 - \Pi_+ z) = (1 - \Pi_- z) \left(\sum_{i=1}^{2^m - 1} V_i^- z^i \right),$$

where

$$V_i^+ = \sum_{j=0}^{i} (-1)^j T_j \Pi_+^{i-j},$$

$$V_i^- = \sum_{j=0}^{i} (-1)^j \Pi_-^{i-j} T_j \in \mathscr{L}_0.$$

Then according to Proposition 1.10 in [**Pa14**, p. 49], the function f_0 belongs to $\mathscr{M}_k^m(Np^{\tilde{m}-1}, \psi)$ with $\tilde{m} = 2^m$, and it turns out to be an eigenfunction of the Frobenius operator $\Pi_+(p)$:

$$f_0 | \Pi_+(p) = \alpha_0(p) f_0.$$

The definition of the operators $\Pi_+(M)$, $\Pi_-(M)$, $V_i^+(q)$, $V_i^-(q) \in \mathscr{L}_{0,q}^m(N)$ and of the numbers $\alpha_0(q)$ $(q \nmid N)$ can be extended by multiplicativity to all positive integers M coprime with N; more precisely, operators $V^+(M)$, $V^-(M)$ are defined by the identities

$$\sum_{M | M_0^{\tilde{m}-1}} M^{-s} V^+(M) = \prod_{q \nmid N} \left[\sum_{i=1}^{\tilde{m}-1} q^{-s} V_i^+(q) \right],$$

$$\sum_{M | M_0^{\tilde{m}-1}} M^{-s} V^-(M) = \prod_{q \nmid N} \left[\sum_{i=1}^{\tilde{m}-1} q^{-s} V_i^-(q) \right],$$

where the notation

$$V^{\pm}(q^i) = V_i^{\pm}(q), \qquad M_0 = \prod_{q \in S} q$$

is used. If we put

$$f_0 = f_{0,S} = \sum_{M|M_0^{\tilde{m}-1}} \alpha_0(M)^{-1} f|V^+(M),$$

we see that

$$f_0 = f_{0,S} \in \mathscr{M}_k^m(NM^{\tilde{m}-1}, \psi),$$

and for all positive integers M with support $S(M)$ in the set S we have that

$$f_0|\Pi_+(M) = \alpha_0(M) f_0.$$

Let

$$f_{0,S}(z) = \sum_{\xi\in B_m} a_0(\xi)e_m(\xi z) \in \mathscr{M}_k^m(NM^{\tilde{m}-1}, \psi) \qquad (\xi \in B_m)$$

be Fourier expansion of the function $f_{0,S}(z)$; then the following multiplicativity property of its Fourier coefficients holds: for all $M \in \mathbf{N}$ with $S(M) \subset S$ we have that

$$a_0(M\xi) = \alpha_0(M)a_0(\xi) \qquad (\xi \in A_m, \xi \geq 0). \tag{8.2}$$

In the notation of **[Pa14,** Chapter 3], we put

$$\hat{\xi}_0 = q_0\xi_o^{-1}, \qquad N_0 = 4q_0CM_0^{\tilde{m}-1}.$$

PROPOSITION. *Let $s \in \mathbf{C}$, $\operatorname{Re}(s) \gg 0$. Then there exists a complex-valued distribution $\mathscr{D}_s$ on G_S which is uniquely determined by its values on Dirichlet characters χ mod M with $S(M) \subset S$ given by*

(8.3)

$$\begin{aligned}
&2a_0(\xi_0)\det\xi_0^{-(s+k-1+\nu)/2}\mathscr{D}_s(\chi)\\
&\quad = \alpha_0(M_0^{\tilde{m}-1}M')^{-1}(CM_0^{\tilde{m}-1}M')^{m(2s+2k-2-m)/4}C^{m(2\nu+m)/4}\\
&\quad\times L_{N_0}\left(s+\frac{m}{2}, \psi\chi_{\xi_0}\overline{\chi}\right)\prod_{i=0}^{m/2-1} L_{N_0}(2s+2i, \psi^2\overline{\chi}^2)\\
&\quad\times \det((2q_0)^{-\frac{1}{2}}\xi_0)^{m/2+\nu} L\left(\frac{s+k-1+\nu}{2}, f_0|V(C), \theta_{2\hat{\xi}_0}^{(\nu)}(\chi_M)|W(N_0M')\right),
\end{aligned}$$

where

$$f_0|V(C)(z) = f_0(Cz),$$

(8.4)

$$\theta_{2\hat{\xi}_0}^{(\nu)}(\chi_M)|W(N_0M')(z) = \det\left(\sqrt{N_0M'z}\right)^{-m/2-\nu}\theta_{2\hat{\xi}_0}^{(\nu)}(\chi_M)(-(N_0M'z)^{-1}),$$

where M, M' are sufficiently large positive integers with the condition $M_0C_\chi|M$, $MM_0C_\chi^2|M'$ so that $S(M) = S(M') = S$ with C_χ being the conductor of the character χ and χ_M the Dirichlet character modulo M induced by χ.

8.3. Now we consider as in **[Pa14,** p. 94], the normalized distributions given by the following formula:

$$\text{(8.5)}\quad \mathscr{D}^{-}_{s,M}(\chi) = \langle f, f\rangle_C^{-1}(2\pi)^{-m(s+k-(m+1)/2)}\prod_{j=1}^{m}\Gamma(s+k-j)\mathscr{D}_{s,M}(\chi),$$

$$\text{(8.6)}\qquad \mathscr{D}^{+}_{s,M}(\chi) = \frac{2\,i^{\delta}\Gamma(s)\cos(\pi(s-\delta)/2)}{(2\pi i)^s}\mathscr{D}^{-}_{s,M}(\chi),$$

where $\mathscr{D}_{s,M}(\chi)$ are the values of the distribution $\mathscr{D}_{s,M}$ given by (8.3) on the profinite group

$$G_S = \mathbf{Z}_S^{\times} = \prod_{q\in S}\mathbf{Z}_q^{\times} \xrightarrow{\sim} \mathrm{Gal}(\mathbf{Q}(S)/\mathbf{Q})$$

with $\mathrm{Gal}(\mathbf{Q}(S)/\mathbf{Q})$ being the Galois group of the maximal abelian extension of $\mathbf{Q}$ unramified outside S and ∞. In definitions (8.5)–(8.6) we use the notation χ_M for the Dirichlet character modulo M induced by χ with support $S(M)\subset S$.

Now we give Archimedean integral representation for the normalized distributions, cf. **[Pa14,** p. 111].

PROPOSITION. *Under the notation and assumption as above the following integral representations for the normalized distributions hold:*

$$\text{(8.7)}\quad \langle f, f\rangle_C\mathscr{D}^{c\pm}_{s,M}(\chi) = \gamma(M')\langle f_0^{\rho}|V(C), F^{c\pm}_{M'}(s,\chi_M)|W(CN_0)\rangle_{CN_0},$$

where

$$\text{(8.8)}\quad \begin{aligned} F^{c+}_{M'}(s,\chi_M) &= \sum_{A_m\ni>0}\ \sum_{C\hat{\xi}_0[h_1]+h_2=M'h} d^{c+}(s,h_1,h_2)e_m(hz),\\ F^{c-}_{M'}(s,\chi_M) &= \sum_{A_m\ni>0}\ \sum_{C\hat{\xi}_0[h_1]+h_2=M'h} d^{c-}(s,h_1,h_2)e_m(hz)\end{aligned}$$

are modular forms from $\mathscr{M}_k^m(CN_0,\psi)$ with cyclotomic Fourier coefficients given by

$$\text{(8.9)}\quad \begin{aligned} d^{c+}(s,h_1,h_2) &= \chi_M(\det h_1)\det h_1^{\nu}\det h_2^{(2s-1)/2}P(h_2,h,s)\\ &\times M(h,\overline{\chi}\chi_{\xi_0}\psi, s+m/2)\times C^{s}_{\overline{\psi}\chi}G(\overline{\psi}\chi)^{-1}(1-(\overline{\chi}\psi)^2c^{-2s})\\ &\times L_N^{+}(s,\omega_{\xi_0,h_2})\prod_{q|N_0}\left\{(1-(\overline{\psi}\chi)(q)q^{s-1})(1-(\overline{\chi}\psi)(q)q^{-s})^{-1}\right\}\\ &\times (q_0C)^{-m(2s+m)/4}\det\xi_0^{(s+k-1+\nu)/2}\det(q_0^{-1/2}\xi_0)^{(2\nu+m)/2};\end{aligned}$$

with $\omega_{\xi_0,h_2} = \overline{\chi}\chi_{\xi_0}\psi\xi_{h_2}$ a Dirichlet character modulo N_0;

$$\text{(8.10)}\quad \begin{aligned} d^{-}(s,h_1,h_2) &= \chi_M(\det h_1)\det h_1^{\nu}P(h_2,h,1-s)\\ &\times(1-(\chi\overline{\psi})^2(c)c^{2(s-1)})L_N^{-}(s,\overline{\chi}\chi_{\xi_0}\psi\xi_{h_2})M(h,\overline{\chi}\chi_{\xi_0}\psi, s+m/2)\\ &\times(q_0C)^{-m(2s+m)/4}\det\xi_0^{(s+k-1+\nu)/2}\det(q_0^{-1/2}\xi_0)^{(2\nu+m)/2};\end{aligned}$$

where $P(v, u, s) \in \mathbf{Q}[v_{ij}, u_{ij}; i \leq j]$ *denotes the polynomial explicitly given by formula* (4.32) *of* **[Pa14,** Chapter 2], *which is defined for all* $s \in \mathbf{Z}$, $1 \leq s \leq k-m-\nu$, $s \equiv \delta \pmod 2$, *does not depend on* f, M, χ, M', *and satisfies the property*

$$P(v, u, s) = \det v^{(k-\nu-m-s)/2} \ (\mathrm{mod}(u_{ij})), \tag{8.11}$$

and where

$$M(h, \chi\chi_{\xi_0}\psi, s) = \prod_{q \in P(h)} M_q(h, \chi\chi_{\xi_0}\psi(q), q^{-s}) \tag{8.12}$$

is the finite Euler product (3.44) *of* **[Pa14,** Chapter 2], *with the product being extended to all prime numbers* q *in the set* $P(h)$ *of primes of* N *and the elementary divisors of the matrix* h, *such that for all of these* q *we have* $M_q(h, t) \in \mathbf{Z}[t]$. *The summation in the inner sum of* (8.8) *is extended over all pairs* (h_1, h_2) *of integral matrices with the conditions*

$$h_1 \in \mathrm{M}_m^+(\mathbf{Z}), \quad h_2 > 0, \quad h_2 \in A_m, \quad C\hat{\xi}_0[h_1] + h_2 = M'h \tag{8.13}$$

(*i.e.,* h_1 *is an integral matrix with positive determinant, not necessarily symmetric,* $h_2 \in A_m$ *is a positive-definite half-integral matrix, and* $C\hat{\xi}_0[h_1]$ *denotes the matrix given by* $C^t h_1 q_0 \xi^{-1} h_1 = C q_0 \xi_0^{-1}[h_1]$).

Now we notice that according to (8.9) and (8.10) the coefficients

$$d^-(s, h_1, h_2) = d^-(\chi_M; s, h_1, h_2)$$

do not depend on modulus M and they define for fixed h_1, h_2 a distribution on $G_S = \mathbf{Z}_S^\times$ with values in $\mathbf{Q}^{\mathrm{ab}}$; these distributions will also be denoted by $d^-(s, h_1, h_2)$. As we shall see these distributions turn out to be bounded measures, and the measures of Theorem 7.3 will be expressed in terms of them.

8.4. Let us consider the $\mathbf{C}$-linear functional

$$\mathscr{L}_f : g \mapsto \frac{\langle f^\rho, g|W(N)\rangle_N}{\langle f, f\rangle_N}, \tag{8.14}$$

on the vector space $\mathscr{M}_k^m(N, \psi)$ defined in **[Pa14]** by (3.51). From the explicit description of $\mathscr{L}_f$ given in **[Pa14,** 3.15] it follows that for an arbitrary

$$g = \sum_{h \geq 0} b(h) e_m(hz) \in \mathscr{M}_k^m(N, \psi)$$

there exist positive matrices $h_1, h_2, \ldots, h_t \in A_m$ and algebraic numbers $\alpha_1, \alpha_2, \ldots, \alpha_t \in \mathbf{Q}(f, \Lambda, \psi)$ from the field $\mathbf{Q}(f, \Lambda, \psi)$ generated by the Fourier coefficients of f and values of the homomorphism Λ and the character ψ such that for all $g \in \mathscr{M}_k^m(N, \psi)$ we have that

$$\mathscr{L}_f(g) = \sum_i \alpha_i b(h_i). \tag{8.15}$$

According to Proposition 4.5 of Chapter 2 of **[Pa14]**, the values of the distributions $\mathscr{D}_s^{c\pm}$ can be represented in terms of the linear form $\mathscr{L}$ as follows:

$$\mathscr{D}_s^{c\pm}(\chi) = \gamma(M')\mathscr{L}(F_{M'}^{c\pm}(s, \chi_M)). \tag{8.16}$$

Combining (8.15) and (8.16) we see that

$$\mathscr{D}_s^{c\pm}(\chi) = \sum_i \alpha_i v^{c\pm}(M'h^{(i)}, s, \chi_M), \tag{8.17}$$

where

$$v^{c\pm}(M'h^{(i)}, s, \chi_M) = \sum_{C\xi_0[h_1]+h_2=M'h} d^{c\pm}(s, h_1, h_2) \tag{8.18}$$

are the Fourier coefficients of the functions $F_{M'}^{c\pm}(s, \chi_M)$, and

$$\gamma(M') = 2^{m(2k-2-m-\kappa)} i^{-m(k-m/2-\nu)} a_0(\xi_0)^{-1} \alpha_0(M_0^{\tilde{m}-1} M'^{-1}) \times (CM_0^{\tilde{m}-1})^{(k-1-m)/2}.$$

In order to prove p-adic properties of the normalized distributions we recall first some p-adic integral representations for the Fourier coefficients $v^{c\pm}(M'h^{(i)}, s, \chi_M)$ by means of Mazur's measure.

Let $\omega \bmod A$ be a fixed primitive Dirichlet character such that $(A, M_0) = 1$ with $M_0 = \prod_{q\in S} q$. Put $\overline{S} = S \cup S(A)$, $\overline{M} = \prod_{q\in\overline{S}} q$. Then for any positive integer c with $(c, \overline{M}) = 1$, $c > 1$, there exist $\mathbf{C}_p$-measures $\mu^+(c, \omega)$, $\mu^-(c, \omega)$ on $\mathbf{Z}_S^\times$ which are uniquely determined by the following conditions: for $s \in \mathbf{Z}$, $s > 0$,

$$i_p^{-1}\left(\int_{\mathbf{Z}_S^\times} \chi x_p^s \, d\mu^+(c, \omega)\right) = (1 - \overline{\chi}\omega(c)c^{-s}) \frac{C_{\omega\overline{\chi}}}{G(\omega\overline{\chi})} \times \prod_{q\in S\setminus S(\chi)} \left\{\frac{1-\chi\overline{\omega}(q)q^{s-1}}{1-\overline{\chi}\omega(q)q^{-s}}\right\}, L_{M_0}^+(s, \overline{\chi}\omega), \tag{8.19}$$

and for $s \in \mathbf{Z}$, $s \leq 0$,

$$i_p^{-1}\left(\int_{\mathbf{Z}_S^\times} \chi x_p^s \, d\mu^-(c, \omega)\right) = (1 - \chi\overline{\omega}(c)c^{s-1}) L_{M_0}^-(s, \overline{\chi}, \omega), \tag{8.20}$$

where

$$L_{M_0}^+(s, \overline{\chi}\omega) = L_{\overline{M}}(s, \overline{\chi}\omega) 2i^\delta \frac{\Gamma(s)\cos(\pi(s-\delta)/2)}{(2\pi)^s} \tag{8.21}$$

$$L_{M_0}^+(s, \overline{\chi}\omega) = L_{\overline{M}}(s, \overline{\chi}\omega) \tag{8.22}$$

are the normalized Dirichlet L-functions with $\delta = 0, 1$, $(-1)^\delta = \overline{\chi}\omega(-1)$. The functions (8.19) and (8.20) satisfy the functional equation

$$L_{M_0}(1-s, \chi\overline{\omega}) = \prod_{q\in S\setminus S(\chi)} \left\{\frac{1-\chi\overline{\omega}(q)q^{s-1}}{1-\overline{\chi}\omega(q)q^{-s}}\right\} L_{M_0}^+(s, \overline{\chi}\omega).$$

Recall that, by the definition of the $\overline{S}$-adic Mazur measure μ^c on $\mathbf{Z}_{\overline{S}}^\times$, (8.19) and (8.20) are given by

$$\int_{\mathbf{Z}_S^\times} x\, d\mu^-(c, \omega) = \int_{\mathbf{Z}_{\overline{S}}^\times} x^{-1}\omega\, d\mu^c,$$

$$\int_{\mathbf{Z}_S^\times} d\mu^+(c, \omega) = \int_{\mathbf{Z}_{\overline{S}}^\times} x\, x_p^{-1}\omega^{-1}\, d\mu^c,$$

where $x \in X_S$, and $X_{\overline{S}}$ is regarded as a subgroup of $X_{\overline{S}}$.

8.5. Let us consider the $\mathbf{C}_p$-linear forms $\mathscr{D}^{C\pm} : \mathscr{C}^{2k-2m}(\mathrm{Gal}_S) \to \mathbf{C}_p$ defined by the conditions

$$\int_{a+(M)} x_p^r\, d\mathscr{D}^{c\pm} = \int_{a+(M)} d\mathscr{D}^{c\pm}_{-k+m+r},$$

for $r = 0, 1, \dots, 2k-2m-1$.

In order to prove (iv) of 7.3 we need to verify the following growth condition for $\mathscr{D}^c$:

$$\sup_{a\in\mathrm{Gal}_S}\left|\int_{a+(m)} (x-a)^r a\mathscr{D}^c\right|_p = o\left(|M|_p^{r-2\,\mathrm{ord}_p\,\alpha_0(p)}\right), \tag{8.23}$$

for $r = 0, 1, \dots, 2k-2m-1$.

From the above discussion and the Archimedean integral representations (8.16) and (8.17) for the distribution $\mathscr{D}^{c\pm}_{-k+m+j}(\chi)$, we obtain

$$\begin{aligned}\int_{a+(M)} (x-a)^r\, d\mathscr{D}^{c\pm} &= \sum_{j=0}^{r}\binom{r}{j}(-a)^{r-j}\int_{a+(M)} d\mathscr{D}^{c\pm}_{-k+m+j}\\ &= \sum_{j=0}^{r}\binom{r}{j}(-a)^{r-j}\frac{1}{\varphi(M)}\sum_{\chi \bmod M}\chi^{-1}(a)\mathscr{D}^{c\pm}_{-k+m+j}(\chi)\\ &= \gamma(M')\sum_{j=0}^{r}\binom{r}{j}(-a)^{r-j}\frac{1}{\varphi(M)}\sum_{\chi \bmod M}\chi^{-1}(a)\mathscr{L}(F^{c\pm}_{M'}(s, \chi_M)).\end{aligned} \tag{8.24}$$

In order to prove (8.23) it suffices to check it for the number

$$A = \gamma(M')\sum_{j=0}^{r}\binom{r}{j}(-a)^{r-j}\frac{1}{\varphi(M)}\sum_{\chi \bmod M}\chi^{-1}(a)v^{c\pm}(M'n, -k+m+j, \chi).$$

On the other hand, we know that $v^{c\pm}(\cdot)$ is congruent modulo M'' ($M'|M''$) to a linear combination of summands of the form

$$\chi(\alpha)\alpha^{j-k+m}\int_{\mathbf{Z}_{S(N_0)}^\times} \chi x_p^{j-k+m}\, d\mu^\pm(c, \cdot)$$

and this is equal to

$$\int_{\mathbf{Z}_{S(N_0)}^\times} \chi(\alpha x)(\alpha x)^{j-k+m}\, d\mu^\pm(c, \cdot),$$

where $\alpha \in \mathbf{Z}_{S(N_0)}^{\times}$. In this way we have

$$A = \gamma(M')\sum_{j=0}^{r}\binom{r}{j}(-a)^{r-j}\int_{\mathbf{Z}_{S(N_0)}^{\times}}\frac{1}{\varphi(M)}\sum_{\chi \bmod M}\chi^{-1}(a)\chi(\alpha x)(\alpha x)^{j-k+m}\,d\mu^{\pm}(c,\cdot)$$
$$= \gamma(M')\int_{x\equiv a\alpha^{-1} \bmod M}\sum_{j=0}^{r}\binom{r}{j}(-a)^{r-j}(\alpha x)^{j-k+m}\,d\mu^{\pm}(c,\cdot)$$
$$= \gamma(M')\alpha^{r-k+m}\int_{x\equiv a\alpha^{-1} \bmod M}x^{1-k}(x-a\alpha^{-1})^{r}\,d\mu^{\pm}(c,\cdot).$$

Since $x^{m-k}\mu^{\pm}(c,\cdot)$ is a bounded measure, the last integral has order $o(|M|_p^r)$. Now we choose $M' = M^2$; then from the definition of $\gamma(M')$ we get $\gamma(M') = o(|M|_p^{-2\,\mathrm{ord}_p\alpha(p)})$, establishing (iv) of 7.3.

Note that (v) of 7.3 follows from (iv) because of the uniqueness properties of h-admissible measures (see §4)) which imply the uniqueness of the non-Archimedean standard zeta functions. Statement (iii) is essentially the content of the main theorem in **[Pa14**, Theorem B, p. 4].

The equality (i) also follows from Proposition 2.3 of **[Pa14**, p. 89] in the case of the S-complete conductor, and for arbitrary χ it is proved by a direct calculation of the corresponding p-adic integral

$$\int_{\mathrm{Gal}_S}\chi\,d\mathscr{D}_{-k+m+j}^{c\pm}$$

as in **[Pa10]**.

Appendix. p-adic families of Galois representations attached to motives, and their L-functions

Hida **[Hi2]** has constructed interesting families of Galois representations of the type

$$\rho_p : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_p[[T]]), \qquad G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}),$$

which are nonramified outside p. These representations have the following property: if we consider the homomorphisms

$$\mathbf{Z}_p[[T]] \xrightarrow{s_k} \mathbf{Z}_p, \qquad 1+T \mapsto (1+p)^{k-1},$$

then we obtain a family of Galois representations

$$\rho_p^{(k)} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_p),$$

which is parametrized by $k \in \mathbf{Z}$, and for $k = 2, 3, \ldots$, these representations are equivalent over $\mathbf{Q}_p$ to the p-adic representations of Deligne attached to modular forms of weight k. This means that the representations of Hida are obtained by the p-adic interpolation of Deligne's representations. A geometric interpretation of Hida's representations was given by Mazur and Wiles **[Maz–W3]**, cf. **[Maz3]**. For example, for the modular form Δ of weight 12 Hida has constructed a representation

$$\rho_{p,\Delta} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_p[[T]]),$$

as an example of his general theory, where the prime number p has the property $\tau(p) \not\equiv 0 \pmod p$ (e.g., $p < 2041$, $p \neq 2, 3, 5$, and 7).

Also, Hida has generalized his construction to the case of Hilbert modular forms.

In this appendix we would like to describe a conjectural generalization of his construction to arbitrary motives, and to formulate a certain conjecture. The important case, in which this conjecture can be verified, corresponds to Hecke characters of CM-type. We describe certain p-adic families, which are in general bigger than those obtained by the cyclotomic twist. We stress the fact that the corresponding p-adic L-functions depend analytically on the parameter of these p-adic Galois representations.

As above we fix embeddings $i_\infty : \overline{\mathbf{Q}} \to \mathbf{C}$, $i_p : \overline{\mathbf{Q}} \to \mathbf{C}_p$ and we shall often regard algebraic numbers (via these embeddings) as both complex and p-adic numbers, where $\mathbf{C}_p$ is the Tate field.

A1. Motives and L-functions over a totally real field. Let F be a totally real number field of degree $n = [F : \mathbf{Q}]$ and let T be another number field which is assumed to be a subfield of $\mathbf{C}$. By a motive M over F of with coefficients in T we shall mean a collection of objects $M_{B,\sigma} = M_\sigma$, M_{DR}, M_λ, $I_{\infty,\sigma}$, $I_{\lambda,\sigma}$, where σ runs over the set J_F of all complex embeddings of F,

M_σ is the Betti realization of M (with respect to the embedding $\sigma \in J_F$) which is a vector space over T of dimension d endowed for real $\sigma \in J_F$ with a T-rational involution ρ_σ;

M_{DR} is the de Rham realization of M, a free $(T \otimes F)$-module of rank d, endowed with a decreasing filtration $\{F^i_{\mathrm{DR}}(M) \subset M_{\mathrm{DR}} \mid i \in \mathbf{Z}\}$ of $(T \otimes F)$-modules (which may not be free in some cases when $F \neq \mathbf{Q}$);

M_λ is the λ-adic realization of M at a finite place λ of the coefficient field T (a T_λ-vector space of degree d over T_λ, a completion of T at λ) which is a Galois module over $G_F = \mathrm{Gal}(\overline{F}/F)$ so that we have a compatible system of λ-adic representations denoted by

$$r_{M,\lambda} = r_\lambda : G_F \to \mathrm{GL}(M_\lambda).$$

Also,

$$I_{\infty,\sigma} : M_\sigma \otimes_T \mathbf{C} \to M_{\mathrm{DR}} \otimes_{\sigma(F)\cdot T} \mathbf{C}$$

is the complex comparison isomorphism of complex vector spaces for each $\sigma \in J_F$, and

$$I_{\lambda,\sigma} : M_\sigma \otimes_T T_\lambda \to M_\lambda$$

is the λ-adic comparison isomorphism of T_λ-vector spaces. It is assumed in the notation that the complex vector space $M_\sigma \otimes_{\mathbf{Q}} \mathbf{C}$ is decomposed in the

Hodge bigraduation

$$M_\sigma \otimes_T \mathbf{C} = \bigoplus_{i,j} M_\sigma^{i,j}$$

in which $\rho_\sigma(M_\sigma^{i,j}) \subset M_\sigma^{j,i}$ for $\sigma \in J_F$ and the Hodge numbers

$$h_\sigma(i, j) = h_\sigma(i, j, M) = \dim_{\mathbf{C}} M_\sigma^{i,j}$$

do not depend on σ. Moreover,

$$I_{\infty,\sigma}\left(\bigoplus_{i'\geq i} M_\sigma^{i',j}\right) = F_{\mathrm{DR}}^i(M) \otimes_{F,\sigma} \mathbf{C}.$$

Also, $I_{\lambda,\sigma}$ takes ρ_σ to the r_λ-image of the Galois automorphism which is denoted by the same symbol $\rho_\sigma \in G_F$ and corresponds to the complex conjugation of $\mathbf{C}$ under an embedding of $\overline{F}$ to $\mathbf{C}$ extending σ. We assume that M is pure of weight w (i.e., $i + j = w$).

The L-function $L(M, s)$ of M is defined as the following Euler product:

$$L(M, s) = \prod_{\mathfrak{p}} L_{\mathfrak{p}}(M, \mathcal{N}\mathfrak{p}^{-s}),$$

extended over all maximal ideals $\mathfrak{p}$ of the maximal order $\mathcal{O}_F$ of F and where

$$\begin{aligned} L_{\mathfrak{p}}(M, X)^{-1} &= \det(1 - X \cdot r_\lambda(\mathrm{Fr}_{\mathfrak{p}}^{-1}) \mid M_\lambda^{I_{\mathfrak{p}}}) \\ &= (1 - \alpha^{(1)}(\mathfrak{p})X) \cdot (1 - \alpha^{(2)}(\mathfrak{p})X) \cdot \cdots \cdot (1 - \alpha^{(d)}(\mathfrak{p})X) \\ &= 1 + A_1(\mathfrak{p})X + \cdots + A_d(\mathfrak{p})X^d; \end{aligned}$$

here $\mathcal{N}\mathfrak{p}$ is the norm of $\mathfrak{p}$ and $\mathrm{Fr}_{\mathfrak{p}} \in G_F$ is the Frobenius element at $\mathfrak{p}$, defined modulo conjugation and modulo the inertia subgroup $I_{\mathfrak{p}} \subset G_{\mathfrak{p}} \subset G_F$ of the decomposition group $G_{\mathfrak{p}}$ (of any extension of $\mathfrak{p}$ to $\overline{F}$). We make the standard hypothesis that the coefficients of $L_{\mathfrak{p}}(M, X)^{-1}$ belong to T, and that they are independent of λ coprime to $\mathcal{N}\mathfrak{p}$. Therefore we can and we shall regard this polynomial both over $\mathbf{C}$ and over $\mathbf{C}_p$. We shall need the following twist operation: for an arbitrary motive M over F with coefficients in T, an integer m, and a Hecke character χ of finite order, one can define the twist $N = M(m)(\chi)$, which is again a motive over F with the coefficient field $T(\chi)$ of the same rank d and weight w so that we have

$$L(N, s) = \prod_{\mathfrak{p}} L_{\mathfrak{p}}(M, \chi(\mathfrak{p})\mathcal{N}\mathfrak{p}^{-s-n}).$$

A2. The group of Hida and the algebra of Iwasawa–Hida. Now let us fix a motive M with coefficients in $T = \mathbf{Q}(\langle a(\mathfrak{n})_{\mathfrak{n}}\rangle)$ of rank d and of weight w, and let $\mathrm{End}_T M$ denote the endomorphism algebra of M (i.e., the algebra of T-linear endomorphisms of any M_σ, which commute with the Galois action of $G_{K'}$ under the comparison isomorphisms, where K' is an appropriate final extension of K). Let

$$\mathrm{Gal}_p = \mathrm{Gal}(F^{\mathrm{ab}}_{p,\infty}/F)$$

denote the Galois group of the maximal abelian extension $F^{\mathrm{ab}}_{p,\infty}$ of F unramified outside primes of F above p and ∞. Define $\mathscr{O}_{F,T,p} = \mathscr{O}_F \otimes \mathscr{O}_T \otimes \mathbf{Z}_p$.

DEFINITION. The *group of Hida* $\mathrm{GH}_M = \mathrm{GH}_{M,p}$ is the product

$$\mathrm{GH}_M = \mathrm{End}_T M^\times(\mathscr{O}_{F,T,p}) \times \mathrm{Gal}_p,$$

where $\mathrm{End}_T M^\times$ denotes the algebraic T-group of invertible elements of $\mathrm{End}_T M$ (it is implicitly supposed that the group $\mathrm{End}_T M^\times$ possesses an $\mathscr{O}_T$-integral structure given by an appropriate choice of an $\mathscr{O}_T$-lattice).

Consider next the $\mathbf{C}_p$-analytic Lie group

$$\mathscr{X}_{M,p} = \mathrm{Hom}_{\mathrm{contin}}(\mathrm{GH}_M, \mathbf{C}_p^\times)$$

consisting of all continuous characters of the Hida group GH_M, which contains the $\mathbf{C}_p$-analytic Lie group

$$\mathscr{X}_p = \mathrm{Hom}_{\mathrm{contin}}(\mathrm{Gal}_p, \mathbf{C}_p^\times)$$

consisting of all continuous characters of the Galois group Gal_p (via the projection of GH_M onto Gal_p.

The group $\mathscr{X}_{M,p}$ contains the discrete subgroup $\mathscr{A}$ of arithmetical characters of the type

$$\chi \cdot \eta \cdot \mathscr{N} x_p^m = (\chi, \eta, m),$$

where $\chi \in \mathscr{X}^{\mathrm{tors}}_{M,p}$ is a character of finite order of GH_M, η is a T-algebraic character of $\mathrm{End}_T M^\times(\mathscr{O}_{F,T,p})$, $m \in \mathbf{Z}$, and $\mathscr{N} x_p$ denotes the following natural norm homomorphism

$$\mathscr{N} x_p : \mathrm{Gal}_p \to \mathrm{Gal}(\mathbf{Q}^{\mathrm{ab}}_{p,\infty}/\mathbf{Q}) \cong \mathbf{Z}_p^\times \to \mathbf{C}_p^\times, \qquad \mathscr{N} x_p \in \mathscr{X}_p.$$

DEFINITION. The *algebra of Iwasawa–Hida* $I_M = I_{M,p}$ of M at p is the completed group ring $\mathscr{O}_p[[GH_M]]$, where $\mathscr{O}_p$ denotes the ring of integers of the Tate field $\mathbf{C}_p$.

Note that this definition is completely analogous to the usual definition of the Iwasawa algebra Λ as the completed group ring $\mathbf{Z}_p[[\mathbf{Z}_p]]$ if we take into account that $\mathbf{Z}_p$ coincides with the factor group of $\mathbf{Z}_p^\times$ modulo its torsion subgroup.

Now for each arithmetic point $P = (\chi, \eta, m) \in \mathscr{A}$ we have a homomorphism $\nu_P : GH_M \to \mathscr{O}_p$ defined by the corresponding group homomorphism $P : GH_M \to \mathscr{O}_p^\times \subset \mathbf{C}_p^\times$.

For an I_M-module N and $P \in \mathscr{A}$ we put

$$N_P = N \otimes_{I_M, \nu_P} \mathscr{O}_p$$

("reduction of N modulo P", or a fiber of N at P).

Therefore, for each Galois representation $r_N : G_F \to \mathrm{GL}(N)$ its reduction $r_{N_P} = r \bmod P$ is defined as the natural composition: $G_F \to \mathrm{GL}(N) \to \mathrm{GL}(N_P)$.

A3. A conjecture on the existence of p-adic families of Galois representations attached to motives. Note first that the fixed embeddings $T \hookrightarrow \mathbf{C}$, $i_\infty : \overline{\mathbf{Q}} \to \mathbf{C}$, $i_p : \overline{\mathbf{Q}} \to \mathbf{C}_p$ define a place $\lambda(p)$ of T attached to the corresponding composition $T \hookrightarrow \overline{\mathbf{Q}} \xrightarrow{i_p} \mathbf{C}_p$.

CONJECTURE I. *For every M over F of rank d with coefficients in T there exists a free I_M-module M_I of the same rank d, a Galois representation*

$$r_I : G_F \to \mathrm{GL}(M_I),$$

an infinite subset $\mathscr{A}' \subset \mathscr{A}$ of "positive" characters, and a distinguished point $P_0 \in \mathscr{A}$ such that

(a) *the reduced Galois representation*

$$r_{I,P_0} : G_F \to \mathrm{GL}(M_{I,P_0})$$

is equivalent over $\mathbf{C}_p$ to the $\lambda(p)$-adic representation $r_{M,\lambda(p)}$ of M at the distinguished place $\lambda(p)$;

(b) *for every $P \in \mathscr{A}'$ there exists a motive M_P over F of the same rank d such that its $\lambda(p)$-adic Galois representation is equivalent over $\mathbf{C}_p$ to the reduction*

$$r_{I,P} : G_F \to \mathrm{GL}(M_{I,P}).$$

We call the module M_I the *realization of Iwasawa* of M.

A4. A generalization of the Hasse invariant for a motive. We define the generalized *Hasse invariant* of a motive in terms of the *Newton polygons* and the *Hodge polygons* of a motive. Properties of these polygons are closely related to the notions of a p-ordinary and a p-admissible motive.

Now we are going to define the Newton polygon

$$P_{\mathrm{Newton},\sigma}(u) = P_{\mathrm{Newton},\sigma}(u, M)$$

and the Hodge polygon

$$P_{\mathrm{Hodge},\sigma}(u) = P_{\mathrm{Hodge},\sigma}(u, M)$$

attached to M, σ. First for $\mathfrak{p} = \mathfrak{p}(\sigma)$ we consider (using i_∞) the local $\mathfrak{p}$-polynomial

$$\begin{aligned} L_\mathfrak{p}(M, X)^{-1} &= 1 + A_1(\mathfrak{p})X + \cdots + A_d(\mathfrak{p})X^d \\ &= (1 - \alpha^{(1)}(\mathfrak{p})X) \cdot (1 - \alpha^{(2)}(\mathfrak{p})X) \cdot \cdots \cdot (1 - \alpha^{(d)}(\mathfrak{p})X), \end{aligned}$$

and we assume that its inverse roots are indexed in such a way that

$$\mathrm{ord}_p \alpha^{(1)}(\mathfrak{p}) \le \mathrm{ord}_p \alpha^{(2)}(\mathfrak{p}) \le \cdots \le \mathrm{ord}_p \alpha^{(d)}(\mathfrak{p}).$$

DEFINITION. The *Newton polygon* $P_{\text{Newton},\sigma}(u)$ $(0 \le u \le d)$ of M at $\mathfrak{p} = \mathfrak{p}(\sigma)$ is the convex hull of the points $(i, \text{ord}_p A_i(\mathfrak{p}))$ $(i = 0, 1, \ldots, d)$.

The important property of the Newton polygon is that the length of the horizontal segment of slope $i \in \mathbf{Q}$ is equal to the number of the inverse roots $\alpha^{(j)}(\mathfrak{p})$ such that $\text{ord}_p \alpha^{(j)}(\mathfrak{p}) = i$ (note that i may not necessarily be an integer, but this will be the case for the p-ordinary motives below).

The *Hodge polygon* $P_{\text{Hodge},\sigma}(u)$ $(0 \le u \le d)$ of M at σ is defined using the Hodge decomposition of the d-dimensional $\mathbf{C}$-vector space

$$M_\sigma = M_\sigma \otimes_T \mathbf{C} = \bigoplus_{i,j} M_\sigma^{i,j}$$

where $M_\sigma^{i,j}$ is a $\mathbf{C}$-subspace. Note that the dimension $h_\sigma(i, j) = \dim_{\mathbf{C}} M_\sigma^{i,j}$ may depend on σ.

DEFINITION. The *Hodge polygon* $P_{\text{Hodge},\sigma}(u)$ is a function $[0, d] \to \mathbf{R}$ whose graph consists of segments passing through the points

$$(0, 0), \ldots, \left(\sum_{i' \le i} h_\sigma(i', j), \sum_{i' \le i} h_\sigma i' h_\sigma(i', j)\right),$$

so that the length of the horizontal segment of the slope $i \in \mathbf{Z}$ is equal to the dimension $h_\sigma(i, j)$.

Now we recall the definition of a p-ordinary motive in the simplest case $F = T = \mathbf{Q}$ (see **[Co, Co–PeRi]**). We assume that M is pure of weight w and rank d. Let G_p be the decomposition group (of the place $\lambda(p)$ in T over p) and $\psi_p : G_p \to \mathbf{Z}_p^\times$ be the cyclotomic character of G_p. Then M is called p ordinary at p if the following conditions are satisfied:

(i) the inertia group $I_p \subset G_p$ acts trivially on each of the ℓ-adic realizations M_ℓ for $\ell \ne p$;

(ii) there exists a decreasing filtration $F_p^i V$ on $V = M_p = M_B \otimes \mathbf{Q}_p$ of $\mathbf{Q}_p$-subspaces that are stable under the action of G_p such that for all $i \in \mathbf{Z}$ the group G_p acts on $F_p^i V / F_p^{i+1} V$ via some power of the cyclotomic character, say $\psi_p^{-e_i}$. Then $e_1(M) \ge \cdots \ge e_t(M)$ and the following properties take place:

(a) $\dim_{\mathbf{Q}_p} F_p^i V / F_p^{i+1} V = h(e_i, w - e_i)$;

(b) The Hodge polygon and the Newton polygon of M coincide:

$$P_{\text{Newton}}(u) = P_{\text{Hodge}}(u).$$

If, furthermore, M is critical at $s = 0$, then it is easy to verify that the number d_p of the inverse roots $\alpha^{(j)}(p)$ with $\text{ord}_p \alpha^{(j)}(p) < 0$ is equal to $d^+ = d^+(M)$ of M_σ^+.

In the general case (of a motive M over F with coefficients in T) the notion of a p-ordinary motive can be defined using the restriction of the ground field F to $\mathbf{Q}$ and the restriction of the coefficient field T to $\mathbf{Q}$ (the last operation corresponds to forgetting of the T-module structure on the realizations of M). In this way we get a motive M' over $\mathbf{Q}$ with coefficients in $\mathbf{Q}$ of the same weight w and rank $\mathrm{rk}(M') = [F:\mathbf{Q}][T:\mathbf{Q}]\cdot d$.

However, it turns out that the notion of a p-ordinary motive is too restrictive, and we introduce the following weaker version of it.

DEFINITION. The motive M over F with coefficients in T is called *admissible at* p if, for all $\sigma \in J_F$,

$$P_{\mathrm{Newton},\sigma}(d_\sigma^+) = P_{\mathrm{Hodge},\sigma}(d_\sigma^+).$$

Here $d_\sigma^+ = d_\sigma^+(M)$ is the dimension of M_σ, $\sigma \in J_F$.

In the general case we use the following vector quantity $h = (h_\sigma)_\sigma$ which is defined in terms of the difference between the Newton polygon and the Hodge polygon of M:

$$h_\sigma = P_{\mathrm{Newton},\sigma}(d_\sigma^+) - P_{\mathrm{Hodge},\sigma}(d_\sigma^+).$$

We call the vector $h = h(M) = (h_\sigma)_\sigma$ the *Hasse invariant* of M at p. Note the following important properties of the quantity h:

(i) $h = h(M)$ does not change if we replace M by its Tate twist.

(ii) $h = h(M)$ does not change if we replace M by its twist $M = M(\chi)$ with a Hecke character χ of finite order whose conductor is prime to $\mathfrak{p}$.

(iii) $h = h(M)$ does not change if we replace M by its dual $M^\vee$.

In the next section we state in terms of this quantity a general conjecture on p-adic L-functions.

A5. A conjecture on the existence of certain families of p-adic L-functions. We are going to describe families of p-adic L-functions as certain analytic functions on the total analytic space, the $\mathbf{C}_p$-analytic Lie group

$$\mathscr{X}_{M,p} = \mathrm{Hom}_{\mathrm{contin}}(\mathrm{GH}_M,\ \mathbf{C}_p^\times),$$

which contain the $\mathbf{C}_p$-analytic Lie subgroup (the cyclotomic line) $\mathscr{X}_p \subset \mathscr{X}_{M,p}$:

$$\mathscr{X}_p = \mathrm{Hom}_{\mathrm{contin}}(\mathrm{Gal}_p,\ \mathbf{C}_p^\times).$$

In order to do this we need a modified L-function of a motive over F. Following Coates this modified L-function has a form appropriate for further use in the p-adic construction. First we multiply $L(M, s)$ by an appropriate factor at infinity and define

$$\Lambda_{(\infty)}(M,\ s) = E_\infty(M,\ s)L(M, s)$$

as $\Lambda_{(\infty)}(\tau, R_{F/\mathbf{Q}}M, \rho, s)$ in the notation of Coates **[Co]** with $\rho = i$ so that $E_\infty(M, s) = E_\infty(\tau, R_{F/\mathbf{Q}}M, \rho, s)$ is the modified Γ-factor at infinity which actually does not depend on the fixed embedding τ of T into $\mathbf{C}$. Also we put

$$\Omega^\nu(M) = (\Omega^\nu(M)^{(\tau)}) = c^\nu(RM)(2\pi i)^{r(RM)} \in (T\otimes\mathbf{C})^\times$$

where $\nu = (-1)^m$,

$$r(RM) = \sum_{j<0} jh(i, j, R_{F/\mathbf{Q}}M) = \sum_{j<0} jh(i, j, M),$$

$n = [F : \mathbf{Q}]$, $c^\nu(RM) = c^\nu(R_{F/\mathbf{Q}}M)$ is the period of $R_{F/\mathbf{Q}}M$. Note that the quantity $r(M)$ has a natural geometric interpretation as the minimum of the Hodge polygon $P_{\mathrm{Hodge}}(M)$.

We define

$$\begin{aligned}\Lambda_{(p,\infty)}&(M(m)(\chi), s)\\ &= (G(\chi)^{-1}D_F^{1/2})^{d^{\epsilon_0}(M(m)(\chi))} \prod_{\mathfrak{p}|p} A_\mathfrak{p}(M(m)(\chi), s) \cdot \Lambda_{(\infty)}(M(m)(\chi), s),\end{aligned}$$

where

$$A_\mathfrak{p}(M(\chi), s) = \begin{cases} \displaystyle\prod_{i=d^++1}^{d} (1 - \chi(\mathfrak{p})\alpha^{(i)}(\mathfrak{p})\mathcal{N}\mathfrak{p}^{-s}) \prod_{i=1}^{d^+}(1 - \chi^{-1}(\mathfrak{p})\alpha^{(i)}(\mathfrak{p})^{-1}\mathcal{N}\mathfrak{p}^{s-1}) & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } \mathfrak{p} \nmid c(\chi), & \\ \displaystyle\prod_{i=1}^{d^+} \left(\frac{\mathcal{N}^{\mathfrak{p}}}{\alpha^{(i)}(\mathfrak{p})}\right)^{\mathrm{ord}_\mathfrak{p} c(\chi)}, \quad \text{otherwise.} & \end{cases}$$

Let $\mathscr{A}$ be the discrete subgroup $\mathscr{A}$ of arithmetical characters,

$$\chi \cdot \eta \cdot \mathcal{N} x_p^m = (\chi, \eta, m) \in \mathscr{A},$$

$\mathscr{A}' \subset \mathscr{A}$ the subset of "positive" characters, and $P_0 \in \mathscr{A}$ a distinguished point of Conjecture I. Here the subset of "positive" characters $\mathscr{A}'$ is a certain infinite subset of the lattice of algebraic characters of the algebraic group $\mathrm{End}_T M^\times = \mathrm{End}_T M^\times$ defined by an appropriate choice of a "positive" basis of the lattice. Let $\mathscr{A}'' \subset \mathscr{A}'$ be the subset of critical elements, which consists of those P for which the corresponding motives M_P are critical (at $s = 0$).

Now we are ready to formulate

CONJECTURE II. *For a canonical choice of periods $\Omega(P) \in \mathbf{C}^\times$ for $P \in \mathscr{A}''$ there exists a $\mathbf{C}_p$-meromorphic function*

$$\mathscr{L}_M : \mathscr{X}_{M,p} \to \mathbf{C}_p$$

with the properties:

(i)

$$\mathscr{L}_M(P) = \frac{\Lambda_{p,(\infty)}(M(m)(\chi), 0)}{\Omega(P)}$$

for almost all $P \in \mathscr{A}''$.

(ii) *For arithmetic points of type*

$$P = (\chi, \eta, m) \in \mathscr{A}''$$

with η *fixed there exists a finite set* $\Xi \subset \mathscr{X}_{M,p}$ *of p-adic characters and positive integers* $n(\xi)$ *(for* $\xi \in \Xi$*) such that, for any* $g_0 \in \mathrm{Gal}_p$*, the function*

$$\prod_{\xi \in \Xi} (x(g_0) - \xi(g_0))^{n(\xi)} \mathscr{L}_M(x \cdot P)$$

is holomorphic on $\mathscr{X}_p$.

(iii) *For arithmetic points of type*

$$P = (\chi, \eta, m) \in \mathscr{A}''$$

with η *fixed the function in* (ii) *is bounded if and only if the Hasse invariant* $h(P) = h(M_P)$ *vanishes.*

(iv) *In the general case the function* $\mathscr{L}_M(P \cdot x)$ *of* $x \in \mathscr{X}_p$ *is of logarithmic growth type* $o(\log \mathscr{N}(\cdot)^{h}{}_0$ *with* $h_0 = [\max_\sigma h_\sigma] + 1$.

Don Blasius has suggested to us the following modification of Conjecture II, based on the theorem of Katz on p-adic L-functions of CM-type, and on the theory of p-adic periods **[Bl3]**.

CONJECTURE II$'$. *There exists a certain choice of complex periods* $\Omega_\infty(P) \in \mathbf{C}^\times$ *and p-adic periods* $\Omega_p(P) \in \mathbf{C}_p^\times$ *for all* $P \in \mathscr{A}''$ *such that "the ratio"* $\Omega_\infty(P)/\Omega_p(P)$ *is canonically defined, and there exists a* $\mathbf{C}_p$*-meromorphic function* $\mathscr{L}_M : \mathscr{X}_{M,p} \to \mathbf{C}_p$ *with the properties*:

(i)

$$\frac{L_M}{\Omega_p(P)} = \frac{\Lambda_{p,(\infty)}(M(m)(\chi), 0)}{\Omega(P)}$$

for almost all $P \in \mathscr{A}''$.

(ii) *For arithmetic points of type* $P = (\chi, \eta, m) \in \mathscr{A}''$ *with* η *fixed there exists a finite set* $\Xi \subset \mathscr{X}_{M,p}$ *of p-adic characters and positive integers* $n(\xi)$ *(for* $\xi \in \Xi$*) such that, for any* $g_0 \in \mathrm{Gal}_p$*, the function*

$$\prod_{\xi \in \Xi} (x(g_0) - \xi(g_0))^{n(\xi)} \mathscr{L}_M(x \cdot P)$$

is holomorphic on $\mathscr{X}_p$.

(iii) *For arithmetic points of type* $P = (\chi, \eta, m) \in \mathscr{A}''$ *with* η *fixed the function in* (ii) *is bounded if and only if the Hasse invariant* $h(P) = h(M_P)$ *vanishes.*

(iv) *In the general case the function* $\mathscr{L}_M(P \cdot x)$ *of* $x \in \mathscr{X}_p$ *is of logarithmic growth type* $o(\log \mathscr{N}(\cdot)^{h_0})$ *with* $h_0 = [\max_\sigma h_\sigma] + 1$.

A6. Example: Hecke characters of CM-type

A6.1. Let $K \supset F$ be a totally imaginary quadratic extension, and $\eta : \mathbf{A}_K^\times / K^\times \to \mathbf{C}^\times$ be an algebraic Hecke's Grössencharakter such that

$$\eta((\alpha)) = \left(\frac{\alpha^{\phi_1}}{|\alpha^{\phi_1}|}\right)^{w_1} \cdots \left(\frac{\alpha^{\phi_n}}{|\alpha^{\phi_n}|}\right)^{w_n} \cdot \mathscr{N}(\alpha)^{w_0/2-1}$$

for $\alpha \in K$, $\alpha \equiv 1 \pmod{\mathfrak{c}(\eta)}$, where $\Sigma = \{\sigma_i : K \to \mathbf{C}\}$ is a fixed CM-type of K, w_i are positive integers, $w_0 = \max_i w_i$. Then there exists a Hilbert modular form $\mathbf{f}$ of weight $k = (w_1 + 1, \ldots, w_n + 1)$ such that $L(s, \mathbf{f}) = L(s, \eta)$, and $M(\mathbf{f})$ coincides with the motive $M(\eta) = R_{K/F}[\eta]$ obtained by restriction of scalars from the motive $[\eta]$ (the last motive exists as an object of the category of motives of CM-type; see **[Bl1]**).

The Hodge structure of $M(\eta)_\sigma$ has the form

$$((w_0 - w_\sigma)/2, (w_0 + w_\sigma)/2) + ((w_0 + w_\sigma)/2, (w_0 - w_\sigma)/2).$$

Let

$$\mathfrak{p} = \mathfrak{p}_\sigma = \begin{cases} \mathfrak{P}\mathfrak{P}', & \text{if } \mathfrak{p} \text{ splits in } K, \\ \mathfrak{P}, & \text{if } \mathfrak{p} \text{ is inert in } K. \end{cases}$$

Then the local factor of $L(M(\eta), s)$ is given by

$$L_\mathfrak{p}(M(\eta), X)^{-1} = \begin{cases} (1 - \eta(\mathfrak{P})X)(1 - \eta(\mathfrak{P}')X), & \text{if } \mathfrak{p} \text{ splits in } K, \\ (1 - \eta(\mathfrak{P})^2), & \text{if } \mathfrak{p} \text{ is inert in } K. \end{cases}$$

Therefore the generalized Hasse invariant $h = (h_\sigma)_\sigma$ of $M(\eta)$ is given by

$$h_\sigma = \begin{cases} 0, & \text{if } \mathfrak{p} \text{ splits in } K, \\ w_0/2, & \text{if } \mathfrak{p} \text{ is inert in } K. \end{cases}$$

In the additive notation the type of η can be written in the following form:

$$\sum_\sigma \frac{w_\sigma}{2}(\sigma - \bar{\sigma}) + \frac{w_0}{2}\sum_\sigma(\sigma + \bar{\sigma}) = \sum_\sigma d_\sigma(\sigma - \bar{\sigma}) + m_0 \sum_\sigma \sigma,$$

where $m_0 = w_0$, $d_\sigma = (w_\sigma - w_0)/2$. Using a shift one sees that the point $s = m$ is critical for $L(s, \eta)$ iff $s = 0$ is critical for the character $\lambda(\mathfrak{a}) = \eta(\mathfrak{a})\mathcal{N}\mathfrak{a}^{-m}$.

Then one sees that the character η of the type $\sum_\sigma d_\sigma(\sigma - \bar{\sigma}) + m_0 \sum_\sigma \sigma$ is critical at 0 iff

$$(*) \qquad m_0, d_\sigma \geq 0 \quad \text{or} \quad m_0 \leq 1,\ d_\sigma \geq 1 - m_0 \qquad \text{for all } \sigma).$$

In order to state the thorem of Katz on p-adic L-functions of CM-fields (in a simplified form) we let $\mathfrak{C} \subset \mathcal{O}_K$ denote an integer ideal of the maximal order of K, and $G_\infty(\mathfrak{C})$ be the ray class group of K of conductor $\mathfrak{C}p^\infty$.

For each CM-type Σ one can canonically choose constants

$$\Omega_\infty = (\Omega_\infty(\sigma))_{\sigma \in \Sigma} \in (\mathbf{C}^\times)^n,$$
$$\Omega_p = (\Omega_p(\sigma))_{\sigma \in \Sigma} \in (\mathbf{C}_p^\times)^n$$

(complex and p-adic periods).

The theorem of Katz states that under the assumption $h = 0$ there exists a bounded p-adic measure μ on $G_\infty(\mathfrak{C})$ such that for all critical characters λ of conductor dividing $\mathfrak{C}p^\infty$ the value of the p-adic integral

$$\frac{\int_{G_\infty(\mathfrak{C})} \hat{\lambda}\, d\mu}{\Omega_p^{m_0\Sigma + 2d}},$$

essentially coincides with the normalized special value

$$\frac{\Lambda_{(p,\infty)}(\lambda, 0)}{\Omega_{\infty}^{m_0\Sigma+2d}},$$

where $\hat{\lambda}$ denotes the p-adic avatar of λ.

This theorem provides an example of a p-adic family of Conjectures I and II, because $\mathrm{End}_T(M)$ is essentially K^x, and by class field theory GH_M is related to $G_\infty(\mathfrak{C})$.

A6.2. Families of Hida of Hilbert modular forms. In this case we start from a motive $M(\mathbf{f})$ attached to a (general) Hilbert modular form and obtain essentially the group $\mathrm{GH}_{\mathbf{f}} = \mathcal{O}^{\times}_{F,T,p} \times \mathrm{Gal}_p$ whose characters parametrize "the weights" of Hilbert modular forms in the corresponding family.

Acknowledgment

This paper is an extended version of talks given during visits of the author to the universities Paris XIII and Grenoble I in 1991–92, and to the Mathematical Institute in Cologne (February 1992). It is a great pleasure for the author to express his deep gratitude to Professors D. Barsky, G. Christol, R. Gillard, and U. Jannsen for the hospitality, the support and very helpful discussions. The author is very graeteful to John Coates, who discussed the general conjecture of §6 in his talk in Seattle, and who suggested to us some essential improvments. We are pleased to regard the article as an illustration and some further development of his ideas.

REFERENCES

[An1] A. N. Andrianov, *On zeta functions of Rankin type associated with Siegel modular forms*, Lecture Notes in Math., vol. 627, Springer-Verlag, Berlin and New York, 1977, pp. 325–338.

[An2] ——, *On the decomposition of Hecke polynomials for the symplectic group of degree n*, Mat. Sb. **104** (1978), 291–341. (Russian)

[An3] ——, *Multiplicative arithmetic of Siegel modular forms*, Uspekhi Mat. Nauk **34** (1979), 67–135. (Russian)

[An4] ——, *Modular descent and the Saito–Kurokawa conjecture*, Invent. Math. **53** (1979), 267–280.

[An-K] A. N. Andrianov and V. L. Kalinin, *On analytic properties of standard zeta functions of Siegel modular forms*, Mat. Sb. **106** (1978), 323–339. (Russian)

[Am-V] Y. Amice and J. Velu, *Distributions p-adiques associées aux séries de Hecke*, Journées Arithmétiques de Bordeaux (Conf. Univ. Bordeaux, 1974), Astérisque no. 24/25, Soc. Math. France, Paris 1975, pp. 119–131.

[Ar] B. Arnaud, *Interpolation p-adique d'un produit de Rankin*, C. R. Acad. Sci. Paris Sér. I **299** (1984), 527–530.

[Bl1] D. Blasius, *On the critical values of Hecke L-series*, Ann. of Math. (2) **124** (1986), 23–63.

[Bl2] ——, *Appendix to Orloff critical values of certain tensor product L-functions*, Invent. Math. **90** (1987) 181–188.

[Bö1] S. Böcherer, *Über die Funktionalgleichung automorpher L-Funktionen zur Siegelscher Modulgruppe*, J. Reine Angew. Math. **362** (1985), 146–168.

[Bö2] ——, *Über die Fourier–Jacobi Entwickelung Siegelscher Eisensteinreihen*. I, II. Math. Z. **183** (1983), 21–46; **189** (1985), 81–100.

[Cass-N] P. Cassou-Nogues, *Valeurs aux entiers négatifs des fonctions zeta et fonctions zeta p-adiques*, Invent. Math. **51** (1979), 29–59.

[Co] J. Coates, *On p-adic L-functions*, Sém. Bourbaki, 40ème année, 1987–88, no. 701, Astérisque, 1989, pp. 177–178.

[Co–PeRi] J. Coates and B. Perrin-Riou, *On p-adic L-functions attached to motives over* $\mathbf{Q}$, Adv. Stud. Pure Math. **17** (1989), 23–54.

[Co–Sch] J. Coates and C.-G. Schmidt, *Iwasawa theory for the symmetric square of an elliptic curve*, J. Reine Angew. Math. **375/376** (1987), 104–156.

[De1] P. Deligne, *Formes modulaires et représentations l-adiques*, Sém. Bourbaki 1968/69, Exp. no. 335, Lecture Notes in Math., vol. 179, Springer-Verlag, Berlin and New York, 1971, pp. 139–172.

[De2] ——, *La conjecture de Weil.* I, Inst. Hautes Études Sci. Publ. Math. **43** (1974), 273–307.

[De3] ——, *Valeurs de fonctions L et periodes d'intégrales*, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 313–342.

[De-R] P. Deligne and K. A. Ribet, *Values of abelian L-functions at negative integers over totally real fields*, Invent. Math. **59** (1980), 227–286.

[Fe] P. Feit, *Poles and residues of Eisenstein series for symplectic and unitary groups*, Mem. Amer. Math. Soc. **61** (1986), no. 346.

[Fr1] E. Freitag, *Die Wirkung von Heckeoperatoren auf Thetareihen mit harmonischen Koeffizienten*, Math. Ann. **258** (1982), 419–440.

[Fr2] ——, *Siegelsche Modulfunktionen*, Springer-Verlag, Berlin and New York, 1983.

[Ga] P. B. Garrett, *Pullbacks of Eisenstein series: Applications*, Automorphic Forms of Several Variables (Taniguchi Sympos., 1983), Birkhäuser, Boston and Basel, 1984.

[Ge–PSh] S. Gelbart, I. I. Piatetski-Shapiro, and S. Rallis, *Explicit constructions of automorphic L-functions*, Lecture Notes in Math., vol. 1254, Springer-Verlag, Berlin and New York, 1987.

[HSh] Shai Haran, *p-Adic L-functions for modular forms*, Compositio Math. **62** (1986), 31–46.

[Ha1] M. Harris, *Special values of zeta functions attached to Siegel modular forms*, Ann. Sci. École Norm. Sup. (4) **14** (1981), 77–120.

[Ha2] ——, *The rationality of holomorphic Eisenstein series*, Invent. Math. **63** (1981), 305–310.

[Ha3] ——, *Arithmetical vector bundles and automorphic forms on Shimura varieties.* I, Invent. Math. **59** (1985), 151–189.

[Ha4] ——, *Cohomology of arithmetic groups and automorphic forms*, Proc. Conf. Luminy, France 1989, Lecture Notes in Math., vol. 1447, Springer-Verlag, Berlin and New York, 1990, pp. 155–202.

[Hi1] H. Hida, *p-Adic measure attached to the zeta functions associated with two elliptic cusp forms.* I, Invent. Math. **79** (1985), 159–195.

[Hi2] ——, *Galois representations into* $\mathrm{GL}_2[[\mathbf{Z}_p]]$ *attached to ordinary cusp forms*, Invent. Math. **85** (1986), 545–613.

[Hi3] ——, *On p-adic L-functions of* $\mathrm{GL}(2)\times\mathrm{GL}(2)$ *over totally real fields*, preprint, 1990.

[H-PSh] R. Howe and I. I. Piatetski-Shapiro, *A counterexample to the "Generalized Ramanujan conjecture" for (quasi-) split groups*, Automorphic Forms, Representations and L-functions, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 315–322.

[Iw] K. Iwasawa, *Lectures on p-adic L-functions*, Ann. of Math. Stud., no. 74, Princeton Univ. Press, Princeton, NJ, 1972.

[Ja] U. Jannsen, *Mixed motives and algebraic K-theory*, Lecture Notes in Math., vol. 1400, Berlin and New York, 1990.

[Ka1] N. M. Katz, *p-adic interpolation of real analytic Eisenstein series*, Ann. of Math. (2) **104** (1976), 459–571.

[Ka2] ——, *The Eisenstein measure and p-adic interpolation*, Amer. J. Math. **99** (1977), 238–311.

[Ka3] ——, *p-adic L-functions for CM-fields*, Invent. Math. **48** (1978), 199–297.

[Kl1] H. Klingen, *Über die Werte Dedekindscher Zetafunktionen*, Math. Ann. **145** (1962), 265–272.

[Kl2] ——, *Über den arithmetischen Charakter der Fourier-koeffizienten von Modulformen*, Math. Ann. **147** (1962), 176–188.

[Kl3] ——, *Zum Darstellungssatz für Siegelsche Modulformen*, Math. Z. **102** (1967), 30–43.

[Kl4] ——, *Über Poincarésche Reihen zur Siegelschen Modulgruppe*, Math. Ann. **168** (1967), 157–170.

[Kl5] ——, *Über Poincarésche Reihen vom Exponentialtyp*, Math. Ann. **234** (1978), 145–157.

[Ko1] N. Koblitz, *p-adic numbers , p-adic analysis and zeta functions*, 2nd ed., Springer-Verlag, Berlin and New York, 1984.

[Ko2] ——, *p-adic analysis: A short course on recent work*, London Math. Soc. Lecture Notes Ser., no. 46, Cambridge Univ. Press, London and New York, 1980.

[Koe] M. Koecher, *Zur Theorie der Modulformen n-ten Grades*. I. II, Math. Z. **59** (1954), 399–416.

[Koj] H. Kojima, *On construction of Siegel modular forms of degree two*, Math. Soc. Japan **34** (1982), 393–412.

[Ku-Le] T. Kubota and H.-W. Leopoldt, *Eine p-adische Theorie der Zetawerte*, J. Reine Angew. Math. **214/215** (1964), 328–339.

[Kur] N. Kurokawa, *Examples of eigenvalues of Hecke operators on Siegel cusp forms of degree two*, Invent. Math. **49** (1978), 149–165.

[Kur–Miz] N. Kurokawa and S. Mizumoto, *On Eisenstein series of degree two*, Proc. Japan Acad. A **57** (1981), 134–139.

[Man1] Yu. I. Manin, *Cyclotomic fields and modular curves*, Uspekhi Mat. Nauk **26** (1971), 7–78. (Russian)

[Man2] ——, *Cusp forms and zeta functions of modular curves*, Izv. Akad. Nauk. Ser. Mat. **36** (1972), 19–66. (Russian)

[Man3] ——, *Explicit formulas for the eigenvalues of Hecke operators*, Acta Arith. **24** (1973), 239–249.

[Man4] ——, *Periods of cusp forms and p-adic Hecke series*, Mat. Sb. **92** (1973), 378–401. (Russian)

[Man5] ——, *The values of p-adic Hecke series at integer points of the critical strip*, Mat. Sb. **93** (1974), 621–626. (Russian)

[Man6] ——, *Non-Archimedean integration and p-adic L-functions of Jacquet-Langlands*, Uspekhi Mat. Nauk **31** (1976), 5–54. (Russian)

[Man7] ——, *Modular forms and number theory*, Proc. Internat. Congress Math. (Helsinki, 1978), pp. 177–186.

[Man-Pa] Yu. I. Manin and A. A. Panchishkin, *Convolutions of Hecke series and their values at integral points*, Mat. Sb. **104** (1977), 617–651. (Russian)

[Maz1] B. Mazur, *On the special values of L-functions*, Invent. Math. **55** (1979), 207–240.

[Maz2] ——, *Modular curves and arithmetic*, Proc. Internat. Congress Math. (Warszawa, 1982), North-Holland, Amsterdam, 1984, pp. 185–211.

[Maz3] ——, *Deforming Galois representations*, Galois Groups over **Q** (Y. Ihara, K. Ribet, and J.-P. Serre, eds.), Springer-Verlag, Berlin and New York, 1989.

[Maz-SD] B. Mazur and H. P. F. Swinnerton–Dyer, *Arithmetic of Weil curves*, Invent. Math. **25** (1974), 1–61.

[Maz–W1] B. Mazur and A. Wiles, *Analogies between function fields and number fields*, Amer. J. Math. **105** (1983), 507–521.

[Maz–W2] ——, *Class fields of Abelian extensions of* **Q**, Invent. Math. **76** (1984), 179–330.

[Maz–W3] ——, *On p-adic analytic families of Galois representations*, Compositio Math. **59** (1986), 231–264.

[Miy] T. Miyake, *On automorphic forms on* GL_2 *and Hecke operators*, Ann. of Math. **94** (1971), 174–189.

[Oda] T. Oda, *Periods of Hilbert modular surfaces*, Progr. Math., vol. 19, Birkhäuser, Boston, 1982.

[Oh] M. Ohta, *On the zeta-functions of an Abelian scheme over the Shimura curve*, Japan J. Math. **9** (1983), 1–26.

[Pa1] A. A. Panchishkin, *Symmetric squares of Hecke series and their values at integral points*, Mat. Sb. **108** (1979), 393–417. (Russian)

[Pa2] ——, *On p-adic Hecke series*, Algebra (A. I. Kostrikin, ed.), Moscow Univ. Press, 1980, pp. 68–71. (Russian)

[Pa3] ——, *Complex valued measures attached to Euler products*, Trudy Sem. Petrovsk. **7** (1981), 239–244. (Russian)

[Pa4] ——, *Modular forms*, Algebra, Topology, Geometry, vol. 19, VINITI, Moscow, 1981, pp. 135–180. (Russian)

[Pa5] ——, *Local measures attached to Euler products in number fields*, Algebra (A. I. Kostrikin, ed.), Moscow Univ. Press, 1982, pp. 119–138. (Russian)

[Pa6] ——, *Automorphic forms and the functoriality principle*, Automorphic Forms, Representations and L-functions, "Mir", Moscow, 1984, pp. 249–286. (Russian)

[Pa7] ——, *Le prolongement p-adique analytique de fonctions L de Rankin*, C. R. Acad. Sci. Paris Sér. **294** (1982), 51–53, 227–230.

[Pa8] ——, *A functional equation of the non-Archimedean Rankin convolution*, Duke Math. J. **54** (1987), 77–89.

[Pa9] ——, *Non-Archimedean convolutions of Hilbert modular forms*, Abstracts 19th USSR Algebraic Conf. (Lvov), vol. 1, 1987, p. 211.

[Pa10] ——, *Non-Archimedean Rankin L-functions and their functional equations*, Izv. Akad. Nauk. Ser. Mat. **52** (1988), 336–354. (Russian)

[Pa11] ——, *Convolutions of Hilbert modular forms and their non-Achimedean analogues*, Mat. Sb. **136** (1988), 574–587. (Russian)

[Pa12] ——, *Non-Archimedean automorphic zeta-functions*, Moscow Univ. Press, 1988.

[Pa13] ——, *Convolutions of Hilbert modular forms, motives and p-adic zeta functions*, preprint MPI, no. 43, Bonn, 1990.

[Pa14] ——, *Non-Archimedean L-functions of Siegel and Hilbert modular forms*, Lecture Notes in Math., vol. 1471, Springer-Verlag, Berlin and New York, 1991.

[PSh] I. I. Piatetski-Shapiro, *Automorphic functions and the geometry of classical domains*, Fizmatgiz, Moscow, 1961; English transl., Gordon and Breach, New York, 1969.

[PSh-R] I. I. Piatetski-Shapiro and S. Rallis, *L-functions of automorphic forms on simple classical groups*, Modular Forms Sympos. (Durham, 1983), Wiley, Chichester, 1984.

[Rag1] S. Raghavan, *Modular forms of degree n and representations by quadratic forms*, Ann. of Math. (2) **70** (1959), 446–477.

[Rag2] ——, *Estimates of coefficients of modular forms and generalized modular relations*, Automorphic Forms, Representation Theory and Arithmetic (Papers Colloq., Bombay, 1979), Springer-Verlag, Berlin and New York, 1981, pp. 247–254.

[Ran1] R. A. Rankin, *Contribution to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions*. I, II, Proc. Cambridge Philos. Soc. **35** (1939), 351–372.

[Ran2] ——, *The scalar product of modular forms*, Proc. London Math. Soc. (3) **35** (1939), 351–372.

[Ro-Tu] J. D. Rogawski and J. B. Tunnel, *On Artin L-functions associated to Hilbert modular forms*, Invent. Math. **74** (1983), 1–42.

[Schm] C.-G. Schmidt, *The p-adic L-functions attached to Rankin convolutions of modular forms*, J. Reine Angew. Math. **368** (1986), 201–220.

[Scho] A. J. Scholl, *Motives for modular forms*, preprint.

[Shi1] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton Univ. Press, Princeton, NJ, 1971.

[Shi2] ——, *On the holomorphy of certain Dirichlet series*, Proc. London Math. Soc. (3) **31** (1975), 79–98.

[Shi3] ——, *The special values of the zeta functions associated with cusp forms*, Comm. Pure Appl. Math. **29** (1976), 783–804.

[Shi4] ——, *On the periods of modular forms*, Math. Ann. **229** (1977), 211–221.

[Shi5] ——, *On certain reciprocity laws for theta functions and modular forms*, Acta Math. **141** (1978), 35–71.

[Shi6] ——, *The special values of zeta functions associated with Hilbert modular forms*, Duke Math. J. **45** (1978), 637–679.

[Shi7] ——, *Algebraic relations between critical values of zeta functions and inner products*, Amer. J. Math. **105** (1983), 253–285.

[Shi8] ——, *On Eisenstein series*, Duke Math. J. **50** (1983), 417–476.

[St1] J. Sturm, *Special values of zeta functions and Eisenstein series of half integral weight*, Amer. J. Math. **102** (1980), 219–240.

[St2] ——, *The critical values of zeta-functions associated to the symplectic group*, Duke Math. J. **48** (1981), 327–350.

[V1] M. M. Višik, *Non-Archimedean measures associated with Dirichlet series*, Mat. Sb. **99** (1976), 248–260.

[V2] ——, *Non-Archimedean spectral theory*, Modern Problems of Mathematics, vol. 25, VINITI, Moscow, 1984, pp. 51–114.

[Wa] L. Washington, *Introduction to cyclotomic fields*, Springer-Verlag, Berlin and New York, 1982.

INSTITUT FOURIER, RUE MATHÉMATIQUES 100, BP74, SAINT-MARTIN D'HÈRES, 38402 FRANCE

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

A p-adic Property of Hodge Classes on Abelian Varieties

DON BLASIUS

Introduction

(0.1) The Hodge conjecture asserts that every Hodge class is the image under the cycle map of an algebraic cycle. If true, then the Hodge classes possess arithmetic properties which will sometimes admit definitions independent of the conjecture. Thus, the problem of proving such properties unconditionally arises.

In a basic paper **[D1]**, Deligne took an important step along these lines by defining an absolute Hodge class and proving that on an abelian variety every Hodge class is an absolute Hodge class. This result has powerful consequences which stem mostly from the strengthening of the Shimura-Taniyama reciprocity law, in a motivic setting, that it makes possible **[DM, D3]**.

(0.2) In the present paper, we prove another such arithmetic property of Hodge classes on abelian varieties. Let p be a rational prime, and let $\sigma_p : \overline{\mathbf{Q}} \to \overline{\mathbf{Q}_p}$ be an embedding. Let X be a proper smooth variety defined over $\overline{\mathbf{Q}}$, and let $\gamma_B \in H_B^{2j}(X)(j)$ be an absolute Hodge class (see (1.3)). Let $\gamma_p = I_p(\gamma_B) \in H_p^{2j}(X)(j)$ and $\gamma_{\mathrm{DR}} = I_\infty(\gamma_B) \in H_{\mathrm{DR}}^{2j}(X)(j)$ be the images of γ_B in p-adic étale and algebraic De Rham cohomology, respectively, under the comparison maps. Recall that Faltings **[F]** has proven Fontaine's conjecture **[Fo]**: there exists an isomorphism

$$I_{\mathrm{DR}} : H_p^j(\sigma_p X)(k) \otimes B_{\mathrm{DR}} \to H_{\mathrm{DR}}^j(\sigma_p X)(k) \otimes_{\sigma_p(\overline{\mathbf{Q}})} B_{\mathrm{DR}}$$

where B_{DR} is the ring introduced by Fontaine. Consequently, it is natural to make the following definition: an absolute Hodge class γ_B is De Rham if,

---

1991 *Mathematics Subject Classification.* Primary 11F85, 14F20, 14F30, 14F40, 14K15.

This paper is in final form and no version of it will be submitted for publication elsewhere.

Partially supported by NSF grant DMS 90-01878.

for all primes p and all embeddings $\sigma_p : \overline{\mathbf{Q}} \to \overline{\mathbf{Q}_p}$, we have

$$I_{\mathrm{DR}}(\sigma_p \gamma_p) = \sigma_p \gamma_{\mathrm{DR}}.$$

Of course, every algebraic class is a De Rham class.

(0.3) THEOREM. *Let X be an abelian variety defined over $\overline{\mathbf{Q}}$. Then every Hodge class on X is De Rham.*

The purpose of this paper is to prove this result while at the same time sketching a variant proof of Deligne's theorem. The proof of the new result here employs the methods of **[D1]**, replacing the use of the usual De Rham isomorphism I_∞ in **[D1]** by Faltings's I_{DR}. However, at a crucial point in **[D1]**, the argument (proving Priniciple B) employs the Gauss-Manin connection. Since it appears that the behavior of I_{DR} relative to Gauss-Manin is not well understood, we find it necessary to change this part of the argument in the p-adic case. Indeed, we replace the use of Gauss-Manin by an appeal to general facts about the structure of the cohomology of a family of varieties. Apart from this simplification, the proof of the main result here follows very closely the argument of **[D1]**. However, for clarity, we have taken throughout a motivic, i.e., Tannakian viewpoint, which provides a slight change of perspective on Principle A of **[D1]**. For progress concerning Gauss-Manin itself in the p-adic setting, see **[W1]**.

(0.4) As an application of (0.3), we deduce formally in the last section that Hodge classes are *crystalline* in the case of primes of good reduction. This fact has several applications in the case of abelian varieties of CM type **[Co, O1, W2]**. We have neglected here to include applications of (0.3) itself. Nevertheless, one should note that it is an easy exercise to define B_{DR}-valued p-adic periods of CM motives and to develop a formalism of p-adic period relations parametrized by identities of tensor products (monomial relations) which occur between such motives, extending the archimedean formalism given in **[D2]**, **[Sc]**, and **[Sh]**. In fact, one may easily proceed further to obtain a theory of p-adic periods with values in $\mathbf{C}_p$, at least once one has chosen an identification of $\mathbf{C}$ with $\mathbf{C}_p$, and obtained thereby a trivialization of the p-adic Tate module. This follows because the functoriality of I_{DR} enables one to define periods in the graded ring B_{HT} associated to the filtration on B_{DR} in such a way that an eigenperiod for the complex multiplication is supported on a single graded component, necessarily isomorphic to $\mathbf{C}_p(n)$. After invoking the trivialization, this component is identified with $\mathbf{C}_p$.

That such a formalism should exist was proven using Dwork theory in the ordinary case by Gillard **[G]** by a method that imitates Shimura's proof of his monomial relations theorem **[Sh]**.

More systematically, one may define, as in **[DM]**, a category of motives starting from abelian varieties of CM type where the morphisms are given by absolute Hodge cycles that are also De Rham. Then (0.3) says that this category is the same as the category defined using just absolute Hodge cycles as morphisms. In particular, the Taniyama group is also the motivic Galois group of this category [D3, L].

(0.5) Acknowledgment. I thank A. Ogus for conversations that led to substantial simplification and restructuring of the proof and G. Faltings for a conversation concerning the compatibility of the De Rham and crystalline comparison maps. I also thank the Max Planck Institut in Bonn for its hospitality during the preparation of the paper.

1. Cohomologies and comparison maps

(1.1) Let K be a subfield of $\mathbf{C}$, and let X be a smooth projective variety defined over K. On X we have several cohomology functors.

First, we have

$$H_B^*(X) = \bigoplus_{j=0}^{2\dim X} H_B^j(X),$$

the topological cohomology of $X(\mathbf{C})$ with rational coefficients. Each $H_B^j(X)\otimes \mathbf{C}$ is equipped with a Hodge decomposition

$$H_B^j(X) \otimes \mathbf{C} = \bigoplus_{p+q=j} H^{p,q}(X).$$

Next we have

$$H_{\mathrm{DR}}^*(X) = \bigoplus_{j=0}^{2\dim X} H_{\mathrm{DR}}^j(X),$$

the algebraic De Rham cohomology of X. Each $H_{\mathrm{DR}}^j(X)$ is a K vector space equipped with the decreasing (Hodge) filtration $F^* H_{\mathrm{DR}}^j(X)$.

Last, for each rational prime p, we have

$$H_p^*(X) = \bigoplus_{j=0}^{2\dim X} H_p^j(X),$$

the p-étale cohomology of $X \times_K \overline{K}$. Each $H_p^j(X)$ is a $\mathbf{Q}_p$ vector space equipped with a continuous action of $\mathrm{Gal}(\overline{K}/K)$.

Between these cohomology theories we have the graded comparison isomorphisms:

$$I_\infty : H_B^*(X) \otimes \mathbf{C} \to H_{\mathrm{DR}}^* \otimes_K \mathbf{C}$$

and, for p a prime,

$$I_p : H_B^* \otimes \mathbf{Q}_p \to H_p^*(X).$$

Of course, the map I_∞ satisfies

$$I_\infty \left(\bigoplus_{\substack{p+q=j \\ p \ge p_0}} H_B^{p,q} \right) = F_j^{p_0} H_{\mathrm{DR}}^j \otimes_K \mathbf{C}.$$

If $K \subseteq \overline{\mathbf{Q}} \subseteq \mathbf{C}$ is a number field, we obtain further structures. Let $\mathbf{C}_p$ be a completion of an algebraic closure of $\mathbf{Q}_p$, and let σ_p be an embedding of $\overline{\mathbf{Q}}$ in $\mathbf{C}_p$. Let $\widehat{\sigma_p(K)}$ be the topological closure of $\sigma_p(K)$. Let B_{DR} be Fontaine's **[Fo]** **Z**-filtered $\overline{\mathbf{Q}_p}$ algebra. If D_p denotes the group of continuous automorphisms of $\overline{\mathbf{Q}_p}$, then B_{DR} is a D_p module for which the action is semilinear: $\tau(\alpha b) = \tau(\alpha)\tau(b)$, if $\alpha \in \overline{\mathbf{Q}_p}$, $b \in B_{\mathrm{DR}}$, and $\tau \in D_p$. Each element of D_{σ_p} extends uniquely to a continuous automorphism of $\mathbf{C}_p$. Let D_{σ_p} be the subgroup of D_p consisting of the elements that fix $\widehat{\sigma_p(K)}$ pointwise, and let V_p be a finite-dimensional $\mathbf{Q}_p$ vector space on which D_{σ_p} acts continuously. Then $V_p \otimes_{\mathbf{Q}_p} B_{DR}$ acquires a filtration from the filtration on B_{DR} and is naturally a D_{σ_p}-module: one puts $\tau(v \otimes b) = \tau(v) \otimes \tau(b)$.

For a variety X defined over a field $K \subseteq \overline{\mathbf{Q}}$, let $\sigma_p X$ be the conjugate of X by σ_p. Basic to our work is the following result of Faltings, already cited in the Introduction:

THEOREM. *For each prime p, there is a functorial D_{σ_p}-equivariant filtered isomorphism*

$$I_{\mathrm{DR}} : H_p^*(\sigma_p X) \otimes_{\mathbf{Q}_p} B_{\mathrm{DR}} \to H_{\mathrm{DR}}^j(\sigma_p X) \otimes_{\sigma_p(K)} B_{\mathrm{DR}}$$

where on the right-hand side D_{σ_p} acts via the right factor and the filtration is that defined by the tensor product of the filtrations on B_{DR} and $H_{\mathrm{DR}}^j(\sigma_p X)$. The isomorphism I_{DR} is compatible with cycle maps and with extension of the ground field.

(1.2) Tate twists. Let

$$\begin{aligned} \mathbf{Q}_p(1) &= \varprojlim \mu_{p^n} \quad \text{(for each prime } p), \\ \mathbf{Q}_B(1) &= 2\pi i \mathbf{Q} \subset \mathbf{C}, \\ \mathbf{Q}_{\mathrm{DR}}(1) &= \mathbf{Q}. \end{aligned}$$

Let

$$\chi_p : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{Z}_p^* = \mathrm{Aut}(\mathbf{Q}_p(1))$$

be the p-adic cyclotomic character; it gives the natural action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $\mathbf{Q}_p(1)$ which we always regard as a Galois module. Let $\mathbf{Q}_B(1)$ have the unique Hodge structure which is purely of type $(-1, -1)$, and let $\mathbf{Q}_{\mathrm{DR}}(1)$ have the Hodge filtration $F_{\mathrm{DR}}^j \mathbf{Q}_{\mathrm{DR}}(1) = \mathbf{Q}_{\mathrm{DR}}(1)$ for $j \leq -1$ and $F_{\mathrm{DR}}^j \mathbf{Q}_{\mathrm{DR}}(1) = 0$ if $j \geq 0$. Let

$$I_\infty : \mathbf{Q}_B(1) \otimes \mathbf{C} \to \mathbf{Q}_{DR}(1) \otimes \mathbf{C} = \mathbf{C}$$

be defined by

$$I_\infty(\alpha \otimes z) = \alpha z$$

for $\alpha \in (2\pi i)\mathbf{Q}$ and $z \in \mathbf{C}$. Let

$$I_p : \mathbf{Q}_B(1) \otimes \mathbf{Q}_p \to \mathbf{Q}_p(1)$$

be the map defined via the inverse limit of the isomorphisms

$$\exp : (p^{-n} 2\pi i\mathbf{Z})/(2\pi i\mathbf{Z}) \to \mu_{p^n}.$$

We have, for any n, and any finite extension L of $\mathbf{Q}_p$,

$$(B_{\mathrm{DR}} \otimes (\mathbf{Q}_p(1)^{\otimes n}))^{D_{\sigma_p}} = \widehat{\sigma_p(K)}$$

for all K and σ_p, where for $n \le 0$, $\mathbf{Q}_p(1)^{\otimes n} = (\mathbf{Q}_p(1)^{\vee})^{\otimes |n|}$. Let

$$I_{\mathrm{DR}} : \mathbf{Q}_p(1) \otimes B_{\mathrm{DR}} \to B_{\mathrm{DR}}$$

be the B_{DR}-linear extension of this identity. For each subscript $! = B$, DR, p, let $\mathbf{Q}_!(n) = \mathbf{Q}_!(1)^{\otimes n}$, with the same convention as above if $n \le 0$. In this case, let I_∞, I_p, and I_{DR} denote the maps between these objects naturally defined via those just introduced and having the same symbol.

(1.3) Hodge classes. Let

$$H_B^{2n}(X)(n) = H_B^{2n}(X) \otimes \mathbf{Q}_B(n).$$

Then $H_B^{2n}(X)(n)$ carries a Hodge structure of weight 0 defined by

$$H_B(X)^{p,q} = H_B(X)(n)^{p-n,q-n}.$$

A **Hodge class** $\gamma_B \in H_B^{2n}(X)(n)$ is an element of type $(0,0)$. Let

$$\gamma_{\mathrm{DR}} = I_\infty(\gamma_B)$$

and

$$\gamma_p = I_p(\gamma_B).$$

Then γ_B is **absolutely Hodge** if, for any automorphism τ of $\mathbf{C}$, there exists a Hodge class

$$\gamma_B(\tau) \in H_B^{2n}(\tau X)(n)$$

such that

$$I_\infty(\gamma_B(\tau)) = \tau\gamma_{\mathrm{DR}}$$

and

$$I_p(\gamma_B(\tau)) = \tau\gamma_p$$

for each p. In this case, we define $\tau\gamma_B = \gamma_B(\tau)$. Note that if γ_B is absolutely Hodge then $\gamma_{\mathrm{DR}} \in H_{\mathrm{DR}}^{2n}(X)(n) \otimes_K L$ for a finite extension L of K. Indeed, $\mathrm{Aut}(\mathbf{C}/K)$ acts on the finite-dimensional rational vector subspace of $H_p^{2n}(X)(n)$ generated by the absolute Hodge cycles. If the image were infinite, it would necessarily be uncountable, which is impossible. Since $\tau\gamma_{\mathrm{DR}} = \gamma_{\mathrm{DR}}$ if and only if $\tau\gamma_p = \gamma_p$, γ_{DR} is defined over a finite extension. The smallest such extension is called the field of definition of γ_B; it is the field defined, via Galois theory, by the stabilizer of γ_B in $\mathrm{Gal}(\overline{K}/K)$.

Suppose now that K is a number field and γ_B is an absolute Hodge cycle. As in the Introduction, we say that γ_B is **De Rham** if, for all p and for all embeddings $\sigma_p : \overline{\mathbf{Q}} \to \mathbf{C}_p$,

$$I_{\mathrm{DR}}(\sigma_p \gamma_p) = \sigma_p \gamma_{\mathrm{DR}}.$$

Let X and Y be smooth connected projective varieties defined over $K \subseteq \mathbf{C}$. Suppose that X has dimension n. Define $\mathrm{Mor}_{\mathrm{H}}(X, Y)$ to be the space of all Hodge classes in $H_B^{2n}(X \times Y)(n)$, and let $\mathrm{Mor}_{\mathrm{AH}}(X, Y)$ be the space of all absolute Hodge classes in $H_B^{2n}(X \times Y)(n)$. If K is again a number field, let $\mathrm{Mor}_{\mathrm{DR}}(X, Y)$ be the space of all De Rham classes in $H_B^{2n}(X \times Y)(n)$. Finally, for X and Y not necessarily connected, define $\mathrm{Mor}_?(X, Y) = \bigoplus_{i,j} \mathrm{Mor}_?(X_i, Y_j)$ where the collections $\{X_i\}$ and $\{Y_j\}$ are the connected components of X and Y and $? = \mathrm{H}, \mathrm{AH}$, or DR.

2. Motives

(2.1) Effective motives. We now briefly sketch the construction of some categories of motives. See [DM] for more details on the formalism. Let $\mathscr{C}$ be a subcategory of the category of smooth projective varieties defined over a given field $L \subseteq \mathbf{C}$. Assume that $\mathscr{C}$ is closed under disjoint unions and products. Define $\otimes_?^* \mathscr{C}$ ($? = \mathrm{H}, \mathrm{AH}, \mathrm{DR}$) to be the category with objects the symbols $h(X)$ for X in $\mathscr{C}$, and with morphisms given by $\mathrm{Hom}(h(X), h(Y)) = \mathrm{Mor}_?(X, Y)$. Then $\otimes_?^* \mathscr{C}$ is a $\mathbf{Q}$-linear category for which we put $h(X) \oplus h(Y) = h(X \coprod Y)$ and $h(X) \otimes h(Y) = h(X \times Y)$. Let $\otimes_?^+ \mathscr{C}$ be the category whose objects are pairs (M, p) with $M \in \otimes_?^* \mathscr{C}$ and p an idempotent element in $\mathrm{End}(M)$. The morphisms are

$$\mathrm{Hom}_?((M_1, p_1), (M_2, p_2)) = \{f : M_1 \to M_2 \mid f \circ p_1 = p_2 \circ f\}/\sim$$

where

$$\sim = \{f : M_1 \to M_2 \mid f \circ p_1 = p_2 \circ f = 0\}.$$

We put

$$(M_1, p_1) \otimes (M_2, p_2) = (M_1 \otimes M_2, p_1 \otimes p_2).$$

The definition is so arranged that any $\mathbf{Q}$-linear functor

$$\omega : \otimes_?^* \mathscr{C} \to \{\mathbf{Q}\text{-vector spaces}\}$$

extends to $\otimes_?^+ \mathscr{C}$ and we have

$$\omega((M, p)) = \mathrm{Im}(\omega(p)).$$

Especially, the rational Hodge structure-valued functor H_B^* on $\mathscr{C}$ extends to a functor ω_B on $\otimes_?^+ \mathscr{C}$, and we put, for $M = (h(X), p)$,

$$M_B = p(H_B^*(X)).$$

(2.2) Remark on tensor structure. Note that the canonical decomposition

$$H_B^*(X) = \bigoplus_{j=0}^{2\dim X} H_B^j(X)$$

provides a family of idempotents $\{p_j\}$ so that $h(X) = \bigoplus_{j=0}^{2\dim X} h^j$ with $h^j = (h(X), p_j)$ and $h_B^j = H_B^j(X)$. The functor sending M to M_B is faithful and **Q**-linear. However, it is not a tensor functor (cf. **[DM]**) since the commutativity isomorphism $c^* : h(X) \otimes h(Y) \to h(Y) \otimes h(X)$ given by the natural permutation isomorphism $X \times Y \to Y \times X$ sends $\gamma = \gamma_X \otimes \gamma_Y \in H_B^j(X) \otimes H_B^k(Y)$ to $c_B^*(\gamma) = (-1)^{jk}\gamma_Y \otimes \gamma_X \in H_B^j(Y) \otimes H_B^k(X)$. Hence, one replaces $c^* = \sum_{j,k} c_{j,k}^*$ by $c = \sum_{j,k} c_{j,k}$ where $c_{j,k} = (-1)^{jk} c_{j,k}^*$. With this change of the commutativity isomorphisms on $\otimes_?^+\mathscr{C}$, the category becomes a tensor category and H_B^* becomes a tensor functor.

(2.3) Motives. Suppose that $\mathscr{C}$ contains a curve Γ so that $\otimes_?^+\mathscr{C}$ contains $L = h^2(\Gamma)$. Then $L_B \cong \mathbf{Q}_B(-1)$, and, for any $n \geq 0$, $\mathrm{Hom}(M, N)$ is canonically isomorphic to $\mathrm{Hom}(M \otimes L^n, N \otimes L^n)$ via $\phi \to \phi \otimes 1$ for $\phi \in \mathrm{Hom}(M, N)$. Let $\otimes_?\mathscr{C}$ denote the category obtained by inverting L: an object of $\otimes_?\mathscr{C}$ is a pair (M, n) with $M \in \otimes_?^+\mathscr{C}$ and $n \geq 0$, and the morphisms are given by

$$\mathrm{Hom}((M, m), (N, n)) = \mathrm{Hom}(M \otimes L^{N-m}, N \otimes L^{N-n})$$

for any $N \geq m, n$. For $n \leq 0$, put $(M, n) = M \otimes L^{|n|}$. It is conventional to define $M(n) = (M, n)$. Then the rule $M(n)_B = M_B(n)$ extends the functor $M \to M_B$ to $\otimes_?\mathscr{C}$. The categories $\otimes_{\mathrm{H}}\mathscr{C}$, $\otimes_{\mathrm{AH}}\mathscr{C}$, and $\otimes_{\mathrm{DR}}\mathscr{C}$ are called the categories of **motives** for **Hodge, Absolute Hodge,** and **De Rham classes**, generated by $\mathscr{C}$, respectively. Note that we have obvious inclusions:

$$\otimes_{\mathrm{DR}}\mathscr{C} \subseteq \otimes_{\mathrm{AH}}\mathscr{C} \subseteq \otimes_{\mathrm{H}}\mathscr{C}.$$

If $\mathscr{C}$ is the category generated by a single variety X, we write $\otimes_{\mathrm{H}}X$, $\otimes_{\mathrm{AH}}X$, and $\otimes_{\mathrm{DR}}X$. Finally, we sometimes write ω_B for the functor which to a motive M attaches its topological cohomology M_B, viewed just as a rational vector space.

(2.4) PROPOSITION. *The categories $\otimes_{\mathrm{H}}\mathscr{C}$, $\otimes_{\mathrm{AH}}\mathscr{C}$, and $\otimes_{\mathrm{DR}}\mathscr{C}$ are semi-simple, Tannakian categories for which ω_B is a fiber functor.*

PROOF. This is proved in **[DM,** §6] for the AH case. The other cases are identical.

(2.5) Other realizations. Let $M = (h(X), e)$ be a motive in one of the categories of motives just constructed. Via the comparison isomorphisms I_∞ and I_p, the idempotent class $e = e_B$ defines $e_{\mathrm{DR}} = I_\infty \circ e_B \circ I_\infty^{-1}$ in $\mathrm{End}_{\mathbf{C}}(H_{\mathrm{DR}}^*(X) \otimes_K \mathbf{C})$ and $e_p = I_p \circ e_B \circ I_p^{-1}$ in $\mathrm{End}_{\mathbf{Q}_p}(H_p^*(X))$. Put $M_{\mathrm{DR},\mathbf{C}} =$

$\mathrm{Im}(e_{\mathrm{DR}})$, $M_p = \mathrm{Im}(e_p)$, and extend these functors to all of $\otimes_? \mathscr{C}$ in the evident way, i.e., via the rule $M(n)_B = M_B(n)$, etc.

If $M \in \otimes_{\mathrm{AH}} \mathscr{C}$ and e_B is defined over K, then M_p is a $\mathrm{Gal}(\overline{K}/K)$-module, and $e_{\mathrm{DR}} \in \mathrm{End}_K(H^*_{\mathrm{DR}}(X))$ so that $\mathrm{Im}(e_{\mathrm{DR}}) = M_{\mathrm{DR}}$ is a K-vector space such that $M_{\mathrm{DR}} \otimes_K \mathbf{C} = M_{\mathrm{DR},\mathbf{C}}$. Then M_{DR} carries a K-rational filtration $F^{\cdot} M_{\mathrm{DR}}$ such that

$$I_\infty \left(\bigoplus_{p \geq p_0} M_B^{p,q} \right) = F^{p_0} M_{\mathrm{DR}} \otimes_K \mathbf{C}.$$

If $M \in \otimes_{\mathrm{DR}} \mathscr{C}$, then we also have, for each prime p and each $\sigma_p : \overline{\mathbf{Q}} \to \mathbf{C}_p$, the comparison isomorphism

$$I_{\mathrm{DR}} : (\sigma_p M_p) \otimes_{\mathbf{Q}_p} B_{\mathrm{DR}} \to (\sigma_p M_{\mathrm{DR}}) \otimes_{\widehat{\sigma_p(K)}} B_{\mathrm{DR}}.$$

(2.6) Dual groups. Let

$$\mathscr{G}_? = \mathrm{Aut}^{\otimes}(\omega_B, \otimes_? \mathscr{C})$$

be the group of automorphisms of ω_B that respect the tensor structures (see **[DM]**). Then $\mathscr{G}_?$ is a connected reductive pro-algebraic group defined over $\mathbf{Q}$; it is algebraic if and only if the ring of isomorphism classes of objects of $\otimes_? \mathscr{C}$ is finitely generated. Note that

$$\mathscr{G}_{\mathrm{H}} \subseteq \mathscr{G}_{\mathrm{AH}} \subseteq \mathscr{G}_{\mathrm{DR}}.$$

Each $\mathscr{G}_?$ acts on each M_B ($M \in \otimes_? \mathscr{C}$) via a representation ρ_M and the correspondence $M \to \rho_M$, extended to morphisms, defines an equivalence of categories for which

$$\omega_B(\mathrm{Hom}_?(M, N)) = \mathrm{Hom}_{\mathscr{G}_?}(M_B, N_B).$$

We have evident notions of Hodge (resp. absolute Hodge, resp. De Rham) classes on a motive M in the category $\otimes_{\mathrm{H}} \mathscr{C}$ (resp. $\otimes_{\mathrm{AH}} \mathscr{C}$, resp. $\otimes_{\mathrm{DR}} \mathscr{C}$). Especially, if $\tau \in \mathrm{Aut}(\mathbf{C})$, and $M \in \otimes_? \mathscr{C}$, ($? = \mathrm{AH}$ or DR), and $M = (X, e)(n)$, then $\tau M = (\tau X, \tau e)(n)$ is defined.

(2.7) PROPOSITION. *Let $M \in \otimes_? \mathscr{C}$. Then the subspace $M_B^{\mathscr{G}_?}$ of $\mathscr{G}_?$- invariants in M_B is the subspace of all ?-classes, for $? = \mathrm{H}, \mathrm{AH}$, or* DR.

(2.8) PROOF. We give the proof for $\mathscr{G}_{\mathrm{DR}}$. The other cases are the same. Note first that the space of De Rham classes on M is

$$\{\phi_B(1) \mid \phi \in \mathrm{Hom}(\mathbf{Q}(0), M)\}$$

where $1 \in \mathbf{Q} = \mathbf{Q}(0)_B$. Thus every De Rham class is $\mathscr{G}_{\mathrm{DR}}$-invariant. If $\gamma_B \in M_B$ is fixed by $\mathscr{G}_{\mathrm{DR}}$, then the map $\phi_{\gamma,B} : \mathbf{Q}_B(0) \to M_B$ such that $\phi_{\gamma,B}(1) = \gamma_B$ is $\mathscr{G}_{\mathrm{DR}}$-invariant. Hence it belongs to $\mathrm{Hom}(\mathbf{Q}(0), M)$ and so γ_B is De Rham.

(2.9) Proposition. *Let $\mathcal{A}$ and $\mathcal{B}$ be* **Q**-*linear Tannakian categories with $\mathcal{A} \subseteq \mathcal{B}$. Let $\omega : \mathcal{B} \to \{\mathbf{Q}\text{-vector spaces}\}$ be a fiber functor, and denote its restriction to $\mathcal{A}$ by $\omega_{\mathcal{A}}$. Let $\mathcal{G}_{\mathcal{B}} = \mathrm{Aut}^{\otimes}(\omega, \mathcal{B})$ and $\mathcal{G}_{\mathcal{A}} = \mathrm{Aut}^{\otimes}(\omega_A, \mathcal{A})$ so that $\mathcal{G}_{\mathcal{B}} \subseteq \mathcal{G}_{\mathcal{A}}$. Suppose that*

$$\mathcal{G}_{\mathcal{B}} = \mathcal{G}_{\mathcal{A}}.$$

Then the inclusion of $\mathcal{A}$ into $\mathcal{B}$ is an equivalence of categories.

(2.10) Proof. This is evident: $\mathcal{A}$ and $\mathcal{B}$ are both equivalent to the category of representations of the same group; hence, they are equivalent. That the inclusion of $\mathcal{A}$ into $\mathcal{B}$ defines an equivalence is also clear since under $\omega_{\mathcal{A}}$ the inclusion of the representations of $\mathcal{G}_{\mathcal{A}}$ obtained from $\mathcal{A}$ into those obtained from $\mathcal{B}$ is an equivalence of categories.

(2.11) Proposition (Principle A). *Let X be a smooth projective variety defined over the complex numbers. The following are equivalent:*

(1) $\otimes_{\mathrm{AH}} X = \otimes_{\mathrm{H}} X$.
(2) $\mathcal{G}_{\mathrm{H}} = \mathcal{G}_{\mathrm{AH}}$.
(3) *Every Hodge class in $\otimes_{\mathrm{H}} X$ is absolutely Hodge.*

Suppose that X is defined over a number field. Then the following are equivalent:

(4) $\otimes_{\mathrm{AH}} X = \otimes_{\mathrm{DR}} X$.
(5) $\mathcal{G}_{\mathrm{DR}} = \mathcal{G}_{\mathrm{AH}}$.
(6) *Every absolute Hodge class in $\otimes_{\mathrm{AH}} X$ is De Rham.*

(2.12) Proof. By Proposition (2.7), it is clear that (1) is equivalent to (3), and (4) is equivalent to (6). Furthermore, (3) implies (2), and (6) implies (5). Hence we need only show that (2) implies (1), and (5) implies (4). Let $M \in \otimes_{\mathrm{AH}} X$ with associated representation ρ_M of $\mathcal{G}_{\mathrm{AH}}$. Then M is indecomposable in $\otimes_{\mathrm{AH}} X$ if and only if ρ_M is irreducible, since the category is semisimple. Suppose that, as an element of $\otimes_{\mathrm{H}} X$, $M = M_1 \oplus M_2$. Then the restriction of ρ_M to $\mathcal{G}_{\mathrm{H}}$ is a nontrivial direct sum $\rho_1 \oplus \rho_2$. But since $\mathcal{G}_{\mathrm{H}} = \mathcal{G}_{\mathrm{AH}}$, this cannot happen. Thus each M that is irreducible in $\otimes_{\mathrm{AH}} X$ remains irreducible in $\otimes_{\mathrm{H}} X$. On the other hand, every irreducible object of $\otimes_{\mathrm{H}} X$ is a constituent of a $\otimes_{\mathrm{AH}} X$ irreducible object, by definition of the categories. Hence, $\otimes_{\mathrm{AH}} X$ and $\otimes_{\mathrm{H}} X$ have the same objects. Since

$$\begin{aligned} \omega_B(\mathrm{Hom}_{\mathrm{H}}(M, N)) &= \mathrm{Hom}_{\mathcal{G}_{\mathcal{X}}}(M_B, N_B) = \mathrm{Hom}_{\mathcal{G}_{\mathcal{AX}}}(M_B, N_B) \\ &= \omega_B(\mathrm{Hom}_{\mathrm{AH}}(M, N)), \end{aligned}$$

we conclude that

$$\mathrm{Hom}_{\mathrm{H}}(M, N) = \mathrm{Hom}_{\mathrm{AH}}(M, N)$$

as well. Thus, $\otimes_{\mathrm{AH}} X = \otimes_{\mathrm{H}} X$, as was to be shown.

The proof for the second case is exactly parallel, with "AH" replacing "H" and "DR" replacing "AH". (See **[D1]** for another approach which does not employ Tannakian duality.)

3. Principle B

(3.1) THEOREM. *Let S be a smooth, geometrically connected variety defined over the subfield K of $\mathbf{C}$. Let $\pi : X \to S$ be a smooth proper morphism defined over K. Let $\gamma_B \in H^0(S, \mathbf{R}^{2n}\pi_*\mathbf{Q})(n)$. For $s \in S(L)$, let $\gamma_B(s) \in H_B^{2n}(X_s)(n)$ be the restriction of γ_B to the fiber $X_s = \pi^{-1}(s)$. Let $s_0 \in S(K)$. Then:*

(1) *Suppose $K = \mathbf{C}$. If $\gamma_B(s_0)$ is a Hodge class, then $\gamma_B(s)$ is a Hodge class for all $s \in S(\mathbf{C})$.*
(2) *Suppose $K = \mathbf{C}$. If $\gamma_B(s_0)$ is an absolute Hodge class, then $\gamma_B(s)$ is an absolute Hodge class for all $s \in S(\mathbf{C})$.*
(3) *Suppose $K \subseteq \overline{\mathbf{Q}}$. If $\gamma_B(s_0)$ is De Rham, then $\gamma_B(s)$ is De Rham for all $s \in S(\overline{\mathbf{Q}})$.*

(3.2) PROOF. The Leray spectral sequence degenerates at E_2 and provides a surjection $\alpha : H_B^{2n}(X)(n) \to H^0(S, \mathbf{R}^{2n}\pi_*\mathbf{Q})(n)$ whose kernel we denote K_B. For $s \in S(K)$, the restriction $\beta_s : H^0(S, \mathbf{R}^{2n}\pi_*\mathbf{Q})(n) \to H_B^{2n}(X_s)(n)$ is injective. Let $\widehat{\gamma_B} \in H_B^{2n}(X)(n)$ satisfy $\alpha(\widehat{\gamma_B}) = \gamma_B$. Then $\gamma_B(s) = \beta_s \circ \alpha(\widehat{\gamma_B})$, and the kernel of $\beta_s \circ \alpha : H_B^{2n}(X)(n) \to H_B^{2n}(X_s)(n)$ equals K_B and is independent of $s \in S(K)$. Since $\beta_s \circ \alpha$ is a morphism of mixed Hodge structures, $\beta_s \circ \alpha$ identifies $H_B^{2n}(X)(n)/K_B$ with a pure sub-Hodge structure of $H_B^{2n}(X_s)(n)$ which is independent of s. Let $\overline{\gamma_B}$ be the image of $\widehat{\gamma_B}$ in $H_B^{2n}(X)(n)/K_B$. Since $\beta_s \circ \alpha(\widehat{\gamma_B})$ is a Hodge class, so is $\overline{\gamma_B}$. Hence $\beta_s \circ \alpha(\overline{\gamma_B}) = \gamma_B(s)$ is a Hodge class for all $s \in S(\mathbf{C})$. This proves the first claim.

We now prove the second claim. Let $\widehat{\gamma_{DR}} = I_\infty(\widehat{\gamma_B})$, $\widehat{\gamma_p} = I_p(\widehat{\gamma_B})$, $K_{DR} = I_\infty(K_B \otimes \mathbf{C})$, and $K_p = I_p(K_B \otimes \mathbf{Q}_p)$. Restricting to X_{s_0}, we have $\widehat{\gamma_{DR}}(s_0) = I_\infty(\gamma_B(s_0))$ and $\widehat{\gamma_p}(s_0) = I_p(\gamma_B(s_0))$. Let $\sigma \in \mathrm{Aut}(\mathbf{C})$. Then $\sigma(K_p)$ is the kernel of restriction $H_p^{2n}(\sigma X)(n) \to H_p^{2n}(X_{\sigma(s)})(n)$ and σK_{DR} is the kernel of restriction $H_{DR}^{2n}(\sigma X)(n) \to H_{DR}^{2n}(X_{\sigma(s)})(n)$ for all $s \in S(\mathbf{C})$. Since $\sigma(\widehat{\gamma_p})$ restricts to $\sigma(I_p(\gamma_B(s_0)))$, $\gamma^* = I_p^{-1}(\sigma(\widehat{\gamma_p})) \in H_B^{2n}(\sigma X)(n) \otimes \mathbf{Q}_p$ restricts to $I_p^{-1}(\sigma\gamma_p(s_0)) = \sigma\gamma_B(s_0)$. But if $\phi : V \to W$ is a linear map of rational vector spaces and there exists $v \in V \otimes \mathbf{Q}_p$ such that $\phi(v) = w \in W$, then there exists $v' \in V$ such that $\phi(v') = w$. Hence there exists $\widehat{\gamma_B}(\sigma) \in H_B^{2n}(\sigma X)(n)$ whose restriction to $H_B^{2n}(X_{\sigma(s_0)})(n)$ is $\sigma\gamma_B(s_0)$. Note that (i) $I_p(\widehat{\gamma_B}(\sigma)) - \sigma(\widehat{\gamma_p}) \in \sigma(K_p)$ and (ii) $I_\infty(\widehat{\gamma_B}(\sigma)) - \sigma(\widehat{\gamma_{DR}}) \in \sigma(K_{DR})$. Let $K_B(\sigma)$ denote the kernel of restriction $H_B^{2n}(\sigma X)(n) \to H_B^{2n}(X_{\sigma(s_0)})(n)$. Then $I_p(K_B(\sigma)) \otimes \mathbf{Q}_p =$

$\sigma(K_p)$ and hence $K_B(\sigma)$ is also the kernel of restriction $H_B^{2n}(\sigma X)(n) \to H_B^{2n}(X_{\sigma(s)})(n)$ for all $s \in S(\mathbf{C})$. We now argue as before. Let $\overline{\gamma_B}(\sigma)$ be the image of $\widehat{\gamma_B}(\sigma))$ in $H_B^{2n}(\sigma X)(n)/K_B(\sigma)$. Then $\overline{\gamma_B}(\sigma)$ is a Hodge class, and hence its image $\overline{\gamma_B}(\sigma, s) \in H_B^{2n}(X_{\sigma(s)})(n)$ is a Hodge class for every $s \in S(\mathbf{C})$. Finally, that $I_\infty(\overline{\gamma_B}(\sigma, s)) = \sigma(\gamma_{\mathrm{DR}}(s))$ and $I_p(\overline{\gamma_B}(\sigma, s)) = \sigma(\gamma_p(s))$ follows at once from (i) and (ii) above. This proves the second part.

To prove the third part, let $\sigma_p : \overline{\mathbf{Q}} \to \mathbf{C}_p$ be an embedding and define the maps I_{DR} relative to this σ_p. Let $\overline{X}$ be a smooth compactification of X, defined over K, with $\overline{X} - X$ a union of smooth divisors having normal crossings. By **[D4,** Theorem 4.1.1], the natural map

$$\tilde{\alpha} : H_B^{2n}(\overline{X})(n) \to H^0(S, \mathbf{R}^{2n}\pi_*\mathbf{Q})(n)$$

is surjective. As before, this means that the kernel $\widetilde{K_B}$ of the restriction $\beta_s \circ \tilde{\alpha} : H_B^{2n}(\overline{X})(n) \to H_B^{2n}(X_s)(n)$ is independent of $s \in S(L)$. Let $\widetilde{K_{\mathrm{DR}}} \subseteq H_{\mathrm{DR}}^{2n}(\overline{X})(n)$ be the kernel of restriction to $H_{\mathrm{DR}}^{2n}(X_{s_0})(n)$. Since $\widetilde{K_{\mathrm{DR}}} = I_\infty(K_B \otimes \mathbf{C})$, it is also the kernel of restriction for all $s \in S(L)$.

Let $\overline{\gamma_B} \in H_B^{2n}(\overline{X})(n)$ restrict to γ_B. Then $\overline{\gamma_B}(s_0) = \gamma_B(s_0)$ is a De Rham class by hypothesis. Replacing K by a finite extension, if necessary, we see that the restriction to X_{s_o} of $I_\infty(\overline{\gamma_B})$ equals $I_\infty(\overline{\gamma_B}(s_0)) \in H_{\mathrm{DR}}^{2n}(X_{s_o})(n)$. Hence, there exists $\overline{\gamma_{\mathrm{DR}}} \in H_{\mathrm{DR}}^{2n}(\overline{X})(n)$ such that $\overline{\gamma_{\mathrm{DR}}}(s_0) = I_\infty(\overline{\gamma_B}(s_0))$. Similarly, there exists $\overline{\gamma_p} \in H_p^{2n}(\overline{X})(n)$ such that $\overline{\gamma_p}(s_0) = I_p(\overline{\gamma_B}(s_0))$. By assumption, we have $I_{\mathrm{DR}}(\sigma_p\overline{\gamma_p}(s_0)) = \sigma_p\overline{\gamma_{\mathrm{DR}}}(s_0)$. Hence $I_{\mathrm{DR}}(\sigma_p\overline{\gamma_p}) - \overline{\gamma_{\mathrm{DR}}}$ belongs to $\sigma_p\widetilde{K_{\mathrm{DR}}} \otimes B_{\mathrm{DR}}$. Let $s \in S(L)$. Then

$$I_{\mathrm{DR}}(\sigma_p(\overline{\gamma_p}(s))) = I_{\mathrm{DR}}(\sigma_p\overline{\gamma_p})(s) = \sigma_p(\overline{\gamma_{\mathrm{DR}}}(s))$$

since I_{DR} commutes with restriction and $\sigma_p\widetilde{K_{\mathrm{DR}}}$ restricts to 0. Since p, σ_p, and $s \in S(L)$ are arbitrary, we are done.

4. Completion of the proof

In this section, we review the objects and steps of Deligne's proof of his absolute Hodge cycles theorem, giving also the extension to the p-adic maps I_{DR}. However, our exposition is not fully self-contained and the reader will need to consult **[D1]** to fill in the details.

Let A be an abelian variety. We have attached to A the three Tannakian categories $\otimes_{\mathrm{H}}A$, $\otimes_{\mathrm{AH}}A$, and $\otimes_{\mathrm{DR}}A$, and hence three groups

$$\mathscr{G}_{\mathrm{H}} \subseteq \mathscr{G}_{\mathrm{AH}} \subseteq \mathscr{G}_{\mathrm{DR}}.$$

By Proposition (2.11), it is enough for us to prove

$$\mathscr{G}_{\mathrm{H}} = \mathscr{G}_{\mathrm{AH}} = \mathscr{G}_{\mathrm{DR}}.$$

(4.1) The CM case: a reduction. Let K be a CM field (i.e., a totally imaginary quadratic extension of a totally real number field) which is Galois over $\mathbf{Q}$. A *CM type* of K is a set Φ of complex embeddings of K such that $\Phi \cup \Phi\rho$ is all embeddings and $\Phi \cap \Phi\rho$ is empty. Embedding K into $\mathbf{C}^{\Phi}$ by sending $k \in K$ to $z(k) \in \mathbf{C}^{\Phi}$ defined by $z(k(\sigma)) = \sigma(k)$, the ring of integers $\mathscr{O}_K$ of K becomes a lattice in $\mathbf{C}^{\Phi}$ and $A_{\Phi} = \mathbf{C}^{\Phi}/\mathscr{O}_K$ is an abelian variety. Let S be the set of all CM types. By definition, an abelian variety of CM type is an abelian variety that is isogenous to a product of such A_{Φ}, where K is allowed to vary. All such abelian varieties admit projective models defined over $\overline{\mathbf{Q}}$.

Note that, if $A \subseteq B$ and if the above equality of groups holds for B, then it holds for A, by Proposition (2.11). Furthermore, the equality (i) holds for A if and only if it holds for A^n with any positive integer n, and (ii) holds for A if and only if it holds for any abelian variety isogenous to A. Note finally that if A has an action of K, then A^n carries an action of any field extension of K of degree n. Hence, starting from any A of CM type we can find a Galois CM extension L of $\mathbf{Q}$ such that A is isogenous to a quotient of a power of

$$B = \prod_{\Phi \in S} A_{\Phi}$$

where S denotes the set of all CM types of L. By the above remarks, to prove our claim for A, it is enough to prove it for B. (Cf. **[D1,** p. 65] for more details about this reduction.)

(4.2) Special De Rham classes. Deligne finds three special types of absolute Hodge classes:

[1] Classes of the graphs of endomorphisms given by the evident embedding of L into $\mathrm{End}(A_{\Phi})$ for each Φ.

[2] Let $\sigma \in \mathrm{Gal}(L/\mathbf{Q})$. Via $\sigma : L \to L$, each A_{Φ} becomes also of type $A_{\Phi\sigma}$, and we have a natural isomorphism of A_{Φ} with $A_{\Phi\sigma}$.

[3] Let $T \subseteq S$ have d elements. Let

$$B_T = \prod_{\Phi \in T} A_{\Phi}.$$

Suppose that for the action of L on $H^{10}(B_T)$ each embedding of L occurs with equal multiplicity, necessarily equal to $d/2$. Then

$$\bigwedge_L^d H_B^1(B_T)(d/2) \subseteq H_B^d(B_T)(d/2)$$

consists of De Rham classes.

The classes of types [1] and [2] are De Rham because they are algebraic and a principal point of the argument of **[D1]** is to show that the classes of type [3] are absolutely Hodge using Principle B. To do this, one first constructs a universal family of abelian varieties $\pi : \mathscr{A} \to X$ parametrized by an arithmetic quotient X of the symmetric space attached to a certain unitary group in d variables associated to the quadratic extension defined by L relative to its maximal totally real subfield. This family contains B_T, carries an action of $\mathscr{O}_L$, and has a fiber of the form $A_0^{[L:\mathbf{Q}]}$ where the L action is that defined by an embedding of L into

$$M_{[L:\mathbf{Q}]}(\mathbf{Q}) \subseteq \operatorname{End}(A_0^{[L:\mathbf{Q}]}).$$

The association to $x \in X(\mathbf{C})$ of $\bigwedge_L^d H_B^1(A_x)(d/2)$ is a constant local subsystem on X whose global sections lie in $H^0(X, R^d\pi_*\mathbf{Q})$. Further, at the point x_0 with fiber $A_0^{[L:\mathbf{Q}]}$ all the elements of $\bigwedge_L^d H_B^1(A_{x_0})(d/2)$ are algebraic and, hence, absolutely Hodge and De Rham. In fact, this space is generated over L by the class of the cycle $A_0^{[L:\mathbf{Q}]-1} \times \{0\} \subset A_0^{[L:\mathbf{Q}]}$. Thus, Principle B applies: the classes of type [3] are absolutely Hodge. To see that they are De Rham we need only note also that (i) $\pi : \mathscr{A} \to X$ is defined over a number field, since it is a universal family attached to a moduli problem of PEL type, (ii) the point of X corresponding to B_T is algebraic, since B_T is of CM type, and (iii) the A_0 of Deligne's construction is, up to isogeny, any abelian variety of dimension $d/2$, and hence both A_0 and x_0 can be taken to be defined also over a number field. Thus, Principle B applies and the classes in $\bigwedge_L^d H_B^1(B_T)(d/2)$ are De Rham.

(4.3) Completion of the CM case. Note first that the classes of type [1] force

$$\mathscr{G}_{\mathrm{H}} \subseteq \mathscr{G}_{\mathrm{AH}} \subseteq \mathscr{G}_{\mathrm{DR}} \subseteq \left(\prod_{\Phi} \mathrm{L}^*\right) \times G_m \stackrel{\mathrm{def}}{=} G.$$

Next, observe that $\mathscr{G}_{\mathrm{H}}$ can be explicitly described: it is the $\mathbf{Q}$-Zariski closure in $G(\mathbf{C})$ of the cocharacter $\mu : G_m/\mathbf{C} \to G/\mathbf{C}$ that acts on the $(1, 0)$ classes in $H_B^1(A_\Phi)$ by sending z to z^{-1}, acts on the $(0, 1)$ classes trivially, and projects to the identity on the G_m factor. Since G acts on $H_B^1(B)$ and $\mathbf{Q}(1)_B$ via projection on the first and second factors, respectively, it acts on all tensor expressions

$$(H_B^1(B)^{\otimes n} \otimes H_B^1(B)^{\vee \otimes m})(k) = W$$

and the subspaces of W_{H}, W_{AH}, and W_{DR} of Hodge, absolutely Hodge, and De Rham classes are stable for this action. If $\gamma \in W_{\mathrm{H}} \otimes \overline{\mathbf{Q}}$ (resp. $W_{\mathrm{AH}} \otimes \overline{\mathbf{Q}}$, resp. $W_{\mathrm{DR}} \otimes \overline{\mathbf{Q}}$) transforms according to the character χ of $G/\overline{\mathbf{Q}}$, then χ has trivial restriction to $\mathscr{G}_{\mathrm{H}}$ (resp. $\mathscr{G}_{\mathrm{AH}}$, resp. $\mathscr{G}_{\mathrm{DR}}$). The classes of types [2] and [3] are De Rham and therefore provide an explicit submodule $\mathscr{X}$ of the character group $\mathscr{X}_G$ of G whose elements restrict trivially to $\mathscr{G}_{\mathrm{DR}}$. On the other hand, using the explicit description of $\mathscr{G}_{\mathrm{H}}$ via μ, Deligne shows by linear algebra that any element of $\mathscr{X}_G$ that restricts trivially to $\mathscr{G}_{\mathrm{H}}$ belongs to $\mathscr{X}$. Thus, every character of G that restricts trivially to $\mathscr{G}_{\mathrm{H}}$ also restricts trivially to $\mathscr{G}_{\mathrm{DR}}$. Hence $\mathscr{G}_{\mathrm{DR}} \subseteq \mathscr{G}_{\mathrm{H}}$ and so $\mathscr{G}_{\mathrm{DR}} = \mathscr{G}_{\mathrm{H}}$ and we are done.

(4.4) Completion of the proof. Let A be an abelian variety, not of CM type. The data $(\mathscr{G}_{\mathrm{H}}, \mu)$, with $\mu : G_m/\mathbf{C} \to \mathscr{G}_{\mathrm{H}}/\mathbf{C}$ defined as above for A, define (choosing additional structure, in particular a sufficiently small open compact subgroup U of $\mathscr{G}_{\mathrm{H}}(\mathbf{A}_f)$) a Shimura variety Sh_U which carries a natural family of abelian varieties $\pi : \mathscr{A} \to Sh_U$, such that there exists $s_0 \in Sh_U(\mathbf{C})$ for which $\pi^{-1}(s_0)$ is isogenous to A. Identify $\pi^{-1}(s_0)$ with A. If

$$\gamma_{s_0} \in (H^1_B(A)^{\otimes n} \otimes H^1_B(A)^{\vee\otimes m})(\frac{n-m}{2}) = W_{s_0}$$

is a Hodge class, then this family has the property that γ_{s_0} extends to a global section γ_B of $H^0(Sh_U, R^{m+n}\pi_*\mathbf{Q})$, where we have used the identification $H^1_B(C)^\vee = H^1_B(C)(1)$ for any abelian variety C. Further, since A is not of CM type, $\dim(Sh_U) > 0$ and, by a general principle (cf. **[D1]**), there exists $s_1 \in Sh_U(\mathbf{C})$ such that $\pi^{-1}(s_1)$ is of CM type. Hence, by Principle B, $\gamma_B(s_0)$ is an absolute Hodge class since this is true of $\gamma_B(s_1)$.

To obtain the De Rham result, we must only check that $\pi : \mathscr{A} \to Sh_U$ and s_1 are defined over $\overline{\mathbf{Q}}$. This is clear for Sh_U itself since it is a Shimura variety, and to see it for the family, one may remark either: (1) that $\pi : \mathscr{A} \to Sh_U$ is the universal family attached to a moduli problem defined by absolute Hodge cycles, polarization, and level structure or (2) that there is a natural $\overline{\mathbf{Q}}$-embedding of Sh into $\mathscr{A}_{\delta,n}$, the moduli space of abelian varieties of dimension $n = \frac{d}{2}[L : \mathbf{Q}]$ of polarization degree δ (determined by the degree of the polarization chosen on A) and that $\pi : \mathscr{A} \to Sh_U$ is just the pullback of the universal family over $\mathscr{A}_{\delta,n}$. Finally, it is clear that s_1 is defined over $\overline{\mathbf{Q}}$: from the first viewpoint, it is the modulus point on Sh_U associated to $\pi^{-1}(s_1)$ and its additional structure, and since $\pi^{-1}(s_1)$ is of CM type, this data can have only finitely many distinct isomorphism classes of conjugates under the elements of $\mathrm{Aut}(\mathbf{C})$; from the second viewpoint, the image of s_1 in $\mathscr{A}_{\delta,n}$ is algebraic, by the same argument, and hence so is s_1. Hence, we can apply Principle B to conclude that γ_{s_0} is De Rham.

5. A crystalline consequence

(5.1) Let X be a proper smooth variety defined over the number field K. Suppose that $\sigma_p X \times_{\sigma_p(K)} \widehat{\sigma_p(K)}$ has good reduction. Then

(1) the crystalline cohomology groups $H^j_{\mathrm{cris}}(\sigma_p X)$ are defined for all $j \geq 0$. They are vector spaces over the maximal subextension $W(\sigma_p)$ of $\widehat{\sigma_p(K)}$ that is unramified over $\mathbf{Q}_p$. Each carries a $\mathbf{Q}_p$ linear automorphism Φ that is ϕ semilinear, where ϕ denotes the Frobenius automorphism of $W(\sigma_p)$: $\Phi(\alpha v) = \phi(\alpha)\Phi(v)$ for $v \in H^j_{\mathrm{cris}}(\sigma_p X)$ and $\alpha \in W(\sigma_p)$.

(2) There is a canonical identification

$$H^j_{\text{cris}}(\sigma_p X) \otimes_{W(\sigma_p)} \widehat{\sigma_p(K)} = H^j_{\text{DR}}(\sigma_p X).$$

(3) Let B_{cris} denote the algebra introduced by Fontaine in **[Fo]**. It contains the maximal unramified extension of $\mathbf{Q}_p$ inside $\mathbf{C}_p$, and it carries a D_{σ_p} action extending the natural action on the maximal unramified extension. Further, it carries an automorphism Φ_{cris} that extends the action of Frobenius on the maximal unramified extension and commutes with the action of D_{σ_p}. Then there is B_{cris} linear isomorphism

$$I_{\text{cris}} : H^j_p(\sigma_p X) \otimes_{\mathbf{Q}_p} B_{\text{cris}} \to H^j_{\text{cris}}(\sigma_p X) \otimes_{W(\sigma_p)} B_{\text{cris}}$$

that is D_{σ_p} equivariant. Here we use the same definitions for the D_{σ_p} action on each side as in the De Rham case. The isomorphism is compatible with products and with cycle maps.

(4) The isomorphism I_{cris} satisfies

$$I_{\text{cris}} \circ (1 \otimes \Phi_{\text{cris}}) = (\Phi \otimes \Phi_{\text{cris}}) \circ I_{\text{cris}}.$$

(5) We have

$$I_{\text{cris}} \otimes 1 = I_{\text{DR}}.$$

REMARKS. The first two properties are basic to the theory of crystalline cohomology. The third and fourth are fundamental results of **[F1]**, and the fifth, while not explicitly claimed in **[F1]**, follows easily **[F2]** from the compatibility of the constructions of I_{cris} and I_{DR}.

We extend these assertions to the Tate-twisted case by setting

$$H^j_{\text{cris}}(\sigma_p X)(k) = H^j_{\text{cris}}(\sigma_p X)$$

and by putting

$$\Phi_{H^j_{\text{cris}}(\sigma_p X)(k)} = p^{-k} \Phi_{H^j_{\text{cris}}(\sigma_p X)}.$$

(5.2) Let $\gamma_B \in H^{2j}_B(X)(j)$ be a De Rham class that is defined over the number field K. Let $\sigma_p : K \to \mathbf{C}_p$ be an embedding. We say that γ_B is **crystalline** at σ_p if

(1) X has good reduction at σ_p,

(2) γ_{DR} belongs to the crystalline subspace $H^{2j}_{\text{cris}}(\sigma_p X)(j)$ of $H^{2j}_{\text{DR}}(\sigma_p X)(j)$,

(3) $\Phi(\gamma_{\text{DR}}) = \gamma_{\text{DR}}$.

(5.3) THEOREM. *Let A be an abelian variety defined over K with good reduction at σ_p. Let $\gamma_B \in H^{2j}_B(A)(j)$ be a Hodge class defined over K. Then γ_B is crystalline at σ_p.*

PROOF. We have

$$\sigma_p \gamma_{\text{DR}} = I_{\text{DR}}(\sigma_p \gamma_p) = I_{\text{cris}}(\sigma_p \gamma_p) \in H^{2j}_{\text{cris}}(\sigma_p A)(j),$$

thus proving that γ_{DR} belongs to the crystalline subspace. Since

$$\Phi(\sigma_p\gamma_{\mathrm{DR}}) = \Phi(I_{\mathrm{cris}}(\sigma_p\gamma_p)) = I_{\mathrm{cris}}((1\otimes\Phi_{\mathrm{cris}})(\sigma_p\gamma_p\otimes 1)) = I_{\mathrm{cris}}(\sigma_p\gamma_p) = \sigma_p\gamma_{\mathrm{DR}},$$

we see that (3) holds.

REFERENCES

[Co] P. Colmez, *Périodes des variétés abeliennes à multiplications complexes*, preprint.

[DS] E. DeShalit, *Monomial relations between p-adic periods*, J. Reine Angew. Math. **374** (1987), 193–207.

[D1] P. Deligne, *Hodge cycles on abelian varieties* (*Notes by J. S. Milne*), Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer-Verlag, New York, 1982, pp. 9–100.

[D2] ——, *Valeurs de fonctions L et périodes d'intégrales*, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 313–346.

[D3] ——, *Motifs et groupe de Taniyama*, Hodge Cycles, Motives and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer-Verlag, New York, 1982, pp. 261–279.

[D4] ——, *Théorie de Hodge.* II, Inst. Hautes Études Sci. Publ. Math. **40** (1972), 5–57.

[DM] P. Deligne and J. Milne, *Tannakian categories*, Hodge Cycles, Motives and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer-Verlag, New York, 1982, pp. 100–228.

[F1] G. Faltings, *Crystalline cohomology and p-adic Galois representations*, Algebraic Analysis, Geometry, and Number Theory (J. I. Igusa, ed.), Johns Hopkins Univ. Press, 1990, pp. 25–79.

[F2] ——, personal communication, 1992.

[Fo] J.-M. Fontaine, *Sur certains types de représentations p-adiques du groupe de Galois d'un corps local; construction d'un anneau de Barsotti-Tate*, Ann. of Math. (2) **115** (1982), 529–577.

[G] R. Gillard, *Rélations entre périodes p-adiques*, Invent Math. **93** (1988), 355–381.

[L] R. P. Langlands, *Automorphic representations, Shimura varieties, and motives*, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 205–246.

[O1] A. Ogus, *Hodge cycles and crystalline cohomology*, Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer-Verlag, New York, 1982, pp. 357–414.

[O2] ——, *A p-adic analogue of the Chowla-Selberg formula*, *p*-adic Analysis (F. Baldassari, S. Borch, and B. Dwork, eds.), Lecture Notes in Math., vol. 1454, Springer-Verlag, New York, 1990.

[Sc] N. Schappacher, *Periods of Hecke characters*, Lecture Notes in Math., vol. 1301, Springer-Verlag, New York, 1988.

[Sh] G. Shimura, *Automorphic forms and the periods of abelian varieties*, J. Math. Soc. Japan **31** (1979), 561–592.

[W1] J. P. Wintenberger, *Torseur entre cohomologie étale p-adique et cohomologie cristalline; le cas abelien*, Duke Math. J. **62** (1991), 511–526.

[W2] ——, *Théorème de comparaison p-adique pour les schémas abéliens*, prepublications Paris-sud 91-26, Inst. Hautes Études Sci. Publ. Math. (to appear).

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

E-mail address: `blasius@math.ucla.edu`

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Drinfeld Modules: Cohomology and Special Functions

DAVID GOSS

ABSTRACT. The Betti, v-adic, and de Rham cohomologies of a Drinfeld module and its generalization, T-module, are discussed here. These T-modules are shown to give rise to a category of T-motives which, under mild conditions, possesses a tensor product. We discuss Γ-functions and the L-series associated to T-modules. We show how certain zeta-values are related to the tensor powers of the Carlitz module; this is the theory of Anderson-Thakur. We then present the results of J. Yu which establish transcendency results that are not yet known classically. Finally, we mention the very recent work of L. Denis on rational points of "Fermat families".

Introduction

Let $\mathbb{F}$ be a finite field and X a smooth, projective, geometrically connected variety over $\mathbb{F}$. As is very well known, the construction of "algebraic" cohomology theories for X was necessitated by the conjectures of Weil on the number of $\mathbb{F}$-rational points of X. Once these constructions were fairly well understood, Grothendieck realized that they should, perhaps, have a deeper geometric significance and thus was led to his "motives".

The original proofs of these "Weil conjectures" were for the cases of elliptic curves (Hasse), and curves and abelian varieties (Weil himself). These proofs use the "ℓ-adic Tate module" of ℓ-division points of the abelian variety, for ℓ a prime number not equal to the characteristic of $\mathbb{F}$. In turn, the Tate module performs so well as a first homology group because abelian varieties are essentially linear objects. That is, over $\mathbb{C}$, abelian varieties can be given analytically as a quotient of an affine space by a lattice M which is essentially the first Betti homology H_1. Therefore, $H_1 \otimes \mathbb{Z}_\ell$ can be realized purely algebraically from the group structure of the abelian variety. On the other hand, the proof of the conjectures for curves lies in the association to a curve of its *Jacobian* J. This abelian variety serves as a—for lack of a more descriptive phrase—"living" version of its one-dimensional homology.

1991 *Mathematics Subject Classification*. Primary 11G09; Secondary 11R58.
Partially supported by NSA Grant MDA 904-91-H-0054.
This paper is in final form and no version of it will be submitted for publication elsewhere.

The theorems proved for abelian varieties are thus translated into theorems for curves. Therefore, as is well known, for curves and abelian varieties one is able to work with relatively elementary constructions and the general machinery of ℓ-adic cohomology is not needed.

Now let A be an abelian variety over k where k is a number field. One defines the L-series of A, $L(A, s)$, as an Euler product over all places of k; the finite places are handled by the ℓ-adic Tate modules and the infinite places are handled by de Rham cohomology **[Se1]**. The "glue" that unites all of these theories is their canonical isomorphisms with *Betti* cohomology. While there are other important cohomology theories, these three theories suffice for a working theory. Each cohomology theory is viewed as a concrete "realization" of the "motive" underlying the cohomology.

Now let $\mathbf{A} = \mathbb{F}_r[T]$ be the polynomial ring over the finite field with r elements. Let $\mathbf{k} = \mathbb{F}_r(T)$ and $\mathbf{K} = \mathbb{F}_r((\frac{1}{T}))$ which we view as a complete topological field. As is very elementary to see, $\mathbf{A}$ is a *discrete* subring of $\mathbf{K}$. It therefore makes sense to talk about $\mathbf{A}$-*lattices* sitting inside the algebraic closure $\overline{\mathbf{K}}$ of $\mathbf{K}$. Let M be one such lattice. Then one can form the quotient $\overline{\mathbf{K}}/M$ as a (rigid) analytic space. Now let

$$e(z) = z \prod_{0 \neq \alpha \in M} (1 - z/\alpha);$$

one shows that $e_M(z)$ is both entire and an $\mathbb{F}_r$-linear mapping. Thus, as analytic spaces,

$$\overline{\mathbf{K}}/M \simeq \overline{\mathbf{K}}.$$

(This will be less shocking when the reader realizes that $\overline{\mathbf{K}}/\mathbf{K}$ is *infinite dimensional.*) On the other hand, $\overline{\mathbf{K}}/M$ is an $\mathbf{A}$-module; thus, via $e_M(z)$, $\overline{\mathbf{K}}$ inherits a new $\mathbf{A}$-module structure. This procedure was first used by L. Carlitz **[C1]** in 1935 for certain rank-one lattices M. In 1974, however, Drinfeld **[Dr1]** presented a theory for lattices of all ranks; thus, these new actions are called "Drinfeld modules".

In §1, we discuss Drinfeld modules. We present the equivalence of these objects with lattices in the case of fields L of "generic characteristic". Almost by definition we shall see that Drinfeld modules over L possess the analogs of ℓ-adic and Betti cohomology; indeed the first, "v-adic cohomology" for $v \in \operatorname{Spec}(\mathbf{A})$, arises naturally from division points, and the second, "Betti cohomology", comes from lattices. We shall also show how Drinfeld modules have a "de Rham" theory. This originally arose from an idea of P. Deligne to use "universal additive extensions", but, remarkably, it also can be described through a differential formalism. Thus the basic trilogy of cohomology theories exists for Drinfeld modules.

The simplest Drinfeld module is the "Carlitz module" C, which is also presented in §1. We show how this object is an $\mathbf{A}$-analog of the classical motive $\mathbb{Z}(1)$. Finally, §1 contains some very nice ideas of E.-U. Gekeler on

Drinfeld modules over finite fields in the manner of the classical theory of abelian varieties.

Since Drinfeld modules have good cohomology, it makes sense to try to embed them into a larger theory of "motives". This was begun by G. Anderson **[A1]**. Our purpose here is to try to sketch the general theory of these "T-modules" and "T-motives". Our sources are **[A1, A2, AT1]**, and communications with their authors over the years.

The subject of T-(modules and motives) certainly deserves a substantial undertaking in order to create a definitive reference; due to the constraints of time, this article will **not** meet that standard. Rather, following in the nature of this conference, we will proceed rather loosely at times: Those important concepts that we did not, on occasion, have time enough to flesh out totally (or feel comfortable enough proving) we shall list as "ideas". We then sometimes give "hints" to their proofs hoping to entice the reader to complete the job (or, even, correct any misstatements). On the other hand, statements stated as "theorems" will (as usual) either be proved or referenced. Indeed, the reader should read *all* the material on motives with a rather critical eye. While the author has confidence that the material is substantially correct (and certainly correct in the large picture) there may indeed be errors coming from the difficulties of writing on a subject not one's own, with few sources and little time.

Having said all that, it should be clear to the reader that *all* responsibility for statements, proofs, philosophy, etc., given here rests *solely* with the present author. This is the author's interpretation of the material learned from the above sources. On the other hand, it must also be said that any credit for the concepts and results presented on T-motives truly belongs with the author of **[A1]** and **[A2]**.

In order to define motives, one first generalizes Drinfeld modules to higher-dimensional objects — these are the "T-modules". This is somewhat similar to going from elliptic curves to general abelian varieties. Let E be one such T-module. By dualizing, one sees that the space of morphisms (as algebraic groups) from E to $\mathbb{G}_a$ forms a module over the noncommutative ring $L[T, \tau]$, where τ is the rth power mapping. This module is the "*T-motive*" of E and one shows that this construction gives an anti-equivalence of categories. The motivic formulation gives a very concise way to describe cohomology and, importantly, allows us to introduce a tensor product into the theory (see, e.g. §2.6 for an application to a "Legendre formula").

In §3 we briefly introduce Γ-functions and the L-series of T-modules and Drinfeld modules. We present some of the work of D. Thakur (e.g., **[Th4]**) that puts the Γ-functions on a very solid foundation. We also discuss the appropriate notions of exponentiation and analyticity for the L-series. We show, in Remark 3.14, how one can get L-series which combine both classical motives and the motives discussed here, and we speculate a bit on the meaning of such general functions. Finally, we restrict ourselves to the zeta function $\zeta_{\mathbf{A}}(s)$ of **A**-itself. We discuss some of its properties and present some of the evidence that hints at still deeper structure. We also show how the classical formalism leads to focusing on the relationship between *special values* of $\zeta_{\mathbf{A}}(s)$ and the nth tensor power $C^{\otimes n}$ of C. This relationship is the theory of Anderson-Thakur **[AT1]**. Using it, Jing Yu **[Y3]** (see Theorem 3.19) established transcendency results for $\zeta_{\mathbf{A}}(s)$ at positive integers that are *far stronger* than anything known classically.

The theory of these L-series is still quite young by historical standards. The reader will note that very little is known (or even conjectured) when compared to the material presented at this conference. We hope that this is taken as an invitation to explore new and open areas.

Section 4 discusses how one can proceed for general **A**. Finally, §5 presents some very new work of L. Denis on "Fermat families". These are generalizations of the classical Fermat equations using the T-modules of §2. In the case where the T-module has dimension 1, we obtain the "Fermat equations" of **[Go7]**, and Denis is able to show "Fermat's Last Theorem" in this context. In the case where the T-module has higher dimension, we obtain systems of equations that, as of yet, do not have any classical counterparts.

It is my pleasure to thank Dinesh Thakur for reading early versions of this work and providing a number of helpful comments. I also wish to thank the referee for many helpful comments.

1. Drinfeld modules

1.1. Notation. Let $\mathscr{C}$ be a fixed, smooth, projective, geometrically irreducible curve over the finite field $\mathbb{F}_r$, $r = p^m$. Let $\infty \in \mathscr{C}$ be a closed point. The ring of functions holomorphic outside ∞ is denoted by $\mathbf{A}$. As is well known, $\mathbf{A}$ is a Dedekind domain with finite class group and unit group $\mathbb{F}_r^*$. We let $\mathbf{k}$ be the function field of $\mathscr{C}$ and $\mathbf{K} = \mathbf{k}_\infty =$ the completion of $\mathbf{k}$ at ∞. Let $|?|_\infty$ be the normalized absolute value at ∞.

The following lemma is well known and follows, for example, from the Riemann-Roch theorem. It will, perhaps, be helpful to the reader to work out the trivial case of $\mathbf{A} = \mathbb{F}_r[T]$.

PROPOSITION 1.1.1. *The ring* $\mathbf{A}$ *is a discrete subring of* $\mathbf{K}$. *It is also co-compact* (*i.e.,* $\mathbf{K}/\mathbf{A}$ *is compact*). □

We let "deg" denote the degree mapping on ideals and elements of $\mathbf{A}$. Thus if $a \in \mathbf{A}$, then $\mathbf{A}/(a)$ has $r^{\deg(a)}$ elements. Let

$$d_\infty = \deg(\infty).$$

If a has a pole of order t_a at ∞, then clearly

$$\deg(a) = t_a d_\infty .$$

The reader should keep in mind the analogy with the rational numbers $\mathbb{Q}$; so $\mathbf{A} \sim \mathbb{Z}$, $\mathbf{k} \sim \mathbb{Q}$, $\mathbf{K} \sim \mathbb{R}$, and, of course, $\mathbb{R}/\mathbb{Z} \simeq S^1$ is compact.

We let $\overline{\mathbf{K}}$ be a fixed algebraic closure of $\mathbf{K}$ equipped with the canonical extension of $|?|_\infty$. We let $\bar{\mathbf{k}} \subseteq \overline{\mathbf{K}}$ be the algebraic closure and $\mathbf{k}^s \subseteq \bar{\mathbf{k}}$ the separable closure.

The simplest, and most basic, example is the case $\mathbf{A} = \mathbb{F}_r[T]$, $\mathbf{k} = \mathbb{F}_r(T)$, and $\mathbf{K} = \mathbb{F}_r\left(\left(\frac{1}{T}\right)\right)$. It will be useful to the reader to keep this example in mind during the reading of this chapter.

Let $\mathbf{A}$ again be arbitrary. We let $\alpha_T \in \mathbf{A}$ be a *fixed* element with a pole of order $t_{\alpha_T} > 0$ with t_{α_T} prime to p. We thus obtain an injection

$$\mathbb{F}_r[T] \hookrightarrow \mathbf{A}, \qquad T \mapsto \alpha_T.$$

Since the order of the pole of α_T is prime to p one sees that $\mathbf{k}$ is a *separable* extension of $\mathbb{F}_r(T)$. Moreover, using the injection of $\mathbb{F}_r[T]$ into $\mathbf{A}$, we may define arithmetic objects firstly for $\mathbb{F}_r[T]$, and then for arbitrary $\mathbf{A}$ viewed as a ring of "complex multiplications".

We call a field, L, equipped with a morphism $\imath : \mathbf{A} \to L$, an "$\mathbf{A}$-field". The prime $\wp = \ker \imath$ is called the "characteristic" of L. If $\wp = (0)$, then L is said to have "generic characteristic".

1.2. Drinfeld modules. Let L be an $\mathbf{A}$-field. Let $\mathbb{G}_a$ be the additive group over L. Let $\mathrm{End}_{\mathbb{F}_r}(\mathbb{G}_a/L)$ be the endomorphism ring of $\mathbb{F}_r$-linear endomorphisms of $\mathbb{G}_a$ over L. The primary example is the mapping τ with

$$\tau(x) := \tau_r(x) := x^r \quad (\tau_p(x) := x^p).$$

In fact, one sees that $\mathrm{End}_{\mathbb{F}_r}(\mathbb{G}_a/L)$ is the ring of polynomials

$$f = \sum_{i=0}^{\nu} a_i \tau^i ,$$

with $\{a_i\} \subseteq L$ and τ^i is defined via composition. If $x \in L$, then $f(x) = \sum a_i x^{r^i}$, and we set

$$Df = a_0 = \text{derivative of } f(x) \text{ at } x = 0 .$$

The ring $\mathrm{End}_{\mathbb{F}_r}(\mathbb{G}_a/L)$ will also be denoted $L\{\tau\}$. It is *not* commutative in general because

$$(a\tau)(b\tau) = ab^r\tau^2 .$$

We can now define Drinfeld modules **[Dr1, DH1, K1]**. We begin with the case $\mathbf{A} = \mathbb{F}_r[T]$.

DEFINITION 1.2.1. A *Drinfeld module*, ϕ, of rank $d > 0$ is a homomorphism

$$\phi: A \to L\{\tau\}, \qquad a \mapsto \phi_a,$$

such that

$$\phi_T = \sum_{i=1}^{d} a_i^T \tau^i ,$$

where $\{a_i^T\} \subseteq L$, $a_d^T \neq 0$, and $a_0^T = \imath(T)$.

Let $\alpha \in \mathbf{A}$ have degree ν. The reader may easily check that

$$\phi_\alpha = \iota(\alpha)\tau^0 + \sum_{i=0}^{d\nu} a_i^\alpha \tau^i$$

where $a_{d\nu}^\alpha \neq 0$.

EXAMPLE 1.2.2. We present here the *Carlitz module* C. It is the simplest rank-one Drinfeld module for $\mathbf{A} = \mathbb{F}_r[T]$ and, historically, the first example of a Drinfeld module. It originally arose in the paper **[C1]** (but note that we have changed the definition slightly in order to have Frobenius elements in certain Galois groups correspond to monic primes). The module C is defined over $\mathbf{k}$ (in fact, over $\mathbf{A}$) and is given by

$$C_T = T\tau^0 + \tau .$$

Thus

$$C_{T^2} = T^2\tau^0 + (T^r + T)\tau + \tau^2 ,$$

etc. We shall see later (in Example 1.3.9) that C plays a role analogous to the classical motive $\mathbb{Z}(1)$.

Let ϕ_1, ϕ_2 be two Drinfeld modules for $\mathbb{F}_r[T]$ defined over the $\mathbf{A}$-field L.

DEFINITION 1.2.3. A *morphism* $\rho: \phi_1 \to \phi_2$ over L is an element $\rho \in L\{\tau\}$ such that

$$\rho\phi_1 = \phi_2\rho .$$

We let $\mathrm{Hom}_L(\phi_1, \phi_2)$ be the set of morphisms for ϕ_1 to ϕ_2 defined over L. It is clear that $\mathrm{Hom}_L(\phi_1, \phi_2)$ is an $\mathbf{A}$-module. Moreover, $R = \mathrm{End}_L(\phi)$ is an $\mathbf{A}$-algebra. It comes equipped with a homomorphism

$$R \to L, \qquad f \mapsto Df,$$

which obviously is also a mapping of $\mathbf{A}$-algebras. In general we shall drop the subscripted reference to L when we are interested in the *absolute* morphisms; i.e., defined over some fixed algebraic closure of L.

We can now give the definition of Drinfeld modules for general $\mathbf{A}$ equipped, as above, with the map $\mathbb{F}_r[T] \hookrightarrow \mathbf{A}$, $T \mapsto \alpha_T$. Let L be an $\mathbf{A}$-field via $\iota: \mathbf{A} \to L$. Clearly ι also makes L into an $\mathbb{F}_r[T]$-field.

DEFINITION 1.2.4. A *Drinfeld module* E for $\mathbf{A}$ defined over L is given by the following two pieces of data:

(1) A Drinfeld module ϕ for $\mathbb{F}_r[T]$ of positive rank defined over L.
(2) A homomorphism $\psi: \mathbf{A} \to \mathrm{End}_L(\phi)$, $\alpha \mapsto \psi_\alpha$, extending ϕ such that

$$D \circ \psi_\alpha = \iota(\alpha) .$$

REMARK 1.2.5. With a little thought, the reader can see that E is uniquely determined by the mapping ψ. Thus one sees that to give a Drinfeld module for $\mathbf{A}$ is the same as giving an $\mathbb{F}_r$-linear homomorphism $\psi: \mathbf{A} \to L\{\tau\}$ such that

$$\imath(\alpha) = D \circ \psi_\alpha$$

for $\alpha \in \mathbf{A}$ and such that for *some* $\alpha \in \mathbf{A}$

$$\psi_\alpha \neq \imath(\alpha)\tau^\circ .$$

For the rest of this section, we let $\mathbf{A}$ be arbitrary and $E = (\psi)$ be a Drinfeld $\mathbf{A}$-module over an $\mathbf{A}$-field L. We shall often refer to E as "ψ".

Morphisms of Drinfeld modules are defined exactly as in the $\mathbb{F}_r[T]$-case. Let ρ be a morphism between two Drinfeld modules. One sees that ρ is an *isomorphism* if and only if

$$\rho = \beta\tau^\circ$$

for some $\beta \in L^*$. Nonzero morphisms are called "isogenies". It is trivial to see that isogenies of Drinfeld modules are surjective with finite kernel as for abelian varieties.

Let $\alpha \in \mathbf{A}$. We write

$$\psi_\alpha = \imath(\alpha)\tau^0 + \sum_{j=1}^{\delta(\alpha)} a_j^\alpha \tau^j ,$$

where $a_{\delta(\alpha)}^\alpha \neq 0$. The mapping $\alpha \mapsto \delta(\alpha)$ satisfies

$$\delta(\alpha\beta) = \delta(\alpha) + \delta(\beta) ,$$

and

$$\delta(\alpha + \beta) \leq \max\{\delta(\alpha), \delta(\beta)\} .$$

Let $\wp$ be the characteristic of L. One can find $\lambda \in \mathbf{A} - \wp$ such that $\delta(\lambda) > 0$ and $\psi_\lambda(x)$ is a *separable* polynomial in x. The roots of $\psi_\lambda(x)$ in a separable closure L^s of L form a finite $\mathbf{A}$-module $\psi[\lambda]$. Since $\mathbf{A}$ is a Dedekind domain, we conclude that

$$\psi[\lambda] = (A/\lambda)^d ,$$

for some integer $d > 0$. Thus, from above, we deduce that

$$\delta(\alpha) = d \deg(\alpha)$$

for *all* α. The integer, d, is called the "rank of ψ". It agrees with our earlier definition in the $\mathbb{F}_r[T]$-case.

Let $\overline{L}$ be a fixed algebraic closure of L. Let $I \subseteq \mathbf{A}$ be an ideal.

DEFINITION 1.2.6. We set

$$E[I] = \psi[I] = \{t \in \overline{L} \mid \psi_i(t) = 0 \text{ for } i \in I\} .$$

The elements of $\psi[I]$ are called the *I-division points.* Recall that we set $\wp = \ker(\imath)$.

Proposition 1.2.7. *Let ψ have rank d. Then, as* **A**-*modules,*

$$\psi[I] \simeq A/I^e$$

where $e \le d$. Equality holds if and only if $\wp \nmid I$. □

In the next definition we view $\overline{L}$ as an **A**-module via ψ; i.e., if $a \in \mathbf{A}$ and $l \in \overline{L}$, then we have the action $(a, l) \mapsto \psi_a(l)$.

Definition 1.2.8. Let $v \in \operatorname{Spec}(\mathbf{A})$ be prime to $\wp$. We set

$$T_v(\psi) = \operatorname{Hom}_{\mathbf{A}}(\mathbf{k}_v/\mathbf{A}_v, \overline{L}).$$

It follows trivially from 1.2.7 that for ψ of rank d,

$$T_v(\psi) \simeq \mathbf{A}_v^d,$$

where $\mathbf{A}_v$ is the completion of **A** with respect to v.

The assignment $\psi \mapsto T_v(\psi)$ is easily seen to be a covariant functor. As with elliptic curves, we view $T_v(\psi)$ as a first "v-adic" homology with coefficients in $\mathbf{A}_v$.

Let $\mathbf{k}_v$ be the quotient field of $\mathbf{A}_v$.

Definition 1.2.9. We set

$$H_v^1(\psi, \mathbf{k}_v) = \operatorname{Hom}_{\mathbf{A}_v}(T_v(\psi), \mathbf{k}_v) \simeq \mathbf{k}_v^d.$$

In general, for $i \ge 0$, we set

$$H_v^i(\psi, \mathbf{k}_v) = \bigwedge^i H_v^1(\psi, \mathbf{k}_v).$$

Definition 1.2.9 is obviously based on the analogous *theorem* for abelian varieties. The functors $H_v^i(\psi, \mathbf{k}_v)$ are clearly contravariant.

Finally, we recall the following result. For a proof we refer the reader to **[DH1**, Theorem 4.9] (see also **[Dr1]**).

Proposition 1.2.10. *Let ψ be a rank d Drinfeld module over L. Then* $\operatorname{End}(\psi)$ *is a finitely generated projective* **A** *module of rank $\le d^2$. Moreover,* $\operatorname{End}(\psi) \otimes_{\mathbf{A}} \mathbf{K}$ *is a division algebra.* □

1.3. Analytic theory. We now let our **A**-field L be a finite extension of **K** inside our fixed algebraic closure $\overline{\mathbf{K}}$. Let $L^s \subseteq \overline{\mathbf{K}}$ be the separable closure. We continue to allow **A** to be arbitrary and $\imath$ is the obvious inclusion map.

Let $M \subseteq L^s$ be a finitely generated **A**-submodule of L^s. Since **A** is a Dedekind domain, and M is torsion-free, the module M has a *rank* that may be computed as the dimension of $M \otimes_{\mathbf{A}} \mathbf{k}$.

Definition 1.3.1. An *L-lattice M of rank d* is a discrete, rank d, **A**-submodule of L^s that is $\operatorname{Gal}(L^s/L)$ stable.

We occasionally drop the reference to L and just refer to M as a "lattice".

Definition 1.3.2. Let M be an L-lattice. We set

$$e_M(z) = z \prod_{\substack{\alpha \in M \\ \alpha \neq 0}} (1 - z/\alpha).$$

Since M is discrete, one sees easily that $e_M(z)$ is *entire*; that is, $e_M(z)$ is given by a power series with an infinite radius of convergence. Moreover, by separability, $e_M(z) \in L[[z]]$.

Since M is an **A**-module, we can write M as the increasing union $M = \bigcup M_i$, where M_i is a finite-dimensional $\mathbb{F}_r$-vector space. Put

$$e_{M_i}(z) = z \prod_{\substack{\alpha \in M_i \\ \alpha \neq 0}} (1 - z/\alpha);$$

thus $e_{M_i}(z)$ is a polynomial of finite degree, and

$$e_M(z) = \lim_i e_{M_i}(z).$$

Let $\tilde{e}_{M_i}(z) = \prod_{\alpha \in M_i}(z + \alpha_i) = c_i e_{M_i}(z)$, for some nonzero c_i. Put

$$f_i(z) = \tilde{e}_{M_i}(z + y) - \tilde{e}_{M_i}(y).$$

One checks readily that $f_i(z)$ and $\tilde{e}_{M_i}(z)$ have the same set of roots $= M_i$. Moreover, they are both monic; thus

$$f_i(z) = \tilde{e}_{M_i}(z),$$

and so

$$\tilde{e}_{M_i}(z) + \tilde{e}_{M_i}(y) = \tilde{e}_{M_i}(z + y).$$

Similarly, one see that $\tilde{e}_{M_i}(z)$ is also $\mathbb{F}_r$-linear. We therefore conclude that $e_M(z)$ is an $\mathbb{F}_r$-linear entire function.

Note that

$$e'_M(z) = 1,$$

for all z by additivity.

Now let $\alpha \in \mathbf{A}$. The entire functions

$$e_M(\alpha z) \quad \text{and} \quad e_M(z) \prod_{0 \neq \beta \in \alpha^{-1}M/M} (1 - e_M(z)/e_M(\beta))$$

can easily be seen to have the same divisors. Since we are using non-Archimedean analysis, we conclude that these functions must differ by a *constant*. By taking derivatives, we see that this constant is α, or

$$e_M(\alpha z) = \alpha e_M(z) \prod_{0 \neq \beta \in \alpha^{-1}M/M} \left(1 - \frac{e_M(z)}{e_M(\beta)}\right).$$

Thus we see that the assignment

$$\alpha \mapsto \rho_{M,\alpha}(z) = \alpha z \prod_{0 \neq \beta \in \alpha^{-1}M/M} \left(1 - \frac{z}{e_M(\beta)}\right)$$

is a Drinfeld module of rank d. In other words, we have transferred the **A**-module structure of $\overline{\mathbf{K}}/M$ to $\overline{\mathbf{K}}$ where it is given by polynomial mappings.

Conversely, let ρ be a Drinfeld module of rank d over L. By completing at the origin, one can associate to ρ its formal **k**-module $\hat{\rho}$. Drinfeld **[Dr1]** shows that this formal **k**-module is trivial; i.e., there exists an isomorphism between $\hat{\rho}$ and $\hat{\mathbb{G}}_a$ = formal $\mathbb{G}_a$. Thus one deduces the existence of a formal power series

$$e(z) = z + \sum_{j=1}^{\infty} c_j z^{r^j}, \quad \{c_j\} \subseteq L,$$

such that

$$e(\alpha z) = \hat{\rho}_\alpha(e(z))$$

for $\alpha \in \mathbf{k}$. That $e(z)$ is unique follows from the above functional equation for $e(z)$ as does the fact that $e(z)$ is entire. It is then not too difficult to see that

$$e(z) = z \prod_{\substack{\alpha \in M \\ \alpha \neq 0}} \left(1 - \frac{z}{\alpha}\right),$$

where M is an L-lattice of rank d (and is obviously the kernel of $e(z)$).

DEFINITION 1.3.3. Let M_1, M_2 be two L-lattices of rank d. A *morphism* from M_1 to M_2 is an element $\lambda \in L$ with $\lambda M_1 \subseteq M_2$.

The arguments sketched above then give **[Dr1]**:

THEOREM 1.3.4 (DRINFELD). *Let L be as above. Then the category of Drinfeld modules of rank d over L is isomorphic to the category of L-lattices of rank d.* □

COROLLARY 1.3.5. *Let L be as above and let ψ be a rank d Drinfeld module over L. Then $\mathrm{End}_L(\psi)$ is a commutative domain.*

PROOF. By 1.3.4, $\mathrm{End}_L(\psi)$ is isomorphic to a subring of L. □

In general, results on endomorphisms of Drinfeld modules are what would be expected from the analogy with elliptic curves (see also 1.2.10).

DEFINITION 1.3.6. Let L be as above and let ψ be a rank d Drinfeld module over L corresponding to the lattice M. We set

(1) $H_1(\psi) = M$ and $H_j(\psi) = \bigwedge^j H_1(\psi)$,

(2) $H_j^B(\psi, \mathbf{k}) = H_j(\psi) \otimes \mathbf{k}$,

(3) $H_B^1(\psi, \mathbf{k}) = \mathrm{Hom}_{\mathbf{A}}(H_1(\psi), \mathbf{k})$ and

$$H_B^j = \bigwedge^j H_B^1(\psi, \mathbf{k}).$$

It is easy to see, of course, that there is a perfect pairing

$$H_j^B(\psi, \mathbf{k}) \times H_B^j(\psi, \mathbf{k}) \to \mathbf{k}.$$

Let $v \in \mathrm{Spec}(\mathbf{A})$.

PROPOSITION 1.3.7. *As vector spaces we have* $H^i_v(\psi, \mathbf{k}_v) \simeq H^i_B(\psi, \mathbf{k}) \otimes_{\mathbf{k}} \mathbf{k}_v$.

PROOF. This follows immediately from the analytic description of ψ. □

REMARK 1.3.8. Let L be an arbitrary A-field and ψ a Drinfeld module over L. Let $v \in \operatorname{Spec}(\mathbf{A})$ be prime to the characteristic of L. Then we deduce a *continuous* action of $\operatorname{Gal}(\overline{L}/L)$ on $H^i_v(\psi, \mathbf{k}_v)$ via the action on $T_v(\psi)$. If $\mathbf{K} \subseteq L$, as in 1.3.7, then we also deduce an action of $\operatorname{Gal}(\overline{L}/L)$ on $H^i_B(\psi, \mathbf{k})$ via the action on the lattice of M. Proposition 1.3.7 then becomes an isomorphism of Galois modules.

EXAMPLE 1.3.9. Let $A = \mathbb{F}_r[T]$. We return to the Carlitz module C and finish giving the analogy of it with $\mathbb{Z}(1)$. Recall (1.2.2) that C was defined over $\mathbf{k}$ by

$$C_T = T\tau^\circ + \tau .$$

Let $[i] = T^{r^i} - T$ and, for $i > 0$, let

$$D_i = [i][i-1]^r \cdots [1]^{r^{i-1}} \in \mathbf{A}.$$

It is easy to see that

$$D_i = \prod_{\substack{\deg(n)=i \\ n \text{ monic}}} n .$$

From $e(Tz) = Te(z) + e(z)^r$, $e'(z) = 1$, one computes that the Carlitz exponential, $e(z)$, is equal to

$$z + \sum_{i=1}^{\infty} \frac{z^{r^i}}{D_i} .$$

Moreover, $e(z) = e_M(z)$ for $M = \mathbf{A}\xi$, where

$$\xi = (-T^r)^{1/(r-1)}(1 - T^{1-r})^{1/(r-1)} \prod_{i=1}^{\infty} \left(1 - \frac{[i]}{[i+1]}\right) ,$$

and here $(1-T^{1-r})^{1/(r-1)}$ is the *unique* 1-unit root of $1-T^{1-r}$. The element

$$(1 - T^{1-r})^{1/(1-r)} \prod_{i=1}^{\infty} \left(1 - \frac{[i]}{[i+1]}\right) .$$

is called the "1-unit piece of ξ".

The element ξ is *transcendental* over $\mathbf{k}$ by Wade **[W1]**. From this discussion, the reader may easily see the analogy of ξ with $2\pi i$.

Let f be a monic prime of $\mathbf{A}$. We set

$$\mathbf{k}(f^n) = \mathbf{k}(C[f^n]) ;$$

thus, $\mathbf{k}(f^n)$ is analogous to $\mathbb{Q}(\zeta_{p^n})$, **[C2, Ha2, Go1]**. In fact, the classical arguments show:

(1) $\mathbf{k}(f^n)/\mathbf{k}$ is abelian with $\operatorname{Gal}(\mathbf{k}(f^n)/\mathbf{k}) \simeq (A/f^n)^*$ via the action on $C[f^n]$.

(2) The prime f ramifies totally in $\mathbf{k}(f^n)$. All other finite primes are unramified.
(3) Let g be a monic prime $\neq f$ of $\mathbf{A}$. Let $\sigma_g \in \mathrm{Gal}(\mathbf{k}(f^n)/\mathbf{k}) \simeq (\mathbf{A}/f^n)^*$ be the Frobenius element associated to g. Then

$$\sigma_g = g + (f^n).$$

(4) The subgroup $\mathbb{F}_r^* \subset (\mathbf{A}/f^n)^*$ is the decomposition and inertia group of any infinite prime.

Thus, by §2.1 of **[D2]**, we see the analogy with $\mathbb{Z}(1)$. For more along these lines, see, for example, **[C2, Ha2, AT1, Go1, Go2, Go3, Go10]**.

1.4. Results on v-adic cohomology. Let L be a finite extension of $\mathbf{k}$. Let $\mathscr{O} \subseteq L$ be the ring of $\mathbf{A}$-integers. Let ψ be a fixed Drinfeld module of rank d over L.

We let $S \subseteq \mathrm{Spec}(\mathscr{O})$ be the set of "bad primes" for ψ. For $\wp \notin S$, this means the following: One can find $\alpha \in L^*$ such that

(1) $\psi_1 = \alpha\psi\alpha^{-1}$ has $\wp$*-integral* coefficients, and,
(2) the reduced mapping $\tilde{\psi}_1 : A \to \mathscr{O}/\wp\{\tau\}$ is a Drinfeld module of rank d.

Since $\mathbf{A}$ is finitely generated over $\mathbb{F}_r$, it is clear that almost all primes in $\mathrm{Spec}(\mathscr{O})$ are "good" (i.e., *not* bad). In fact, for almost all primes we can let $\alpha = 1$.

Let $G = \mathrm{Gal}(\overline{L}/L)$. Let $\wp$ be a good prime and v a prime of $\mathbf{A}$ $\neq \wp$. We have the continuous G-action on $H_v^1(\psi, \mathbf{k}_v)$ which is easily seen to be unramified at $\wp$. Conversely, by examining the equations necessary for the existence of α, as above, Takahashi was able to show the converse **[Ta1]**:

THEOREM 1.4.1 (TAKAHASHI). *The Drinfeld module ψ has good reduction at a finite prime $\wp$ if and only if, for $v \neq \wp$, the Galois module $H_v^1(\psi, \mathbf{k}_v)$ is unramified at $\wp$.* □

Clearly Theorem 1.4.1 is the analog of the classical criterion of *Néron-Ogg-Šafarevič* in the theory of abelian varieties. In this regard we also have **[T1]**:

THEOREM 1.4.2 (TAGUCHI). *$H_v^1(\psi, \mathbf{k}_v)$ is a semisimple $\mathbf{k}_v[G]$-module.* □

EXAMPLE 1.4.3. Let $\mathbf{A} = \mathbb{F}_r[T]$, $\mathbf{k} = \mathbb{F}_r(T)$. Let ρ be the rank 2 Drinfeld module given by

$$\rho_T = T\tau^\circ + T^{1-r}\tau + T^{1-r^2}\tau^2.$$

Put $\alpha = T^{-1}$. Then $\alpha\rho_T\alpha^{-1} = T\tau^\circ + \tau + \tau^2$, which obviously has good reduction everywhere.

For more on the reduction of Drinfeld modules, see **[Dr1, Dr2, DH1, Ta1]**. Next we discuss the theory of Drinfeld modules over finite fields. For a good account of these matters, we refer the reader to **[Ge3]** (see also **[Go10]**). We remark, however, that the next two theorems appear in, or are implicit in, Drinfeld's paper **[Dr2]**. Thus, let L now be a finite extension of $\mathbf{A}/\wp$ with $r_1 = r^t$ elements where

$$t = \deg(\wp)[L:\mathbf{A}/\wp].$$

Let ψ be a rank d Drinfeld module over L. Let F be the r_1th power mapping on $\mathbb{G}_a/L$. It is quite easy to check that F gives rise to an endomorphism of ψ.

Our proof of the next result comes from **[Ta1]**:

LEMMA 1.4.4. *Let $\mathscr{L}$ be an $\mathbf{A}$-field and ϕ a Drinfeld module over $\mathscr{L}$. Let α be an endomorphism of ϕ and $T_v(\alpha)$ the action of α on $T_v(\phi)$, $v \neq \operatorname{char}(L)$. Then the characteristic polynomial of $T_v(\phi)$ has coefficients in $\mathbf{A}$ which are independent of v.*

PROOF. The subring $\mathbf{A}[\alpha]$ of $\operatorname{End}(\phi)$ is a commutative domain. By 1.2.10, we see that $\operatorname{End}(\phi) \otimes \mathbf{K}$ is a division ring; so $\mathbf{k}(\alpha) \subseteq \operatorname{End}(\phi) \otimes \mathbf{K}$ must be a CM_∞-field; that is, $\mathbf{k}(\alpha)$ must be a field with a unique prime lying above ∞. Let $R \subseteq \mathbf{k}(\alpha)$ be the ring of $\mathbf{A}$-integers. By passing to an isogenous Drinfeld module if necessary **[Ha1**, Proposition 3.2], we may assume that ϕ has CM via R.

We conclude that $T_v(\phi)$ is a free $(R \otimes \mathbf{A}_v)$-module of finite type. The v-adic representation of α is thus induced by the representation $\alpha: \beta \mapsto \alpha\beta$ on R. The result follows readily. □

Let $Q(x) \in \mathbf{A}[x]$ be the characteristic polynomial of F. The above lemma assures us that $Q(x)$ is a power of the minimal polynomial of F as an endomorphism.

Proceeding in a similar fashion, we have:

THEOREM 1.4.5.

(1) $(Q(0)) = \wp^{[L:\mathbf{A}/\wp]}$.

(2) *Let $\{\alpha_i\}$ be the roots of $Q(x)$ in $\overline{\mathbf{K}}$. Then for all i we have*

$$|\alpha_i| = r_1^{1/d}.$$ □

As we have seen, $\operatorname{Gal}(\overline{L}/L)$ acts on $H_v^1(\psi, \mathbf{k}_v)$, and it is trivial to see that the *geometric Frobenius* acts as F. Thus Theorem 1.4.5 assures us that "$H^1(\psi)$ is pure of weight $1/d$".

THEOREM 1.4.6. *The isogeny class of ψ is uniquely determined by $Q(x)$.* □

The next result is from **[Ge3]** and is reported on in **[Go10]**.

PROPOSITION 1.4.7. *Let $\chi(L, \psi)$ be the Euler characteristic (ideal) of L as a finite $\mathbf{A}$-module via ψ. Then*

$$\chi(L, \psi) = (Q(1)).$$ □

Let $L^{(n)}$ be the unique extension of L of degree n.

DEFINITION 1.4.8. Let $i \geq 0$. We set

$$P_i(x) = \det(1 - Fx \mid H_v^i(\psi, \mathbf{k}_v))$$

for $v \neq \operatorname{char}(L)$.

Clearly

$$P_1(x) = x^d Q(1/x).$$

DEFINITION 1.4.9 (GEKELER). We set

$$Z(\psi, u) = \prod_{0 \leq t \leq d} P_t(u)^{(-1)^{t+1}}.$$

Standard computations with "$u\frac{d}{du}\log$", and Proposition 1.4.7, give

PROPOSITION 1.4.10. *With the above notation,*

$$u\frac{d}{du}\log Z(\psi, u) = \sum_{j \geq 1} \det(1 - F^j \mid H_v^1(\psi, \mathbf{k}_v))u^j,$$

and

$$(\det(1 - F^j \mid H_v^1(\psi, \mathbf{k}_v))) = (\chi(L^{(n)}, \psi)). \qquad \square$$

The reader will easily see the analogy with the classical theory of abelian varieties over finite fields. In **[Go10]**, one can find a proof, due to Gekeler, that $Z(\psi, u)$ (and thus the isogeny class of ψ) is determined by rank ψ and $\sum_{j\geq 1} \det(1 - F^j \mid H_v^1(\psi, \mathbf{k}_v))u^j$. This is not quite obvious since one *cannot* just integrate and exponentiate!

1.5. The de Rham cohomology of a Drinfeld module. In this section, we discuss the de Rham cohomology of a Drinfeld module. These groups arose in the idea of P. Deligne that the theory of "universal additive extensions", which gives a good de Rham theory for elliptic curves and abelian varieties, should *also* work for Drinfeld modules. This was then elaborated on by G. Anderson, J. Yu, and E.-U. Gekeler to give the theory presented here. Our exposition follows **[Ge2]** which, in turn, was heavily influenced by **[A2]**. (See also §2.6.)

It may help the reader to recall a bit of the classical setup. Thus, let E be an elliptic curve over $\mathbb{C}$ with ∞ being the origin of the group law. Let $E^\#$ be the generalized Jacobian with conductor 2∞; thus, if D_∞ denotes the divisors of degree 0 supported on $E - \infty$, then

$$E^\#(\mathbb{C}) \simeq D_\infty \Big/ \left\{ (f) \mid \frac{df}{f} \text{ vanishes at } \infty \right\}.$$

The space $E^\#$ fits into the exact sequence of algebraic groups over $\mathbb{C}$,

$$(\#) \qquad 0 \to V \to E^\# \to E \to 0,$$

where V is the algebraic group associated to the dual $\mathrm{Lie}(E)^\wedge$ (and is isomorphic to $\mathbb{G}_a/\mathbb{C}$). One shows that # is the *universal additive extension* of E in that any extension

$$(*) \qquad 0 \to \mathbb{G}_a \to E^* \to E \to 0$$

of algebraic groups over $\mathbb{C}$ can be deduced uniquely from #. It therefore follows that

$$V \simeq \mathrm{Ext}(E, \mathbb{G}_a)^\wedge.$$

One knows that the first de Rham cohomology group of E, $H^1_{\mathrm{DR}}(E)$, may be described as $H^0(E, \Omega(2\infty))$, i.e., as differential forms with at most a pole of order 2 at ∞. By complex integration, one deduces

$$E^\# \tilde{\to} H^0(E, \Omega(2\infty))^\wedge / H_1(E, \mathbb{Z}),$$

and, thus, an isomorphism of duals

$$\mathrm{Lie}(E^\#)^\wedge = H^0(E, \Omega(2\infty)) = H^1_{\mathrm{DR}}(E).$$

Now let $\mathrm{Ext}^\#(E, \mathbb{G}_a)$ be the group classes of sequences $(*)$ *together with* an additional splitting of Lie algebras:

$$0 \to \mathbb{C} \overset{s}{\leftrightarrows} \mathrm{Lie}(E^*) \to \mathrm{Lie}(E) \to 0.$$

The splitting s gives rise to a map $\mathrm{Ext}^\#(E, \mathbb{G}_a) \to \mathrm{Lie}(E^\#)^\wedge$, which is seen to be bijective; thus we deduce an isomorphism $\mathrm{Ext}^\#(E, \mathbb{G}_a) \tilde{\to} H^1_{\mathrm{DR}}(E)$. It is this description of $H^1_{\mathrm{DR}}(E)$ that provides the clue to our first definition of de Rham cohomology for a Drinfeld module.

Thus, let ψ be a rank d Drinfeld module over an $\mathbf{A}$-field L; we continue to assume that $\mathbf{A}$ is general. We now use the symbol "$\mathbb{G}_a$" to denote the additive group scheme over L together with its action $\iota: \mathbf{A} \to L$.

In the category of schemes with an $\mathbf{A}$-action, we are able to define extensions

$$(*) \qquad 0 \to \mathbb{G}_a \to G^* \to \psi \to 0.$$

We are thus led naturally to our next definition:

Definition 1.5.1. We define $H^1_{\mathrm{DR}}(\psi, L)$ to be the group of short exact sequences $(*)$ *together* with a splitting of the associated sequence of Lie algebras (equipped with the tautological $\mathbf{A}$-action).

We shall see that $H^1_{\mathrm{DR}}(\psi, L)$ is a finite rank L-module which can be shown to commute with base change. These results are based on the following "differential" form of $H^1_{\mathrm{DR}}(\psi, L)$.

Let $M = M(\psi, L)$ be the group $\mathrm{Hom}_{\mathbb{F}_r}(\mathbb{G}_a, \mathbb{G}_a) = L\{\tau\}$ which we will now equip with an $(L \otimes_{\mathbb{F}_r} \mathbf{A})$*-bimodule* structure as follows: The left action of L is via translations and the right action of $\mathbf{A}$ is via ψ.

Let $N = N(\psi, L) \subseteq M(\psi, L)$ be the sub-bimodule of those morphisms with trivial Lie action. Clearly

$$N = L\{\tau\}\tau .$$

DEFINITION 1.5.2.

(1) An $\mathbb{F}_r$-linear *biderivation* of $\mathbf{A}$ into $N(\psi, L)$ is an $\mathbb{F}_r$-linear map

$$\eta : \mathbf{A} \to N, \qquad a \mapsto \eta_a ,$$

such that

$$\eta_{ab} = \iota(a)\eta_b + \eta_a \circ \psi_b .$$

(2) We say that η is *inner* if there exists $m \in M(\psi, L)$ such that

$$\eta_a = \iota(a)m - m \circ \psi_a .$$

(3) We say that η is *strictly inner* or *exact* if $m \in N(\psi, L)$.

Let $D := D(\psi, L)$ be the vector space of biderivations for A to N, D_i the subspace of inner biderivations and $D_{si} \subseteq D_i$ the subspace of strictly inner derivations. In analogy with classical theory, we call the elements of D "differentials of the second kind".

EXAMPLES 1.5.3.

(1) The simplest biderivation is the map $\delta\psi$ with

$$\delta\psi_a = \psi_a - a\tau^0 ,$$

as is easy to see. Multiples of $\delta\psi$ are called "differentials of the first kind".

(2) Let ϕ_1 and ϕ_2 be two Drinfeld modules over L. Let $\rho : \phi_1 \to \phi_2$ be a morphism over L as in Definition 1.2.3 and let $\eta \in D(\phi_2, L)$. Let $\rho^*(\eta)$ be defined by

$$\rho^*(\eta)_a = \eta_a \rho$$

for $a \in \mathbf{A}$. Then, using the commutativity requirements on ρ, one readily sees that $\rho^*(\eta) \in D(\phi_1, L)$. Moreover, if η is inner (resp. strictly inner) then so is $\rho^*(\eta)$.

Let $\eta \in D(\psi, L)$ and let $a \in \mathbf{A}$. Set

$$\psi_a^\eta = \begin{pmatrix} \iota(a) & \eta_a \\ 0 & \psi_a \end{pmatrix} .$$

The fact that η is a biderivation is equivalent to

$$\psi_{ab}^\eta = \psi_a^\eta \psi_b^\eta .$$

We therefore deduce an extension

$$0 \to \mathbb{G}_a \to \mathbb{G}_a \oplus \mathbb{G}_a \to \psi \to 0$$

where the $\mathbf{A}$-action on the middle group is via ψ_a^η and the action on $\mathbb{G}_a$ is via ι. The associated action on Lie algebras is easily seen to be

$$\operatorname{Lie}(\psi_a^\eta) = \begin{pmatrix} \iota(a) & 0 \\ 0 & \iota(a) \end{pmatrix}$$

as $\eta \in N$. We therefore obtain a map $\eta \mapsto [\eta] \in H^1_{\mathrm{DR}}(\psi, L)$ as defined in 1.5.1.

One checks directly that $\eta \in D_{si}$ if and only if $[\eta]$ is trivial and that every extension in $H^1_{\mathrm{DR}}(\psi, L)$ arises from some η. Thus we have **[A2]**:

THEOREM 1.5.4 (ANDERSON). *The map $\eta \to [\eta]$ given above gives rise to an isomorphism between $H^1_{\mathrm{DR}}(\psi, L)$ and $D(\psi, L)/D_{si}(\psi, L)$.* □

One now sees clearly that $H^1_{\mathrm{DR}}(\psi, L)$ is an L-module and, with some work, that $H^1_{\mathrm{DR}}(\psi, L)$ commutes with change of base. The considerations of Example 1.5.3(2) show that H^1_{DR} is a contravariant functor, as the analogy with classical theory necessitates.

PROPOSITION 1.5.5. *Let ψ have rank d over L and let $M(\psi, L)$ and $N(\psi, L)$ be defined as above. Then $M(\psi, L)$ and $N(\psi, L)$ are projective $(L \otimes_{\mathbb{F}_r} \mathbf{A})$-modules of rank d.*

PROOF. We give the proof for M; the result for N follows similarly. Let $\alpha_T \in A$ be chosen as in §1.1 with a pole of order $t_{\alpha_T} > 0$. One sees that $\{\tau^i \mid 0 \leq i < dt_{\alpha_T}\}$ generates M *freely* as an $L[T]$-module. Thus, in particular, M is finitely generated as an $(L \otimes_{\mathbb{F}_r} \mathbf{A})$-module.

If M has torsion as an $(L \otimes_{\mathbb{F}_r} \mathbf{A})$-module then it must also have torsion as an $L[T]$-module. But a simple check shows that this is impossible; so M is a finitely generated projective $(L \otimes_{\mathbb{F}_r} \mathbf{A})$-module. Suppose $\operatorname{rank}_{L\otimes_{\mathbb{F}_r}\mathbf{A}}(M) = \delta$ and let $\{e_1, \dots, e_\delta\}$ be linearly independant elements of M. Let $\{a_1, \dots, a_t\}$, $t = t_{\alpha_T}$, be a basis for $L \otimes \mathbf{A}$ over $L[T]$. Then $\{e_i a_j\}$ freely generate a submodule of M of maximal rank over $L[T]$. From above, we see that $\delta t_{\alpha_T} = dt_{\alpha_T}$ and the result follows. □

Let $\Delta \subseteq \mathbf{A} \otimes_{\mathbb{F}_r} \mathbf{A}$ be the kernel of the map $a \otimes b \mapsto ab$. As is well known, Δ/Δ^2, together with $d: \mathbf{A} \to \Delta/\Delta^2$, $d(a) = a \otimes 1 - 1 \otimes a + \Delta^2$, is *universal* for derivations of $\mathbf{A}$ into $\mathbf{A}$-modules. In a similar fashion, Δ itself, viewed as an $(\mathbf{A} \otimes_{\mathbb{F}_r} \mathbf{A})$-*bimodule* with $d(a) = a \otimes 1 - 1 \otimes a$, is universal for *biderivations.*

Let $\Delta_L = L \otimes_{\mathbf{A}} \Delta \subseteq L \otimes_{\mathbb{F}_r} \mathbf{A}$, where Δ is viewed as a left $\mathbf{A}$-module. We see that

$$D(\psi, L) = \operatorname{Hom}_{L\otimes\mathbf{A}}(\Delta_L, N);$$

thus, by 1.5.5, $D(\psi, L)$ is also projective of rank d. In a similar vein, $D_{si}(\psi, L)$ is the image of $N(\psi, L)$ under the obvious mapping $N(\psi, L) \to \operatorname{Hom}_{L\otimes\mathbf{A}}(\Delta_L, N)$; thus, as in **[Ge2]**,

$$H^1_{\mathrm{DR}}(\psi, L) = D(\psi, L)/D_{si}(\psi, L) = D(\psi, L) \otimes_{L\otimes\mathbf{A}} L.$$

We conclude:

THEOREM 1.5.6. *$H^1_{\mathrm{DR}}(\psi, L)$ is a d-dimensional L-vector space.* □

EXAMPLE 1.5.7. Let $\mathbf{A} = \mathbb{F}_r[T]$ and let ψ be a rank d Drinfeld module. As a biderivation is determined by its action on T, it is simple to see that $H^1_{\mathrm{DR}}(\psi, L)$ has a basis $\{\delta\psi, \eta^{(1)}, \ldots, \eta^{(d-1)}\}$, where $\eta^{(i)}_T = \tau^i$. If $d = 2$, then the analogy with elliptic curves is particularly clear.

Of course, if $H^1_{\mathrm{DR}}(\psi, L)$ is a good de Rham theory, there should be a "cycle integration map" which we now discuss. Thus let L be an $\mathbf{A}$-field which is a finite extension of $\mathbf{K} \subseteq \overline{\mathbf{K}}$. By Theorem 1.3.4, there is an L-lattice M of rank d associated to ψ with

$$e_M(az) = \psi_a(e_M(z))$$

for $a \in \mathbf{A}$.

Since H^1_{DR} commutes with change of base, we see

$$H^1_{\mathrm{DR}}(\psi, \overline{\mathbf{K}}) = \bigcup_{\mathbf{K}_1} H^1_{\mathrm{DR}}(\psi, \mathbf{K}_1)$$

where $\mathbf{K}_1$ runs over all finite extensions of $L \subseteq \overline{\mathbf{K}}$.

Let $L_1 = L(M)$ be the finite extension of L containing M. Also, let $L_{\mathrm{ent}}\{\{\tau\}\}$ be the noncommutative ring of power series in $\tau(z) = z^r$ which are *entire*, i.e., have an infinite radius of convergence.

The idea behind cycle integration now lies in the next lemma. For a proof, see **[Ge1]**.

LEMMA 1.5.8. *Let $\eta \in D(\psi, L)$. Then there exists a unique power series $F_\eta \in L_{\mathrm{ent}}\{\{\tau\}\}\tau$ such that*

$$aF_\eta(z) - F_\eta(az) = \eta_a(e_M(z))$$

for all $a \in \mathbf{A}$. □

The reader should note the change of sign in our choice of F_η from that of **[Ge1]** and **[Ge2]**.

DEFINITION 1.5.9. Let η, M be as above. Let $\gamma \in M$. Then we set

$$\int_\gamma \eta = F_\eta(\gamma).$$

We call $\{F_\eta(\gamma)\}$ the *quasi-periods* of ψ. They are of the "first" and "second" kind depending on η.

The important point to note is that, by 1.5.8, $\gamma \mapsto \int_\gamma \eta$ is an $\mathbf{A}$-*linear* map from M to L_1.

EXAMPLE 1.5.10.

(1) Let $\eta = \delta\psi$. Then it is trivial to see that

$$F_\eta(z) = z - e_M(z).$$

Consequently

$$\int_\gamma \delta\psi = \gamma\,,$$

and the periods of M are contained in the "quasi-periods".

(2) Let η be exact; thus there exists $n \in N(\psi, L)$ with

$$\eta_a = an - n\psi_a\,.$$

We conclude that

$$F_\eta(z) = n(e_M(z)) \Rightarrow F_\eta(\gamma) = 0\,.$$

The converse to 1.5.10(2) is also true. It gives the *de Rham isomorphism* in this context:

THEOREM 1.5.11 [A2]. *Cycle integration gives rise to an isomorphism between*

$$H^1_B(\psi, \overline{\mathbf{K}}) = H^1_B(\psi, \mathbf{k}) \otimes \overline{\mathbf{K}} = \mathrm{Hom}_{\mathbf{A}}(M, \overline{\mathbf{K}})$$

and $H^1_{\mathrm{DR}}(\psi, \overline{\mathbf{K}})$. □

LEMMA 1.5.12 [A2]. *Let* $\{\eta, \gamma\}$ *be as above. Then we have*

(1)

$$\int_\gamma \eta = -\lim_{\deg(a)\to\infty} \eta_a\left(e_M\left(\frac{\gamma}{a}\right)\right)\,,$$

(2)

$$\int_\gamma \eta = -\sum_{n=0}^{\infty} a^n \eta_a\left(e_M\left(\frac{\gamma}{a^{n+1}}\right)\right)\,, \quad \textit{for } \deg(a) > 0\,.$$

□

PROOF.

(1) We have

$$\eta_a e_M(z) = aF_\eta(z) - F_\eta(az)\,.$$

Thus,

$$\eta_a e_M(\gamma/a) = aF_\eta(\gamma/a) - F_\eta(\gamma)\,.$$

Now expand F_η as a power series in z and take the limit.

(2) This follows upon substituting

$$\eta_a(e_M(z)) = aF_\eta(z) - F_\eta(az)$$

into the formula. □

REMARK 1.5.13. Classically one has

$$e^z = \lim_{n\to\infty}(1 + z/n)^n\,,$$

as is shown by taking the logarithm. In a similar fashion, using the logarithm of ψ (i.e., the inverse of $e_M(z)$) one can also show

$$e_M(z) = \lim_{\deg(a)\to\infty} \psi_a(z/a)\,.$$

In the language of the next section, the matrix

$$\Psi_a^\eta = \begin{pmatrix} a & \eta_a \\ 0 & \psi_a \end{pmatrix}, \qquad \eta \in D_\eta(\psi, E),$$

gives rise to a two-dimensional "A-module"; i.e., an algebraic action of **A** on $\mathbb{G}_a \times \mathbb{G}_a$. One then sees that the two-dimensional "exponential" of this object is $\lim_{\deg(a)\to\infty} \Psi_a^\eta(z/a)$.

2. T-modules and T-motives

We now pass to a more general category which contains finite products of Drinfeld modules *and* which also possesses, under certain circumstances, a tensor product. The clue to this construction can be seen in the de Rham theory (§1.5) where we studied $M(\psi, L)$ for a Drinfeld module ψ over an **A**-field L. By 1.5.5, we know that $M(\psi, L)$ is a projective $(L\otimes_{\mathbb{F}_r}\mathbf{A})$-module of the same rank as ψ.

The map $\psi \mapsto M(\psi, L)$ is a contravariant functor. Moreover, the reader may readily see that ψ can be reconstructed out of $M(\psi, L)$ **[A1]** (see also, Remark 2.7 of **[DH1]**). On the other hand, it is clear how to tensor, over $L\otimes_{\mathbb{F}_r}\mathbf{A}$, $M(\psi_1, L)$, and $M(\psi_2, L)$ for two Drinfeld modules ψ_1, ψ_2 — this is the advantage of the dual (i.e., $M(\psi, L)$) viewpoint. When one looks for algebraic objects associated to such tensor products, one then obtains the "T-modules"; higher-dimensional generalizations of Drinfeld modules. The module $M(\psi, L)$ is viewed as the "motive" associated to ψ.

There is another, related, construction associated to ψ using this dual viewpoint. This is Drinfeld's "shtuka" **[Dr3, Dr4, Mu1]**. These "shtuka" are vector bundles on the complete curve $\mathscr{C}$ associated to **k** with certain extra properties. Roughly the idea is that one uses the degree gradation on $M(\psi, L)$ and the dictionary of FAC to obtain the associated bundle on $\mathscr{C}$. Although we will have no more to say about shtuka here, they are also beginning to have an impact on the types of objects talked about in §3 **[Th5]**. Finally, there is yet a third object now being developed which is related to the T-motives and shtuka, the "solitons". We refer the reader to **[A4]**.

Up to now, the main application of the theory of T-modules and T-motives has been to provide important examples, such as the nth tensor power $C^{\otimes n}$ of the Carlitz module **[AT1]**; see Example 2.4.3. These examples then are used to obtain information about special zeta-values **[AT1, Y3, Y5]**; see Theorem 3.19. Moreover, much of the cohomology theory of Drinfeld modules, while certainly expected to generalize, has not yet been fully carried over to T-modules. Thus, our purpose here will be to give an introduction to the theory with the hope of interesting the reader in further carrying out the constructions.

We work in the very basic case of $\mathbf{A} = \mathbb{F}_p[T]$ (see 2.1.7(3)). As with the theory of Drinfeld modules, the general case will appear as complex multiplications (§4). Also, as we are letting $r = p$, unless otherwise specified, the symbol "τ" will now denote the pth power mapping τ_p.

2.1. Basic definitions. Let L be an **A**-field equipped with structure map $\iota: \mathbf{A} \to L$. In order to be able to perform various algorithms, we also assume, unless otherwise stated, that L is *perfect.*

Our first definition is a ring of operators which incorporates both the left $(L \otimes_{\mathbb{F}_p} \mathbf{A})$-action and the τ-action on $M(\psi, L)$.

DEFINITION 2.1.1. Let $L[T, \tau]$ be the (noncommutative, in general) ring freely generated by T, τ with the commutativity requirements: $\tau T = T\tau$, $T\alpha = \alpha T$, $\tau\alpha = \alpha^p\tau$ for $\alpha \in L$.

When discussing $L[T, \tau]$ we will often use the notation $L[\tau]$ for the subring of $L[T, \tau]$ generated by L and τ. Previously, we have used the notation " $L\{\tau\}$ ".

DEFINITION 2.1.2. Let $\theta = \iota(T) \in L$. A *T-motive, M*, is a left $L[T, \tau]$-module which is free and finitely generated as an $L[\tau]$-module and such that

$$(T - \theta)^n M/\tau M = 0$$

for $n \gg 0$. A *morphism* of two such T-motives is an $L[T, \tau]$-linear homomorphism of left $L[T, \tau]$-modules.

For example, if ψ is a Drinfeld module over L, then we know that $M(\psi, L) = L\{\tau\}$ as an $L\{\tau\}$-module. Moreover, the action of T is via ψ_T as before; thus $T = \theta$ on $M/\tau M$. Therefore, $M(\psi, L)$ is a T-motive. In fact this example is typical:

DEFINITION 2.1.3. A *T-module E* is an algebraic group over L which is isomorphic to $\mathbb{G}_a^e$ equipped with an endomorphism over L, $E \xrightarrow{T} E$, such that

$$(T - \theta)^n \operatorname{Lie}(E) = 0$$

for $n \gg 0$ and θ as above. A *morphism* of T-modules is a T-equivariant homomorphism of algebraic groups over L. The *dimension* of the T-module E is e.

The nilpotency condition ($(T-\theta)^n = 0$) is needed in order to have tensor products. Should we need to specify the T-module, we shall sometimes write " α_E " for the action of $\alpha \in \mathbf{A}$. The reader should note that a Drinfeld module ψ for $\mathbb{F}_p[T]$ is a T-module in the obvious fashion. Note also that we are now using the notation " Tx " for " $\psi_T(x)$ ", etc.

The above definitions are given following **[A2]**. They differ from those presented originally in **[A1]**. This difference is discussed in Remark 2.1.8.

The categories of T-modules and T-motives are mirror images of each other as follows: Let $M(E)$ be the set of $\mathbb{F}_p$-linear morphisms of E to $\mathbb{G}_a$ as algebraic groups over L (as we did for Drinfeld modules, §1.5). We make $M(E)$ into a T-motive by

$$(\alpha, m) \mapsto (\alpha m(x): E \to \mathbb{G}_a)\,; \qquad m \in M\,,\ \alpha \in L\,;$$
$$(\tau, m) \mapsto (m(x)^p: E \to \mathbb{G}_a)\,; \qquad m \in M\,;$$
$$(T, m) \mapsto (m(Tx): E \to \mathbb{G}_a)\,, \qquad m \in M\,.$$

Reasoning as in [**A1**], we have

THEOREM 2.1.4.

(1) *The map* $E \mapsto M(E)$ *is an anti-equivalence of categories.*
(2) $\mathrm{Hom}_L(\mathrm{Lie}(E), L) \tilde{\to} M(E)/\tau M(E)$. □

PROPOSITION 2.1.5. $M(E)$ *is free of finite rank over* $L[T]$ *if and only if there is a finite-dimensional* L*-subspace,* V, *of* $\mathrm{Hom}_{\mathbb{F}_p}(E, L)$ *such that*

$$\mathrm{Hom}_{\mathbb{F}_p}(E, \mathbb{G}_a) = \sum_{j=0}^{\infty} V \circ T^j .$$

PROOF. Since L is perfect, one has left and right division algorithms for $L[\tau]$, as well as elementary row operations, etc., for matrices. One is then able to show the following lemma:

LEMMA 2.1.6 [**A1**]. *Let* M *be a left* $L[T, \tau]$*-module which is finitely generated as an* $L[T]$*-module and as an* $L[\tau]$*-module. Then* M *is free over* $L[T]$ *of finite rank if and only if it is free over* $L[\tau]$ *of finite rank.* □

The lemma now implies Proposition 2.1.5. □

EXAMPLES 2.1.7.

(1) The most general example of a T-module is given as follows: Put

$$E := \mathbb{G}_a^e = \begin{pmatrix} \mathbb{G}_a \\ \vdots \\ \mathbb{G}_a \end{pmatrix} ,$$

and let N be a fixed nilpotent matrix. Let

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_e \end{pmatrix} \in E.$$

Set

$$\tau^i(x) = \begin{pmatrix} x_1^{p^i} \\ \vdots \\ x_e^{p^i} \end{pmatrix}$$

and

$$T \mapsto (Tx = (\theta I + N)(x) + g_1 \tau(x) + \cdots + g_n \tau^n(x))$$

where $\{g_1, \dots, g_n\} \subseteq M_e(L)$. We say that E is an *abelian* T*-module* (and $M(E)$ is an *abelian* T*-motive*) if and only if $M(E)$ is finitely generated over $L[T]$. We extend this definition to left $L[T, \tau]$-modules in the obvious fashion. By Lemma 2.1.6, $M(E)$ is then free over $L[T]$. We set $r(E) := \mathrm{rank}(E) := \mathrm{rank}_{L[T]} M(E)$.

(2) In particular, "$\mathbb{G}_a$", with the usual action, is a T-module. Thus $T \mapsto (Tx = \theta x)$. We denote this module by "E_{linear}". The associated motive is $L\{\tau\}$ with $(T, m) \mapsto m(\theta x)$. This is obviously nonabelian.

(3) It is easy to see that a T-module of dimension one *with nonscalar action* is the same thing as a Drinfeld module for $\mathbb{F}_p[T]$ (the nilpotent matrix N must vanish in this instance). Let ψ be a Drinfeld module for $\mathbb{F}_p[T]$ which we suppose has complex multiplication by $\mathbb{F}_r$ extending the scalar $\mathbb{F}_p$-action. It is elementary to show that this $\mathbb{F}_r$-action must also be via scalars. Thus we recover the notion of Drinfeld modules over $\mathbb{F}_r[T]$ for all r.

(4) (This example is not needed for the sequel and should be skipped on a first reading by those who are new to the subject.) The observations of part (3) are unique to dimension one as the following example, in the notation of part (1) above, shows: Let $e = 2$ and let E be the T-module over $L = \mathbb{F}_{p^2}(T)$ given by $n = 2$, $N = g_1 = 0$, $g_2 = I$. We give E complex multiplication by $\mathbb{F}_{p^2}[T]$ by having $\mathbb{F}_{p^2}$ act as scalars. Now let $P_i = I + (-1)^i V_2 \tau$ where $V_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. It is easy to see that $I = P_0 P_1$. Then the T-module E_1 given by conjugating E with P_0, P_1 has nonscalar action by $\mathbb{F}_{p^2} - \mathbb{F}_p$. (I am indebted to the referee for this particularly simple example.) Still, given complex multiplication by $\mathbb{F}_r$ one can ask whether there exists a good system of coordinates [**A1**, p. 472] giving scalar action. (See [**Dr1**, Proposition 5.2] and [**DH1**, Chapter 1, §5] where a similar type of result is established for Drinfeld modules over a ring with nilpotents.)

REMARK 2.1.8. The "T-motives" of [**A1**] correspond to our "abelian T-motives", etc. Thus, 2.1.7(2) is *not* an example of a T-motive in the notation of [**A1**].

2.2. Purity. In this subsection we discuss the notion of "purity" [**A1**, §1.9]. It is used to force eigenvalues of Frobenius elements to be "pure".

Thus, let M be a left $L[T, \tau]$-module. Put

$$M\left(\left(\frac{1}{T}\right)\right) := M \otimes_{L[T]} L\left(\left(\frac{1}{T}\right)\right).$$

We equip $M\left(\left(\frac{1}{T}\right)\right)$ with the left $L[T, \tau]$-module structure given by

$$\tau\left(m \otimes \sum_{j \gg -\infty} a_j T^{-j}\right) := \tau m \otimes \sum_{j \gg -\infty} a_j^p T^{-j}.$$

A "lattice" in $M((\frac{1}{T}))$ is a free $L\left[\left[\frac{1}{T}\right]\right]$-submodule which gives $M((\frac{1}{T}))$ upon tensoring with $L\left(\left(\frac{1}{T}\right)\right)$.

DEFINITION 2.2.1.

(1) The module M is said to be *pure* provided both of the following hold. (1) M is free and finitely generated over $L[T]$; and, (2) there exists an $L[[\frac{1}{T}]]$-lattice $W \subseteq M((\frac{1}{T}))$, and positive integers q and s, such that
$$\tau^s W = T^q W.$$

(2) The T-module E is *pure* if and only if the associated motive $M(E)$ is also.

In particular, a pure T-motive is abelian, etc.

EXAMPLE 2.2.2. Let ψ be a Drinfeld module over L of rank d. Then $M(\psi, L)$ is free over $L[\tau]$ on $m_0 := \mathrm{id}$. We saw in Proposition 1.5.5 that $M(\psi, L)$ is free of rank d over $L[T]$. It is easy to see that a basis is given by $\{\tau^i m_0 \mid 0 \le i \le d-1\}$. Moreover if
$$\psi_T = \theta\tau^0 + \sum_{j=1}^{d} g_j \tau^j ,$$
then
$$Tm_0 = \psi_T(m_0).$$
We can also see that $M(\psi, L)$ is pure in the sense of Definition 2.2.1 as follows: Set
$$W = \bigoplus_{j=0}^{d-1} \tau^j m_0 \otimes L\left[\left[\frac{1}{T}\right]\right].$$

CLAIM. $TW = \tau^d W$.

PROOF. Note that $\{\tau^i m_0 \mid 0 \le i \le d-1\} \subseteq W \subseteq TW$. Now, as L is perfect, we have
$$\begin{aligned} TW &= T\left(\sum \tau^j m_0 \otimes L\left[\left[\frac{1}{T}\right]\right]\right) \\ &= T\left(\sum \tau^j m_0 \otimes \left(1-\frac{\theta}{T}\right) L\left[\left[\frac{1}{T}\right]\right]\right) \\ &= (T-\theta)\left(\sum \tau^j m_0 \otimes L\left[\left[\frac{1}{T}\right]\right]\right) \\ &= \sum_{i=1}^{d} g_i \tau^i \left(\sum \tau^j m_0 \otimes L\left[\left[\frac{1}{T}\right]\right]\right). \end{aligned}$$
One then sees that $\tau^d W \subseteq TW$. The other direction follows in a similar fashion. □

REMARK 2.2.3. A sufficient condition for purity of a T-module is that
$$T^q x = (\theta^q + N_1)x + \cdots + g_s \tau^s(x), \quad s > 0,$$
where N_1 is nilpotent and $\det(g_s) \neq 0$.

PROPOSITION 2.2.4 [**A1**, 1.10.1]. *Let M be a pure left $L[T, \tau]$-module. Then M is free and finitely generated over $L[\tau]$. Moreover, let $M \neq \{0\}$ and q and s as in 2.2.1. Then*

$$q/s = \frac{\operatorname{rank}_{L[\tau]} M}{\operatorname{rank}_{L[T]} M}.$$

PROOF. Let f be a positive integer chosen large enough so that if $W \subseteq M\left(\left(\frac{1}{T}\right)\right)$ is the lattice of 2.2.1, then

$$M + T^f W = M\left(\left(\frac{1}{T}\right)\right).$$

Set $M_j = M \cap T^{(j+f)q} W$ for $j \geq 0$. We have isomorphisms

$$\begin{aligned} M_{j+1}/M_j &\simeq T^{(j+f+1)q} W / T^{(j+f)q} W \\ &= \tau^{(j+f+1)s} W / \tau^{(j+f)s} W. \end{aligned}$$

Thus $T^q M_j + M_j = M_{j+1} = \tau^s M_j + M_j$ for all $j \geq 0$. Thus M is finitely generated over $L[\tau]$. Now use Lemma 2.1.6.

The second part follows in a similar fashion. □

Thus purity forces an $L[T, \tau]$-module to be finitely generated as an $L[T]$-module *and* $L[\tau]$-module. From 2.1.6, 2.1.4(2), and the definitions we find

COROLLARY 2.2.5. *If M is a pure left $L[T, \tau]$-module and satisfies*

$$(T - \theta)^j M/\tau M = 0$$

for $j \gg 0$, then M is an abelian T-motive. □

DEFINITION 2.2.6. Let M be a pure T-motive. Then we set

$$\text{rank of } M := r(M) := \operatorname{rank}_{L[T]} M$$

and

$$\text{weight of } M := w(M) := \operatorname{rank}_{L[\tau]} M / \operatorname{rank}_{L[T]} M$$

for nontrivial M. These definitions are transformed to T-modules in the obvious fashion.

Clearly, $\dim E = \operatorname{rank}_{L[\tau]} M(E) = w(E) r(E)$ for a T-module E.

Suppose now that L is the quotient field of a local ring $\mathscr{O}$ with maximal ideal $\wp$. Note that $\mathscr{O}/\wp$ is also perfect. Using the theory of Drinfeld modules as a guide, the reader will readily work out the details of the reduction of a T-module E (*or* motive $M(E)$) at $\wp$. We say that the reduction is *good* if the rank does not change.

IDEA 2.2.7. *Let E have good reduction at $\wp$. Let $\widetilde{E}$ be the reduced motive. Generalize the theory of Drinfeld modules and show that if E is pure of weight w then so is $\widetilde{E}$.*

Let L now be a finite field. We know that L is a finite extension of $\mathbf{A}/\wp$ for some $\wp \in \operatorname{Spec}(\mathbf{A})$. Let L have p_1-elements where $p_1 = p^t$, $t = \deg(\wp)[L : \mathbf{A}/\wp]$.

Let E be a T-module over L which we assume is *pure*. Let $F = (x \mapsto \tau^t(x))$ be its *Frobenius endomorphism*. We are interested in the characteristic polynomial $Q(x)$ of F acting on the v-adic Tate module for $v \neq \wp \in \text{Spec}(\mathbf{A})$. We will give the construction in our next section. For the moment, however, we ask the reader to make the reasonable assumption that such a space exists.

Let $M = M(E)$ be the motive for E. Let $\overline{L}$ be a fixed algebraic closure for L. Let

$$\overline{M} := \overline{M}(E) := M(E) \otimes_L \overline{L}\,,$$

which we view as an $\overline{L}[T\,,\tau]$ with diagonal τ action; i.e., $\tau(x \otimes y) = \tau(x) \otimes y^p$. Thus $\overline{M}(E)$ is just the original motive viewed over $\overline{L}$. (Equivalently, we could view E as a T-module over $\overline{L}$ and take its motive.) In the next section, we will also see that the v-adic Tate module is "essentially" the dual of $\varprojlim (\overline{M}/v^n\overline{M})^\tau$; i.e., the inverse limit over the the fixed spaces of τ (see Definition 2.3.2). The reader should note that this construction makes sense as v acts *centrally*. The process of taking τ-invariants is a general principle in forming cohomology of T-motives.

Lang's theorem on p-linear automorphisms of finite-dimensional vector spaces assures us that

$$(*) \qquad (\overline{M}/v^n\overline{M})^\tau \otimes_{\mathbb{F}_p} \overline{L} = \overline{M}/v^n\overline{M}$$

as $\overline{L}[T\,,\tau]$-modules with the diagonal τ-action on the tensor product.

Now let $f \in \overline{M}(E)$. By definition, the action of τ on f is $\tau(f) = f^p$. The action of F on f is $f \circ F$. Thus F and τ commute; therefore F gives rise to an endomorphism of $\overline{M}/v^n\overline{M}$ *and* its τ-invariants. So to compute the characteristic polynomial of F on the v-adic Tate module, we need only compute it on $\varprojlim \overline{M}/v^n\overline{M}$ as a module over $\varprojlim \overline{L}[T]/v^n\overline{L}[T]$.

However, F is actually an endomorphism of $\overline{M}$ *itself* as an $\overline{L}[T]$-module. We thus deduce that $Q(x)$ has coefficients in $\overline{L}[T]$. But, from $(*)$ we see that the coefficients must also be invariant of $\text{Gal}(\overline{L}/\mathbb{F}_p)$. Thus, we have shown

PROPOSITION 2.2.8. *$Q(x)$ has coefficients in $\mathbf{A}$ which are invariant of the choice of v.* □

In fact, the reader will see that the above arguements work for any endomorphism of E defined over L.

The proof of the following result will be left to the reader.

LEMMA 2.2.9. *Let K be the field of fractions of a complete discrete valuation ring R with uniformizer π. Let V/K be a finite-dimensional vector space and let $f: V \to V$ be a K-linear endomorphism. The eigenvalues of f are all of absolute value $|\pi|^{q/s}$, for fixed $q\,,\ s \in N$, if and only if there exists an R-lattice $W \subseteq V$ such that*

$$f^s(W) = \pi^q W\,.$$ □

It now remains to compute the absolute values of F in $\overline{\mathbf{K}}$. Arguing as above, we may now work with the module $M\left(\left(\frac{1}{T}\right)\right)$. The key here is to note that on M one has the equality $F = \tau^t$. Thus, using the above lemma and the definition of purity, we have

THEOREM 2.2.10. *The eigenvalues of F are pure of weight $w(M)$. That is, if α is an eigenvalue $\in \overline{\mathbf{k}} \subset \overline{\mathbf{K}}$, then*

$$|\alpha| = p_1^{w(M)}. \qquad \square$$

QUESTION 2.2.11. Does $Q(x)$ determine the isogeny class of E?

2.3. Torsion points. Let L be an A-field which we now assume is *algebraically closed* and thus automatically perfect. Let E be a pure T-module defined over L. Let $f \in \mathbf{A}$ be prime to the characteristic of L. We note, again, that f is central in $L[T, \tau]$. In this section we discuss the torsion points of E from the viewpoint of the associated motive $M := M(E)$ following **[A1]** and **[A2]**.

DEFINITION 2.3.1. Set

$$E[f] = \{e \in E(L) \mid f \cdot e = 0\}.$$

We call $E[f]$ the module of f-torsion points.

Clearly $E[f]$ is an A-module in the obvious fashion.

DEFINITION 2.3.2. Put

$$(M/fM)^\tau = \{m \in M/fM \mid \tau m = m\}.$$

Since f is central, $(M/fM)^\tau$ is an $L[T, \tau]$-module. (The module $(M/fM)^\tau$ was used in our previous section.)

PROPOSITION 2.3.3. *We have, as $\mathbb{F}_p$-vector spaces,*

$$E[f] \xrightarrow{\sim} \mathrm{Hom}((M/fM)^\tau, \mathbb{F}_p).$$

PROOF. Let $\tilde{L}$ be L viewed as an $L[\tau]$-module by evaluation: $(g(\tau), x) \mapsto g(x)$. Then one checks that $E[L] \xrightarrow{\sim} \mathrm{Hom}_{L[\tau]}(M, \tilde{L})$.

Since f is prime to the characteristic of L, M/fM is finite dimensional over L. The map τ on M/fM is a p-linear endomorphism. Thus Lang's theorem (applied to GL_m, $m = \deg(f) \cdot r(E)$) assures us that

$$(M/fM)^\tau \otimes_{\mathbb{F}_p} L \simeq M/fM.$$

Consequently,

$$\begin{aligned} E[f] &= \mathrm{Hom}((M/fM)^\tau, L^\tau) \\ &= \mathrm{Hom}((M/fM)^\tau, \mathbb{F}_p). \end{aligned} \qquad \square$$

COROLLARY 2.3.4 **[A1, 1.8.3]**. *$E[f]$ is a free $\mathbb{F}_p[T]/(f)$-module of rank $r(E)$.*

Proof. We know that M/fM is free of rank $r(E)$ over $A/(f)$. Moreover, as $L[T]$-modules,

$$\mathrm{Hom}_{L[\tau]}(M/fM, \widetilde{L}) \simeq (L[T]/(f))^{r(E)}.$$

The result now follows by duality and the previous result. □

Our next goal is to "refine" the above results to a duality of $\mathbb{F}_p[T]$-modules. Let $\Omega = \mathbb{F}_p[T]dT$ be the module of Kähler differentials of $\mathbb{F}_p[T]$. In the usual fashion, we have the residue map

$$\mathrm{Res}_\infty : \mathbb{F}_p(T) \otimes_{\mathbb{F}_p(T)} \Omega \to \mathbb{F}_p$$

given by taking the residue at ∞.

We equip $V := \mathrm{Hom}_{\mathbb{F}_p}(\mathbb{F}_p[T]/(f), L)$ with the structure of left $L[T, \tau]$-module in the following (now clear) fashion:

$$\begin{aligned} &(x, h) \mapsto xh(a), \qquad \text{for } x \in L, \ h \in V, \\ &(T, h) \mapsto h(Ta); \\ &(\tau, h) \mapsto h(a)^p. \end{aligned}$$

With this definition, there is a map

$$E(f) \to \mathrm{Hom}_{L[T,\tau]}(M/fM, \mathrm{Hom}_{\mathbb{F}_p}(\mathbb{F}_p[T]/(f), L))$$

as follows: Let $e \in E[f]$. Then $e \mapsto (m \mapsto (a \mapsto m(ae)))$. One checks that this is an isomorphism. We can now establish

Theorem 2.3.5. *As an $\mathbb{F}_p[T]$-module, $E[f]$ is canonically isomorphic to*

$$\mathrm{Hom}_{\mathbb{F}_p[T]}\left((M/fM)^\tau, f^{-1}\Omega/\Omega\right).$$

Proof. We have, as before,

$$\begin{aligned} &\mathrm{Hom}_{L[T,\tau]}\left(M/fM, \mathrm{Hom}_{\mathbb{F}_p}(\mathbb{F}_p[T]/(f), L)\right) \\ &\quad = \mathrm{Hom}_{\mathbb{F}_p[T]}\left((M/fM)^\tau, \mathrm{Hom}_{\mathbb{F}_p}(\mathbb{F}_p[T]/(f), \mathbb{F}_p)\right). \end{aligned}$$

The result follows by noting that $\mathbb{F}_p[T]/(f)$ and $f^{-1}\Omega/\Omega$ are perfectly paired by $(a, \omega) \mapsto \mathrm{Res}_\infty(a\omega)$. □

Composing Theorem 2.3.5 with Res_∞ gives back Corollary 2.3.4.

Suppose now that $f = v$ is a prime. We can form the v-adic Tate module, $T_v(E)$, of E in the standard fashion. The above results give

$$T_v(E) = \mathbf{A}_v^{r(E)};$$

which is in agreement with what was established earlier for Drinfeld modules.

Put $\Omega_v = \mathbf{A}_v \otimes \Omega$. Set

$$H_v^1(M) = \varprojlim (M/v^n M)^\tau.$$

Then Theorem 2.3.5 gives the duality

$$\mathrm{Hom}_{\mathbf{A}_v}(H_v^1(M), \Omega_v) = T_v(E);$$

which was *precisely* what was used at the end of §2.2.

2.4. Tensor products. Let M_1, M_2 be two left $L[T, \tau]$-modules.

DEFINITION 2.4.1 [**A1**]. We define $M_1 \otimes M_2$ to be $M_1 \otimes_{L[T]} M_2$ equipped with the diagonal τ-action. This is carried over to the T-modules in the obvious fashion.

PROPOSITION 2.4.2 [**A1**, 1.11.1]. *Let M_1, M_2 be two pure T-motives. Then $M_1 \otimes M_2$ is also a pure T-motive.*

PROOF. One checks that $M_1 \otimes M_2$ is again a pure module. It is a T-motive by Corollary 2.2.5. □

Thus, for example, the tensor product of two Drinfeld modules is a pure T-module (see also [**H1**]). By definition, $r(M_1)r(M_2) = r(M_1 \otimes M_2)$ and $\omega(M_1) + \omega(M_2) = \omega(M_1 \otimes M_2)$.

EXAMPLE 2.4.3. Let C be the Carlitz-module over $\mathbf{k}$ (so $C_T = T\tau^{\circ} + \tau$). Then $C^{\otimes n}$ is the abelian T-module with underlying group $\mathbb{G}_a^n$ and the following action of T: Let N_n be the $n \times n$ matrix

$$\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

with 1's above the diagonal and let V_n be the $n \times n$ matrix

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 1 & \cdots & 0 \end{pmatrix}.$$

Then

$$Tx = (TI_n + N_n)x + V_n\tau(x).$$

Although we are now working over $\mathbb{F}_p[T]$, in fact the above definition gives the nth tensor power of the Carlitz module also for $\mathbb{F}_r[T]$ with $\tau = \tau_r$. Indeed, in this case we are merely doing the tensor product operations over $\mathbb{F}_r$.

The weight of $C^{\otimes n}$ is n and its rank is one; it is pure and gives an instance of Remark 2.2.3 (with $q = n$) as is directly seen. For more on the structure of $C^{\otimes n}$, we refer the reader to [**AT1**].

One has the important fact that

$$H_v^1(M_1 \otimes M_2) = H_v^1(M_1) \otimes_{\mathbf{A}_v} H_v^1(M_2).$$

Note also that $H_v^1(M)$ inherits a Galois module structure from $T_v(E)$. One sees that the above equality is actually an equality of Galois modules. In particular, the Galois (abelian) character associated to $C^{\otimes n}$ is just the nth power of the character associated to C (see Example 1.3.9). This is also established directly by Anderson-Thakur [**AT1**, 1.11.1].

DEFINITION 2.4.4. We set

$$H_v^1(E, \mathbf{k}_v) = \mathrm{Hom}_{\mathbf{A}_v}(T_v(E), \mathbf{k}_v).$$

Higher groups are defined via exterior powers.

Thus $H_v^1(E, \mathbf{k}_v) \otimes_{\mathbf{k}_v} \Omega_v \simeq H_v^1(M)$ commutes with tensor product.

2.5. Uniformizability and Betti cohomology. Let $L \subseteq \overline{\mathbf{K}}$ now be a finite extension of $\mathbf{K}$. Let E be an abelian T-module over L which we view as being defined over $\overline{\mathbf{K}}$. In a similar fashion to what is done with the Drinfeld modules (one-dimensional T-modules), one can associate to E an exponential function $e(x) := e_E(x)$. We will present Anderson's criterion for *surjectivity* of this exponential. This is somewhat analogous to the classical criterion of Riemann forms in the theory of complex abelian varieties. The reader should note, however, that there is a different flavor here than for abelian varieties. Indeed, there exist T-modules which *cannot* be uniformized. (Of course, every abelian variety over $\mathbb{C}$ can be uniformized; however, there are complex tori that are not algebraic.)

Let E be given as in Example 2.1.7, and let t be the dimension of E. One then finds **[A1**, §2.1] matrices $\alpha_j \in M_t(\overline{\mathbf{K}})$ such that if

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_t \end{pmatrix} \quad \text{and} \quad e(x) = \sum \alpha_j \tau^j(x),$$

then

$$e((T+N)x) = T_E e(x);$$

here, as before, T_E is multiplication by T in E. One further finds that $e(x)$ is *entire* (in the obvious sense), $\mathbb{F}_p$*-linear*, and *unique*. This is in complete analogy with Drinfeld modules. But, as mentioned above, one can find examples where e is *not surjective* (see **[A1**, §2.2] for a simple example).

Let $M = M(E)$.

DEFINITION 2.5.1.

(1) Let $\overline{\mathbf{K}}\{T\}$ be the ring of power series in T over $\overline{\mathbf{K}}$ which converge on the open unit disc and whose coefficients generate a finite extension of $\mathbf{K}$. (That is, $f(T) = \sum_{j=0}^{\infty} a_j T^j \in \overline{\mathbf{K}}\{T\}$ if and only if $a_j \to 0$ as $j \to \infty$ and $[\mathbf{K}(\{a_j\}) : \mathbf{K}] < \infty$.)

(2) We put $M\{T\} = M \otimes_{\overline{\mathbf{K}}[T]} \overline{\mathbf{K}}\{T\}$.

We view $M\{T\}$ as a $\overline{\mathbf{K}}[T, \tau]$-module with, as usual,

$$\tau\left(m \otimes \sum a_j T^j\right) = \tau(m) \otimes \sum a_j^p T^j.$$

DEFINITION 2.5.2.

(1) We set

$$H^1(M) := M\{T\}^\tau = \text{fixed space of } \tau.$$

(2) We set $H_1(E) =$ kernel of $e(x)$.

Clearly, $H_1(E)$ is an $\mathbf{A}$-module by the functional equation of $e(x)$.

For the proof of the next result, we refer the reader to [**A1**, Theorem 4].

THEOREM 2.5.3 (ANDERSON). *The following properties of E are equivalent:*

(1) $\operatorname{rank}_{\mathbf{A}}(H_1(E)) = r(E)$.
(2) $e(X)$ *is surjective.*
(3) *The map* $H^1(M) \otimes_{\mathbf{A}} \overline{\mathbf{K}}\{T\} \to M\{T\}$ *is an isomorphism (that is, M is "rigid-analytically trivial").* □

DEFINITION 2.5.4. If any of the equivalent conditions of Theorem 2.5.3 hold, then we say that E is *uniformizable.*

COROLLARY 2.5.5. *Let E_1 and E_2 be two pure T-modules. Let $E_1 \otimes E_2$ be the T-module corresponding to $M(E_1) \otimes M(E_2)$. Then if E_1 and E_2 are uniformizable, so is $E_1 \otimes E_2$. Moreover,* $H^1(M(E_1) \otimes M(E_2)) = H^1(M(E_1)) \otimes_{\mathbf{A}} H^1(M(E_2))$.

PROOF. Use property 3 of Theorem 2.5.3 and basic properties of tensor products. □

PROPOSITION 2.5.6 [**A1**, COROLLARY 2.12.1]. *Assume that E is uniformizable. Then $H_1(E)$ is isomorphic to* $\operatorname{Hom}_{\mathbf{A}}(H^1(M(E)), \Omega)$. □

The reader will note the exact analogy with what was established before for T_v, etc.

DEFINITION 2.5.7. Let E be uniformizable. Then we set

$$H_B^1(E, \mathbf{k}) = \operatorname{Hom}_{\mathbf{A}}(H_1(E), \mathbf{k}).$$

We call $H_B^1(E, \mathbf{k})$ the "Betti cohomology of E". Higher cohomology groups are defined via exterior products as before.

We see that $H_B^1(E, \Omega) := H_B^1(E, \mathbf{k}) \otimes \Omega$ commutes with tensor products. Moreover, the analytic theory gives an isomorphism of vector spaces

$$H_B^1(E, \mathbf{k}) \otimes_{\mathbf{k}} \mathbf{k}_v \simeq H_v^1(E, \mathbf{k}_v),$$

etc. Finally, Theorem 2.5.3 gives

PROPOSITION 2.5.8 [**A1**, 2.12.2]. *The functor $E \mapsto H_1(E)$ on uniformizable abelian T-modules is faithful.* □

2.6. More on cohomology and the Legendre formula. Let L be a perfect $\mathbf{A}$-field; $\iota: \mathbf{A} \to L$, $\theta = \iota(T)$. Let E be a T-module over L with associated motive $M = M(E)$.

DEFINITION 2.6.1.

(1) We set $H_{\mathrm{DR}}^1(E) = H_{\mathrm{DR}}^1(M) = \tau M/(T-\theta)\tau M$. Higher groups are given by exterior product.
(2) The *Hodge filtration* for E (or M) is defined via

$$F^i H_{\mathrm{DR}}^1(M) = \text{image of } (T-\theta)^i M \cap \tau M \subseteq H_{\mathrm{DR}}^1(M).$$

Recall that, by the definition of T-modules, there exists n such that $(T-\theta)^n M/\tau M = 0$. Thus

$$F^{n+1}H^1_{\mathrm{DR}}(E) = \{0\}.$$

The definition given in 2.6.1 *should* be canonically isomorphic to the definition coming from universal extensions as in the Drinfeld module case, §1.5.

EXAMPLE 2.6.2. Let ψ be a Drinfeld module. Then $(T-\theta)M/\tau M = 0$. Thus

$$F^i H^1_{\mathrm{DR}}(\psi) = \{0\}$$

for $i > 1$. (For more in this case, see **[Ge3]**.)

Now let L be an A-field which we do not necessarily require to be perfect.

EXAMPLE 2.6.3. Let ψ be a rank d Drinfeld module over L with

$$\begin{aligned}\psi_T &= \imath(T)\tau^\circ + g_1\tau + \cdots + g_d\tau^d, \qquad g_d \neq 0;\\ &= \theta\tau^\circ + g_1\tau + \cdots + g_d\tau^d.\end{aligned}$$

Let M be the associated motive. We know that

$$M = \bigoplus_{i=0}^{d-1} L[T]\tau^i m_0,$$

where m_0 is the identity morphism. Let

$$N = \bigwedge_{\mathbf{A}}^{d} M(E);$$

so N is rank one. A basis for N is given by $m_0 \wedge \tau m_0 \wedge \cdots \wedge \tau^{d-1}m_0$. As with tensor products, the action of τ on N is diagonal; thus

$$\begin{aligned}\tau(m_0\wedge\cdots\wedge\tau^{d-1}m_0) &= \tau m_0\wedge\cdots\wedge\tau^d m_0 = (-1)^{d-1}\tau^d m_0\wedge\cdots\wedge\tau^{d-1}m_0\\ &= \frac{(-1)^{d-1}}{g_d}(T-\theta)(m_0\wedge\tau m_0\wedge\cdots\wedge\tau^{d-1}m_0).\end{aligned}$$

Thus N is the motive associated to the rank-one Drinfeld module $\psi^\wedge$ with

$$\psi^\wedge_T = \theta\tau^0 + (-1)^{d-1}g_d\tau.$$

The Drinfeld module $\psi^\wedge$ is called the "determinant module" associated to ψ. From §2.4, we see that this gives a (*noncanonical*) isomorphism of $\wedge^d T_v(\psi)$ and $T_v(\psi^\wedge)$, for $v \neq \mathrm{char}(L)$, as Galois modules.

We have worked above in the basic case of $\mathbb{F}_p[T]$; however, the result holds true for arbitrary r.

Now let L be perfect and E a pure T-module over L with motive $M = M(E)$. Let $d = r(E)$. Then $N = \bigwedge^d_{\mathbf{A}} M(E)$ is a rank-one motive.

IDEA 2.6.4. *N is a tensor power of a rank-one Drinfeld module.*

HINT. L is perfect. So one has left and right division algorithms.

For more, see **[T2]** and **[T3]**.

DEFINITION 2.6.5. Let $L(0)$ be the left $L[T, \tau]$-module $L[T]$ equipped with the τ-action $\tau(\sum a_i T^i) = \sum a_i^p T^i$.

Clearly if M is a left $L[T, \tau]$-module, then $M \otimes L(0) \simeq M$. Moreover, one sees that

$$L(0) \simeq L[T, \tau]/L[T, \tau](\tau - 1).$$

REMARK 2.6.6. If we replace L by $\mathbf{A}$ itself in the definition of $L(0)$ we obtain "$\mathbf{A}(0)$" defined over $\mathbf{A}$.

Now let our perfect field L be a subfield of $\overline{\mathbf{K}}$. Let M_1 and M_2 be two T-motives over L. We assume that these motives are both *pure* and *uniformizable.* From the discussion in the previous two sections, we see that $M_1 \otimes M_2$ is *also* both pure and uniformizable.

Let M be a pure uniformizable motive associated to a T-module E. Let $\Omega/\mathbf{k} = \mathbb{F}_p(T)dt$ and let W be the lattice of E. As in §2.5, put

$$\omega(M) := H^1(M) = \mathrm{Hom}_{\mathbf{A}}(W, \Omega/\mathbf{k}) = H^1_B(E, \Omega/\mathbf{k}) \in \mathrm{Vect}_{\mathbf{k}},$$

where $\mathrm{Vect}_{\mathbf{k}}$ is the category of finite-dimensional $\mathbf{k}$-vector spaces. Moreover, define

$$\omega(L(0)) := H^1(L(0)) := \mathbf{k}.$$

For the concepts of "Tannakian categories" we refer the reader to **[Mi1]** and **[DMOS]**.

IDEA 2.6.7. *The isogeny classes of pure, uniformizable T-motives over L, and $L(0)$, generate a Tannakian category with identity element $\mathbb{1} = L(0)$ and fibre functor ω.*

HINT. We need to embed the category of pure uniformizable motives, with its tensor product, into one with duals. Step 1: Formally invert *all* motives associated to rank-one Drinfeld modules. Step 2: Define

$$M\hat{} = \left(\bigwedge^{d-1} M\right) \otimes \left(\bigwedge^{d} M\right)\hat{}. \qquad \square$$

Assuming all details to be eventually worked out without obstructions, let us denote the Tannakian category of 2.6.7 by $\mathscr{M}(L)$. The reader should note that, unlike classical theory, no additional assumptions should be needed for the construction. The general theory of Tannakian categories then provides the existence of an affine group scheme G over $\mathbf{k}$ and an equivalence of categories between $\mathscr{M}(L)$ and the category $\mathrm{Rep}_{\mathbf{k}}(G)$ of finite-dimensional $\mathbf{k}$-representations of G. One can then ask, e.g., if G is reduced, etc. For more in the case of "potential complex multiplication", we refer the reader to **[A3]**.

The reader should note that the Drinfeld modules generate a Tannakian subcategory of $\mathscr{M}(L)$ which should contain many interesting examples.

Let us call the "Artin motives":= $\mathscr{M}_{\mathbf{A}}(L) :=$ the Tannakian subcategory of $\mathscr{M}(L)$ generated by those $M \in \mathscr{M}(L)$ such that

$$M = \tau M.$$

Let $G = \mathrm{Gal}(L^s/L)$ and let $\mathrm{Rep}_{\mathbf{k}}(G)$ be the Tannakian category of finite-dimensional $\mathbf{k}$-representations of G.

IDEA 2.6.8. *There is an equivalence of categories between the Artin motives and* $\mathrm{Rep}_{\mathbf{k}}(G)$.

HINT. Let M be as above and put $\overline{M} = M \otimes_L \overline{L}$ with the diagonal action of τ. There should be an isomorphism

$$M \simeq \bigoplus_{d=1}^{t} \overline{L}(0)$$

for some t. We thus obtain a representation of G on $\mathbf{A}^t$ which determines M. This should give the equivalence. □

We finish this section with an example that shows how one may use the tensor formalism to obtain a *Legendre formula* [**A2**] for the periods and quasi-periods of §1.5. For comparison with classical theory, we will work in the rank-two case; however, the reader may easily see how to work in general. We will also work with arbitrary $\mathbb{F}_r[T]$; so τ will now be the rth power mapping, τ_r.

Thus, let $L \subseteq \overline{\mathbf{K}}$ be a finite extension of $\mathbf{K}$ which may now be arbitrary. Let ψ be a rank-two Drinfeld module over L given by $\psi_T = T\tau^0 + g_1\tau + g_2\tau^2$, $g_2 \neq 0$. Let φ be the rank-one Drinfeld module given by $\varphi_T = T\tau^0 - g_2\tau$; so φ is the determinant module of ψ.

As we saw in Example 1.5.7, $H^1_{\mathrm{DR}}(\psi, L)$ has a basis consisting of $\{\delta\psi, \eta\}$ where $\eta_T = \tau$ and $H^1_{\mathrm{DR}}(\varphi, L)$ has basis $\{\delta\varphi\}$. We let Γ_ψ and Γ_φ denote the associated lattices; the associated exponential functions are denoted $e_\psi(x)$ and $e_\varphi(x)$. Let $\gamma \in \Gamma_\psi$ and $\gamma' \in \Gamma_\varphi$ be fixed elements.

We begin with some formalism related to the integration of Definition 1.5.9 in a manner analogous to the classical theory of differential equations. Our first task is to define certain operators on power series. Thus, let $h(x) = \sum a_i x^i \in L[[x]]$. We define the power series $h^{\tau^j}(x)$ by

$$h^{\tau^j}(x) = \sum a_i^{r^j} x^i.$$

If $\alpha(\tau) = \sum b_j \tau^j \in L[\tau]$, then the power series $h^{\alpha(\tau)}(x)$ is defined by linearity; i.e.,

$$h^{\alpha(\tau)}(x) = \sum b_j h^{\tau^j}(x).$$

We now define the power series

$$f_\gamma(x) = -\sum_{n=0}^{\infty} e_\psi\left(\frac{\gamma}{T^{n+1}}\right)x^n\,,$$

and, similarly,

$$g_{\gamma'}(x) = -\sum_{n=0}^{\infty} e_\varphi\left(\frac{\gamma}{T^{n+1}}\right)x^n\,.$$

From Lemma 1.5.12(2), one sees immediately that

$$\int_\gamma \eta = f_\gamma^\tau(T)\,.$$

However, the situation is a little more complicated with $\delta\psi$. Let us set

$$\omega_\gamma(x) = f_\gamma(x)\,dx\,,$$

and

$$\omega_{\gamma'}(x) = g_{\gamma'}(x)\,dx\,.$$

One then finds, with a little manipulation, that 1.5.12(2) implies

$$\int_\gamma \delta\psi = \gamma = \lim_{x\to T}(x-T)f_\gamma(x) = \operatorname*{Res}_{x=T}\omega_\gamma(x)\,,$$

and,

$$\int_{\gamma'} \delta\varphi = \gamma' = \operatorname*{Res}_{x=T}\omega_{\gamma'}(x)\,.$$

One then checks readily that f and g satisfy the ("τ-differential") equations:

(1) $$g_{\gamma'}^{\varphi_T}(x) = x g_{\gamma'}(x)\,,$$

and,

(2) $$f_\gamma^{\psi_T}(x) = x f_\gamma(x)\,.$$

Let γ_1, γ_2 be two elements of Γ_ψ. Set

$$W_{\gamma_1,\gamma_2}(x) = \det\begin{pmatrix} f_{\gamma_1}(x) & f_{\gamma_1}^\tau(x) \\ f_{\gamma_2}(x) & f_{\gamma_2}^\tau(x)\end{pmatrix}\,.$$

This $W_{\gamma_1,\gamma_2}(x)$ is a kind of "Wronskian". As before, one computes

$$\operatorname*{Res}_{x=T} W_{\gamma_1,\gamma_2}(x)\,dx = \det\begin{pmatrix} \int_{\gamma_1}\delta\psi & \int_{\gamma_1}\eta \\ \int_{\gamma_2}\delta\psi & \int_{\gamma_2}\eta\end{pmatrix}\,.$$

Since $f_\gamma^{\psi_T}(x) = x f_\gamma(x)$, one finds

$$W_{\gamma_1,\gamma_2}^{\varphi_T}(x) = x W_{\gamma_1,\gamma_2}(x)\,;$$

i.e., $W_{\gamma_1,\gamma_2}(x)$ also satisfies (1). One then deduces the existence of $\gamma_3 \in \Gamma_\varphi$ with

$$W_{\gamma_1,\gamma_2}(x) = g_{\gamma_3}(x)\,.$$

Suppose now that $\{\gamma_1, \gamma_2\}$ generate Γ_ψ. Then, as $\mathrm{Hom}_{\mathbf{A}}(\Gamma_?, \Omega)$ commutes with tensor product, γ_3 must *generate* Γ_φ. Putting all of this together, we have

$$\begin{aligned}\operatorname*{Res}_{x=T} W_{\gamma_1,\gamma_2}(x)\,dx &= \operatorname*{Res}_{x=T} g_{\gamma_3}(x)\,dx \\ &= \int_{\gamma_3} \delta\varphi \\ &= \gamma_3 \neq 0 \\ &= \det\begin{pmatrix} \int_{\gamma_1}\delta\varphi & \int_{\gamma_1}\eta \\ \int_{\gamma_2}\delta\varphi & \int_{\gamma_2}\eta \end{pmatrix},\end{aligned}$$

which is precisely the searched for *Legendre relation*!

For another approach, we refer the reader to **[Ge1]**.

3. L-series and Γ-functions

We now assume that $\mathbf{A} = \mathbb{F}_r[T]$.

Our goal here is modest. We present the definitions of L-series of motives and Γ-functions, and briefly discuss their main properties. Included in this will be the work of Thakur on Γ-functions and the work of Anderson-Thakur and Yu on zeta-values. For much more, see **[A1, AT1, Th1, Th2, Th4, Th6, Y3, Go1, Go5, Go10]**. For the beginnings of an "automorphic" theory of measures that one hopes will eventually have an impact on these functions, see **[Te1, Te2, Go11]**.

Let L_1 be a finite extension of $\mathbf{k} \subset \bar{\mathbf{k}} \subset \bar{\mathbf{K}}$. Let $L \subseteq \bar{\mathbf{k}}$ be the *perfection* of L_1. Thus there is a canonical isomorphism between $\mathrm{Gal}(\bar{\mathbf{k}}/L)$ and $\mathrm{Gal}(\bar{\mathbf{k}}/L_1)$. Moreover, as is easy to see, there is a bijection between $\mathrm{Spec}(\mathcal{O}_L)$ and $\mathrm{Spec}(\mathcal{O}_{L_1})$ where $\mathcal{O}_L$, $\mathcal{O}_{L_1}$ are the rings of $\mathbf{A}$-integers. Let $\wp \in \mathrm{Spec}(\mathcal{O}_L) = \mathrm{Spec}(\mathcal{O}_{L_1})$. Then there exists a *monic* prime $f \in \mathbf{A}$ such that

$$\mathcal{O}_L/\wp = \mathcal{O}_{L_1}/\wp$$

is finite dimensional over $\mathbf{A}/(f)$. Set

$$\alpha = [\mathcal{O}_L/\wp : \mathbf{A}/(f)].$$

DEFINITION 3.1. We set $n\wp = f^\alpha$. We call f^α the *norm of* $\wp$.

Let $n \in \mathbf{A}$ now be monic of degree d.

DEFINITION 3.2. We set

$$\langle n \rangle = n/T^d.$$

Thus, $\langle nm \rangle = \langle n \rangle \langle m \rangle$. Note also that $\langle n \rangle$ is a 1-unit inside $\mathbf{K}$. In particular, by the binomial theorem, we may raise $\langle n \rangle$ to the yth power for $y \in \mathbb{Z}_p$.

DEFINITION 3.3. We set

$$S_\infty = \bar{\mathbf{K}}^* \times \mathbb{Z}_p.$$

We view S_∞ as a topological group whose operation is written *additively*.

Let $s = (x, y) \in S_\infty$ and $n \in \mathbf{A}$ as above.

3.4. DEFINITION. We set

$$n^s = x^d \langle n \rangle^y .$$

Clearly, $(nm)^s = n^s m^s$ and $n^{s_0+s_1} = n^{s_0} n^{s_1}$. Moreover, let $i \in \mathbb{Z}$ and $s = (T^i, i) \in S_\infty$. Then

$$n^s = n^i .$$

We thus denote (T^i, i) by "i". In this fashion, $\mathbb{Z}$ is a *discrete* subgroup of S_∞. The reader should note that the definition of $\langle n \rangle$ in 3.2 uses the uniformizer $\pi = \frac{1}{T}$ at ∞ on $\mathbb{P}^1$. Actually, however, one may create a theory using *any* uniformizer π at ∞ as long as $T\pi$ is a 1-unit. If we are willing to vary our notion of "monic", *any* uniformizer can be used.

Let M now be a pure motive (e.g., a Drinfeld module); in particular, we will allow M to be $L(0)$. Let $\wp$ be a prime of *good reduction* for M and let $M_\wp$ denote the reduction at $\wp$. Let $Q_\wp(x)$ be the characteristic polynomial of the Frobenius at $\wp$ as in §2.2. By 2.2.8 and 2.2.10 (assuming all goes well!), we see that $Q_\wp(x) \in \mathbf{A}[x]$ with pure roots of weight $w(M)$. Set

$$P_\wp(x) = x^{\deg Q_\wp(x)} Q_\wp(1/x) \ ;$$

so $P_\wp(0) = 1$.

DEFINITION 3.5. We set

$$L^{\text{good}}(M, s) = \prod_{\wp \text{ good}} P_\wp(n\wp^{-s})^{-1} .$$

PROPOSITION 3.6. *The Euler product* $L^{\text{good}}(M, s)$ *converges for all* $s = (x, y) \in S_\infty$ *with* $|x| \gg 0$.

PROOF. This is immediate once 2.2.10 is established. □

The subset of S_∞ given by the above proposition is called the "half-plane of convergence for $L^{\text{good}}(M, s)$".

REMARK 3.7. One can use the classical formalism of inertia subgroups to extend the Euler product definition of $L^{\text{good}}(M, s)$ to include *all* finite primes of L. We call this extended function "$L(M, s)$".

Next let $v \in \text{Spec}(\mathbf{A})$ which is assumed monic. We now give a v-adic version of $L(M, s)$.

DEFINITION 3.8. Set

$$\begin{aligned} S_v &= \varprojlim_j \mathbb{Z}/(r^j(r^{\deg v} - 1)) \\ &\simeq \mathbb{Z}_p \times \mathbb{Z}/(r^{\deg v} - 1) . \end{aligned}$$

Let $n \in \mathbf{A}$ be a monic element that is *prime* to v. Let $s_v \in S_v$. Using classical p-adic theory as a guide, it is easy to define n^{s_v} with the usual obvious properties.

Let $\mathbf{k}_v$ be the completion of $\mathbf{k}$ at v and $\overline{\mathbf{k}}_v$ a fixed algebraic closure. Let $(x_v, s_v) \in \overline{\mathbf{k}}_v^* \times S_v$ and $P_\wp(x) \in \mathbf{A}[x]$ as above.

DEFINITION 3.9. We set

$$L_v^{\text{good}}(M, x_v, s_v) = \prod_{\substack{\wp \text{ good} \\ (\wp, v)=1}} P_\wp(x_v^{-\deg n\wp} \cdot n\wp^{-s_v})^{-1}.$$

PROPOSITION 3.10. *The Euler product* $L_v^{\text{good}}(M, x_v, s_v)$ *converges for all* (x_v, s_v) *with* $|x_v|_v > 1$.

PROOF. This is obvious since the coefficients of $P_\wp(x)$ are in $\mathbf{A}$. □

As before, it is clear how to add in the bad primes that do not divide v.

EXAMPLE 3.11.

(1) Let $L_1 = \mathbf{k}$ and $M = L(0)$. It is then easy to see that

$$L(M, s) = \prod_{\wp \in \text{Spec}(\mathbf{A})} (1 - n\wp^{-s})^{-1}$$
$$= \prod_{f \text{ monic prime}} (1 - f^{-s})^{-1}.$$

We call this function "$\zeta_{\mathbf{A}}(s)$".

(2) Let C be the Carlitz module. Then, one computes that

$$L(C, s) = \zeta_{\mathbf{A}}(s+1).$$

(3) The reader will easily work out the obvious v-adic analogs of (1) and (2).

It is clear that

$$\zeta_{\mathbf{A}}(s) = \sum_{n \text{ monic}} n^{-s}$$
$$= \sum_{j=0}^{\infty} x^{-j} \left(\sum_{\deg n = j} \langle n \rangle^{-y} \right).$$

This last formula is the key to the analytic properties of $\zeta_{\mathbf{A}}(s)$. For fixed y, we can investigate the analytic properties of the power series in x^{-1}. For instance, the "order of zero" at a point $(x_0, y_0) \in S_\infty$ is defined to be the order of the power series in x, $\zeta_{\mathbf{A}}(x, y_0)$ at $x = x_0$, etc. Similarly, one introduces such power series into the v-adic theory.

For the proofs of the following results, which depend only on rather elementary estimates, we refer the reader to **[Go2, Go10]**.

THEOREM 3.12.

(1) $\zeta_{\mathbf{A}}(s)$ *can be continued to a continuous function on all of* S_∞ *such that, for fixed* $y \in \mathbb{Z}_p$, $\zeta_{\mathbf{A}}(s)$ *is given by an* **entire** *power series in* x^{-1}.

(2) *$\zeta_{\mathbf{A},v}(s)$ can be continued to a continuous function on all of $\overline{\mathbf{K}}_v^* \times S_v$ such that, for fixed $s_v \in S_v$, $\zeta_{\mathbf{A},v}(x_v, s_v)$ is given by an entire power series in x_v^{-1}.* □

Let $j \geq 0$ be an integer. Set

$$z(j, x) := \zeta_{\mathbf{A}}(T^{-j}x, -j) := \sum_{t=0}^{\infty} x^{-t} \left(\sum_{\substack{n \in \mathbf{A} \\ \deg n = t}} n^j \right).$$

By Theorem 3.12.1, $z(j, x)$ is also entire. But it also has $\mathbf{A}$-coefficients. Thus the inside sum must vanish for $t \gg 0$, and so $z(j, x) \in \mathbf{A}[x^{-1}]$. In *particular,*

$$\zeta_{\mathbf{A}}(-j) = z(j, 1) \in \mathbf{A}.$$

The above algebraicity result is the *key* to relating $\zeta_{\mathbf{A}}(s)$ and $\zeta_{\mathbf{A},v}(x_v, s_v)$. Indeed one now proves:

THEOREM 3.13. *The functions $(1 - x_v^{-\deg v} v^j) z(j, x_v)$ interpolate v-adically to $\zeta_{\mathbf{A},v}(x_v, s_v)$.* □

We say that $\zeta_{\mathbf{A}}(s)$ is an "essentially algebraic entire function"; the word "essentially" is used as continuity implies that $\zeta_{\mathbf{A}}(s)$ is determined by the *polynomials with* $\mathbf{A}$*-coefficients* $\{z(j, x)\} \subseteq \mathbf{A}[x^{-1}]$, $j \geq 0$. One "hopes" that *all* L-series of motives are also essentially algebraic entire functions. This is known in the case of *complex multiplication* **[Go4, Go10]**. The idea is that the L-series of motives with complex multiplication factor according to "Hecke characters" in the classical fashion; the reader will have little difficulty defining such characters once it is realized that their values must be in $\overline{\mathbf{K}}$. Thus the analyticity is reduced to the abelian case where it follows from known techniques as in 3.12.

One can also ask if the functions $L(M, s)$ satisfy *functional equations.* There is indeed evidence that these functions possess a deeper structure, but the situation seems to be *very* nonclassical. Below we present some of this evidence. For more, especially some computations that *look* suspiciously like they *might* belong to a functional equation, we refer the reader to **[Go10]**.

REMARKS 3.14. For simplicity, in these remarks, we assume that $\mathbf{A}$ is the basic ring $\mathbb{F}_p[T]$. (1) Let X be an integral scheme of finite type over Spec($\mathbf{A}$) and let $\widetilde{X}$ be its perfection. Let M be a family of pure T-motives over $\widetilde{X}$. For each good (in terms of the reduction of M) closed point $z \in \widetilde{X}$, let $P_z(x)$ be the inverse characteristic polynomial of Frobenius acting on the fibre of M over z. We set

$$L^{\text{good}}(M/X, s) = \prod_z P_z(nz^{-s})^{-1} \quad \text{etc.},$$

where nz is defined in the obvious fashion. In this way, we get L-series that merge both classical and function field motives! Indeed, if X is a scheme over $\mathbf{A}/\wp$, $\wp \in \mathrm{Spec}(\mathbf{A})$, and M is the constant motive $X(0)/X$ (whose fibres are just the motive $L(0)$) then $L(X(0)/X, s)$ is just the reduction modulo p of the classical zeta function of X (in the variable u) where one substitutes $n\wp^{-s}$ for u. The reader is invited to use classical theory and the theory of this section (such as Theorem 3.12) as a guide towards conjecturing reasonable properties of such general functions. As to what information is contained in such functions, only future research will tell (but see part (2) just below!). (2) Recall, from Definition 2.1.1, that the T-motives over a field L are modules over the noncommutative ring $L[T, \tau]$ (for a family of motives over X, one replaces L with the structure sheaf of X). Inside $L[T, \tau]$ lies the *commutative* subring $\mathbb{F}_p[T, \tau]$. On the other hand, there are very natural isomorphisms

$$\mathbb{F}_p[T, \tau] \simeq \mathbb{F}_p[T, x] \simeq \mathbf{A}[x] ;$$

where x is an indeterminate. Now, IF the functions of part (1) are essentially algebraic, then they will, by definition, produce elements of $\mathbf{A}[x^{-1}]$ at negative integers. Thus, for example, we can view the "numerators" of these elements as operators on the T-motive (or a family of T-motives).

Recall that, in Example 1.3.9, we discussed the period ξ of the Carlitz-module. We then have

THEOREM 3.15.

(1) *Let* $j > 0$ *with* $j \equiv 0 \ (r-1)$. *Then* $0 \neq \zeta_{\mathbf{A}}(j)/\xi^j \in \mathbf{k}$.

(2) *Let* j *be as in* (1). *Then,* $\zeta_{\mathbf{A}, v}(1, j) = 0$.

(3) *Let* $j > 0$ *now be an arbitrary positive integer. Then*

$$\zeta_{\mathbf{A}}(-j) = 0 \iff j \equiv 0 \ (r-1).$$

These zeroes are of first order (*in the sense given above*).

(4) $z(0, x) = 1$. □

(Actually, v-adic continuity implies (3) $\Rightarrow$ (2).) The results of Theorem 3.15 are clearly in line with what one would expect from the analogy with the Riemann zeta function. Part (1) was established by L. Carlitz in 1935 and independently, but much later, by the present author. The zeroes given in part (3) of the theorem are the "trivial zeroes" of $\zeta_{\mathbf{A}}(s)$. The zeroes of part (2) are the "v-adic trivial zeroes at positive integers". For the arithmetic associated to these zeta-values we refer the reader to **[Go1, Go10]**. We note, however, that these zeta-values are related to the abelian extensions $\mathbf{k}(f)$ (see Example 1.3.9) in somewhat the same manner as Bernoulli numbers are related to cyclotomic number fields. As in the classical case, this relationship also arises from *class number formulae* and *cyclotomic units* (i.e., nontrivial quotients of f-division points). In particular, there is even an analog of the *Kummer-Vandiver conjecture* **[Go8**, p. 181; **Go10**, §5.7] involving the class of nonnegative integers called "magic numbers". (A positive integer i is a magic number if it is either 0 or all its r-adic digits, *except perhaps* the highest, are $r-1$ — so $i = cr^j + (r^j - 1)$ for some j and $0 < c < r$.)

In the classical theory of mixed motives, **[D2]**, one is interested in extensions of $\mathbb{Z}(0)$ by $\mathbb{Z}(n)$. One conjectures deep relationships between such extensions, **K**-groups, and special values of Dedekind zeta functions. We have seen how to construct **A**-versions of $\mathbb{Z}(0)$ and $\mathbb{Z}(1)$ (these **A**-versions are $L(0)$, and $L(1) \approx$ Carlitz module over L up to isogeny). Thus one is led to look for relationships between extensions of such T-motives and the zeta-values at integers *not* covered by Theorem 3.15. This leads naturally to the work of Anderson-Thakur and Yu.

IDEA 3.16. *Let L be an algebraically closed* **A**-*field. Let E be an abelian T-module over L and let M be the associated motive. Then*

$$E(L) = \operatorname{Ext}^1_{L[T,\tau]}(M, L(0)).$$

HINT. Consider $L\left(\left(\frac{1}{T}\right)\right)$, $L\left[\left[\frac{1}{T}\right]\right]$, and $L[T]$ to be $L[T, \tau]$-modules where

$$\tau\left(\sum a_i T^i\right) = \sum a_i^r T^i.$$

Let $\widetilde{L}$ be L considered as an $L[\tau]$-module with $\tau(\alpha) = \alpha^r$, $\alpha \in L$.

Now, by definition, we have

$$E(L) = \operatorname{Hom}_{L[\tau]}(M, \widetilde{L}),$$

as in the proof of Proposition 2.3.3. Moreover, by duality tricks such as in the proof of Theorem 2.3.5, we can see

$$\begin{aligned} E(L) &= \operatorname{Hom}_{L[T,\tau]}\left(M, L\left[\left[\frac{1}{T}\right]\right]/L\right) \\ &= \operatorname{Hom}_{L[T,\tau]}\left(M, L\left(\left(\frac{1}{T}\right)\right)/L[T]\right). \end{aligned}$$

Now show that

$$\operatorname{Ext}^1_{L[T,\tau]}\left(M, L\left(\left(\frac{1}{T}\right)\right)\right) = 0$$

(this may follow from the filtration by L-subspaces of M given by Proposition 2.1.5). Moreover, show also that $\operatorname{Hom}_{L[T,\tau]}\left(M, L\left(\left(\frac{1}{T}\right)\right)\right)$ vanishes. The result now follows by the associated long exact sequence. □

Thus classical theory leads to the tensor powers, $C^{\otimes m}$, of the Carlitz-module as given in Example 2.4.3. We know, from general theory, that $C^{\otimes m}$ is pure of rank one and uniformizable.

Let

$$e_m(x), \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix},$$

be the exponential of $C^{\otimes m}$; so $e_1(x)$ is just the exponential of the Carlitz-module.

The following result is found in **[AT1]**.

THEOREM 3.17 (ANDERSON-THAKUR).

(1) *$H_1(C^{\otimes m})$ is generated by $(\omega_1, \ldots, \omega_m)$ with $\omega_m = \xi^m$; where, as in Example 1.3.9, ξ is the period of the Carlitz-module.*
(2) *There exists a special point $(\ell_1, \ldots, \ell_m) \in \overline{\mathbf{K}}^m$ such that*

$$e_m(\ell_1, \ldots, \ell_m) = (y_1, \ldots, y_m) \in (\mathbf{k}^s)^m.$$

(3) $\ell_m = \zeta_{\mathbf{A}}(m)$.
(4) *If $m \not\equiv 0 \bmod(r-1)$, then $(y_1, \ldots, y_m)$ is* **not** *a torsion point for $C^{\otimes m}$.* □

EXAMPLE 3.18. Put $L_0 = 1$ and, for $j > 1$,

$$L_j = (T^{r^j} - T)\cdots(T^r - T).$$

Then one shows that L_j is the least common multiple of all monics of degree j. As an illustration of Theorem 3.17, we have the formula of *Carlitz-Thakur*:

$$\zeta_{\mathbf{A}}(1) = \sum_{j=0}^{\infty} \frac{(-1)^j}{L_j}, \qquad e_1(\zeta_{\mathbf{A}}(1)) = 1.$$

Moreover, as in **[Th4]**, one can show that $\zeta_{\mathbf{A}}(1)$ is also a function field version of *Euler's constant.*

Using 3.17, Jing Yu **[Y3]** was able to establish analogs of the results of *Hermite-Lindemann* and show the following result (see **[Th1]** for the first discussion of these issues):

THEOREM 3.19 (YU).

(1) *$\zeta_{\mathbf{A}}(j)$ is transcendental over* $\mathbf{k}$ *for all $j \geq 1$.*
(2) *Let $j \not\equiv 0 \bmod(r-1)$. Then $\zeta_{\mathbf{A}}(j)/\xi^j$ is also transcendental over* $\mathbf{k}$.
(3) *Let $j \not\equiv 0 \bmod(r-1)$. Then $\zeta_{\mathbf{A},v}(1, j)$ is transcendental over* $\mathbf{k}$. □

Let $i > 0$ be divisible by $r - 1$. We know from 3.15(2) that $\zeta_{\mathbf{A},v}(1, i) = 0$. For a few special v, Thakur **[Go10**, §5.2**]** establishes simplicity of this zero. For general v, the techniques of **[AT1]** allow us to express the first derivative (i.e., $\frac{d}{dx_v}$) of $\zeta_{\mathbf{A},v}(x_v, s_v)$ at $(1, i)$ as a linear combination of logarithms. Perhaps the techniques of Yu **[Y3]** may be extended to show the nonvanishing of these derivatives, and thus the simplicity of the v-adic trivial zeroes at positive integers for all v.

Next we turn our attention to Γ-functions; some references for the material presented here are **[Go8]** and **[Th4]**. Recall that in Example 1.3.9, we defined the elements $D_i \in \mathbf{A}$ for $i > 0$. We now extend this definition to $i = 0$ by setting $D_0 = 1$. Let j be a nonnegative integer that we write r-adically as

$$j = \sum_{t=0}^{\nu} a_t r^t, \qquad 0 \leq t < r.$$

We then have the following notion of "factorial":

DEFINITION 3.20 (CARLITZ). We set

$$\Pi(j) = \prod_{t=0}^{\nu} D_t^{a_t}.$$

The process of creating $\{\Pi(j)\}$ out of $\{D_t\}$ via the r-adic digits of j is a general feature of the theory; e.g., besides factorials and Γ-functions (see below), there are the "Gauss sums" of **[Th2]** and **[Th5]**. (These Gauss sums are sums over finite fields of multiplicative characters multiplied by "additive characters" coming from the exponential of the Carlitz module C.) Moreover, recall that $\frac{1}{D_t}$ is the coefficient of z^{r^t} in the Carlitz exponential. The use of the digits also allows us to find the "missing" coefficients $\frac{1}{\Pi(j)}$ of arbitrary z^j.

The Carlitz factorial was shown by Sinnott many years ago to have divisibility properties at primes of $\mathbf{A}$ in *exact* accordance with what is known for $n!$ classically.

The next lemma can easily be seen by the closed form formula for D_t of 1.3.9; it also follows from the interpretation of D_t as the product of all monics of degree t.

LEMMA 3.21. *We have* $\langle D_t \rangle \to 1 \in \mathbf{K}$ *as* $t \to \infty$. □

DEFINITION 3.22. Let $y \in \mathbb{Z}_p$ be written r-adically as $y = \sum_{t=0}^{\infty} a_t r^t$ with $0 \le t < r$. Then we set

(1) $\Pi_\infty(y) = \prod_{t=0}^{\infty} \langle D_t \rangle^{a_t}$, and
(2) $\Gamma_\infty(y) = \Pi_\infty(y-1)$.

The function $\Gamma_\infty(y)$ clearly is continuous from $\mathbb{Z}_p \to \mathbf{K}$. Thakur **[Th3]** has shown that $\Gamma_\infty(y)$ possesses many of the properties of Euler's Γ-function. These include a functional equation, multiplication formulae, etc. The proofs of these theorems are remarkable in that they are formal consequences of certain results on r-adic expansions of elements of $\mathbb{Z}_p$. For instance,

$$\Gamma_\infty(y)\Gamma_\infty(1-y) = \Gamma_\infty(0)$$

(which is readily seen from the corresponding result on $\Pi_\infty(y)$) while $\Gamma_\infty(0)$ is calculated as the 1-unit piece of the period of the Carlitz module — in particular, it differs from this period by an algebraic element. Let $v \in \operatorname{Spec}(\mathbf{A})$. Then, following the formalism of the classical p-adic Γ-function, one can also interpolate the Carlitz factorials to continuous functions from $\mathbb{Z}_p \to \mathbf{A}_v^*$. There are similar results as above (functional equations, etc.) for these functions, except that $\Gamma_v(0) = \pm 1$. There are also connections with Gauss sums **[Th2]** in the manner of the Gross-Koblitz theorem.

The function $\Gamma_\infty(y)$ is called the "arithmetic Γ-function" because it is related to the abelian extensions of $\mathbf{k}$ obtained by adjoining constant field extensions. On the other hand, the abelian extensions of $\mathbf{k}$, $\mathbf{k}(f)$, are "geometric" in that they contain no constant field extensions. We now define a "geometric Γ-function" related to these extensions.

DEFINITION 3.23. Let j be a nonnegative integer. Let $\mathbf{A}(j) = \{\alpha \in \mathbf{A} \mid \deg(\alpha) < j\}$. Then we set

$$e_j(x) = \prod_{\alpha \in \mathbf{A}(j)} (x + \alpha).$$

As we saw in the discussion of the exponentials of Drinfeld modules, the polynomials $\{e_j(x)\}$ are $\mathbb{F}_r$-linear. Moreover, clearly $e_0(x) = x$.

LEMMA 3.24. *Let j be as above. Then*

$$1 + \frac{e_j(x)}{D_j} = \prod_{\substack{n \text{ monic} \\ \deg(n)=j}} \left(1 + \frac{x}{n}\right).$$

PROOF. Let n be monic of degree j. Then it is easy to see that $D_j = e_j(n)$. The result now follows directly as both sides have the same roots and constant term. □

DEFINITION 3.25.

(1) Let $(x, y) \in \overline{\mathbf{K}} \times \mathbb{Z}_p$, with $y = \sum_{t=0}^{\infty} a_t r^t$, $0 \leq a_t < r$. Then we set

$$g(x, y) = \prod_{t=0}^{\infty} \left(1 + \frac{e_t(x)}{D_t}\right)^{a_t}.$$

(2) We set $\Pi_0(x, y) = g(x, y)^{-1}$ and $\Gamma_0(x, y) = \Pi_0(x, y - 1)$.

The function, $\Gamma_0(x, y)$, is the geometric Γ-function. Note that it is defined on a slightly larger domain than S_∞. In **[Th4]** one can find the v-adic versions of $\Gamma_0(x, y)$ as well as its multiplication formulae, Chowla-Selberg formulae, etc. The function $\Gamma_0(x, y)$ has a relationship to the exponential, $e(x)$, of the Carlitz module somewhat analogous to the relationship between Euler's Γ-function and $\sin(x)$. This is based on the following result which follows easily from Lemma 3.24:

LEMMA 3.26. *We have*

$$\prod_{\zeta \in \mathbb{F}_r^*} g(\zeta x, (1 - r)^{-1}) = \frac{e(\xi x)}{\xi x}. \qquad \square$$

Finally, we merely put the two definitions together to obtain the "total" Γ-function (or, rather, THE Γ-function):

$$\Gamma(x, y) := \frac{1}{x} \Gamma_\infty(y) \Gamma_0(x, y).$$

From the above discussion, we see that $\Gamma(x, y)$ also has a functional equation, etc.

In the function field theory, one often sees that analogs of classical objects, such as Γ-functions and zeta functions, remarkably "break apart" into smaller components (such as the above factorization for $\Gamma(x, y)$). Function fields are somehow "large" and so there is room for such things as distinct arithmetic and geometric "cyclotomic" abelian extensions as explained above (see **[Go10]** for more, such as geometric and arithmetic conditions for cyclicity — along the lines of the Kummer-Vandiver conjecture — etc.). On the other hand, number fields seem much smaller; in the classical theory objects play *many* roles at once. In fact, this phenomenon is highlighted in the new paper **[Th5]** of Thakur. In this work, Thakur looks at Gauss sums over general $\mathbf{A}$. He finds that the relationship between the analogs of $\{D_i\}$ and the appropriate exponential functions now breaks down. He thus constructs yet a third (!) Γ-function based on the coefficients of the exponentials and not $\{D_i\}$. These functions are then used in a Gross-Koblitz result for the Gauss sums!

One is left with trying to give $\zeta_{\mathbf{A}}(s)$ a relationship with $\Gamma(x, y)$ and a "functional equation" of some sort. At the present time, one has some tantalizing evidence in this regard that we now summarize; see **[Go10]** for a more thorough exposition. Thus, let j be a nonnegative integer written r-adically as $j = \sum_{t=0}^{\nu} a_t r^t$. We set

$$G_j(x) = \prod_{t=0}^{\nu} e_t(x)^{a_t}.$$

Now let $v \in \operatorname{Spec}(\mathbf{A})$ and let $f(x) : \mathbf{A}_v \to \mathbf{A}_v$ be a continuous function. Then, one has a version of *Mahler's Theorem* (due to C. Wagner) that allows us to uniquely expand $f(x)$ as

$$f(x) = \sum_{j=0}^{\infty} b_j \frac{G_j(x)}{\Pi(j)},$$

where $\{b_j\} \subset \mathbf{A}_v$ and $b_j \to 0$ as $j \to \infty$. Conversely, any such series gives rise to a continuous function.

Now, as in Iwasawa theory, this description of continuous functions gives rise to a dual description of v-adic measures with values in $\mathbf{A}_v$; the notion of a v-adic measure, μ, is the obvious translation of the notion of a p-adic measure. By using Riemann sums, one can actually integrate continuous functions against such measures. In classical p-adic Iwasawa theory, one obtains a description of measures in terms of formal power series. For $\mathbf{A}_v$, it is quite remarkable that one obtains instead a description using formal **divided power** series **[Go6, Go9, Go10, Th3]**. Thus to μ we associate the formal divided power series

$$f_\mu(Z) = \sum_{j=0}^{\infty} c_j \frac{Z^j}{j!},$$

where

$$c_j = \int_{\mathbf{A}_v} \frac{G_j(t)}{\Pi(j)}\, d\mu(t).$$

The measure μ is then uniquely given by $f_\mu(Z)$.

Basic to the theory of v-adic measures is the notion of *Dirac measure*, δ_α, associated to $\alpha \in \mathbf{A}_v$. Thus if $f: \mathbf{A}_v \to \mathbf{A}_v$ is a continuous function, then, by definition,

$$\int_{\mathbf{A}_v} f(x)\, d\delta_\alpha = f(\alpha).$$

A first connection between the theory of the gamma and zeta functions can now be seen. Indeed, the divided power series associated to δ_α is

$$f^\alpha(Z) = \sum \frac{G_j(\alpha)}{\Pi(j)} \frac{Z^j}{j!},$$

and **[Go6]** one can show that

$$g(\alpha, y) = \sum \frac{G_j(\alpha)}{\Pi(j)} \binom{y}{j}.$$

As happens classically with the Riemann zeta function, it turns out that the values of $\zeta_{\mathbf{A}}(s)$ at negative integers give rise to v-adic measures; these measures are actually constructed out of the Dirac measures given above. In fact the functions $\{z(j, x)\}$ themselves give rise to a *family* of measures μ_x. One can therefore ask about the description of these measures as divided power series. A priori this description appears to be extremely complicated, though it can be seen to be independent of v. Let us write

$$f_{\mu_x}(Z) = \sum m(x, j) \frac{Z^j}{j!}.$$

We then have the following result **[Th3]**:

THEOREM 3.27 (THAKUR). *We have*

(1) $m(x, j) \neq 0$ *if and only if* j *is a magic number.*
(2) *Let* $0 < j = cr^t + (r^t - 1)$, $0 < c < r$, *be a magic number. Then we have*

$$m(x, j) = \begin{cases} (-1)^t x^{-t}, & \text{if } c < r-1, \\ (-1)^t x^{-t}(1 - x^{-1}), & \text{if } c = r-1. \end{cases}$$

(3) $1 = m(x, 0)$. □

Thakur's result is a quite hopeful sign, though what is implied by this result is not yet known. Still, it links two phenomena that have no obvious relationship with each other. In **[Go6]** and **[Go9]**, the connection with divided power series is further explored using the interpretation of divided power series as hyperdifferential operators. However, much remains to be accomplished.

Finally, using Teitelbaum's modular measures **[Te1, Te2]**, one can obtain "one-variable" functions (i.e., functions from $\mathbb{Z}_p$ to $\mathbf{K}$) with classical style functional equations **[Go11]**. It remains to be seen if these functions may be related to the two-variable functions presented here.

4. Complements

Suppose now that L is an $\mathbb{F}_p[T]$-field that is *not* perfect. How can one develop a theory of motives over L? There are two possible solutions. The first is the one that we have used implicitly and consists of simply embedding L into its perfection.

There is a drawback to using the perfection, however. Indeed, let R be the $\mathbb{F}_p[T]$-integers in the perfection of $\mathbb{F}_p(T)$. Then one sees that R is *not* Noetherian. Thus, the second approach involves working only over L itself. But we have now lost the use of the division algorithms, for the mapping

$$\tau: L[\tau] \to L[\tau], \qquad \tau(\alpha) = \alpha^p ,$$

does *not* have image $L[\tau]\tau$. As a consequence, one finds the existence of *potentially additive* groups over L; i.e., group schemes H over L such that H is *not* isomorphic to $\mathbb{G}_a^d$ but $H \otimes L^{\text{perf}}$ is where L^{perf} is the perfection.

Let M be a left $L[\tau]$-module. We say that M is *potentially free* if $M \otimes L^{\text{perf}}$, with $\tau(m \otimes \alpha) = \tau(m) \otimes \alpha^p$, is a free $L^{\text{perf}}[\tau]$-module. Descent then establishes **[A2]** an anti-equivalence between potentially additive groups over L and potentially-free, finitely generated $L[\tau]$-modules. The anti-equivalence is given by the functor $\operatorname{Hom}_L(H, \mathbb{G}_\alpha)$ with the evident $L[\tau]$-module structure. One then can begin to build a theory of "motives" based on these more general objects over L itself.

Finally, it remains to discuss the situation of general $\mathbf{A}$. As in §1.1, we choose $\alpha_T \in \mathbf{A}$ with a pole at ∞ of order prime to p. Then the map

$$\mathbb{F}_p[T] \hookrightarrow \mathbf{A}, \qquad T \mapsto \alpha_T$$

makes $\mathbf{k}$ into a separable extension of $\mathbb{F}_p(T)$. Recall that $\mathscr{C}$ is the curve associated to $\mathbf{k}$.

Let Δ be the kernel of the map $\mathbf{A} \otimes_{\mathbb{F}_p} \mathbf{A} \to \mathbf{A}$, $a \otimes b \mapsto ab$.

DEFINITION 4.1. Let L be a perfect $\mathbf{A}$-field (thus, also, an $\mathbb{F}_p[T]$-field). An $\mathbf{A}$*-motive* over L consists of a T-motive M and an injection $\lambda: \mathbf{A} \hookrightarrow \operatorname{End}_L(M)$ extending the T-action. Let $\mathbf{A} \otimes_{\mathbb{F}_p} \mathbf{A}$ act on $M/\tau M$ (on the left as scalars and on the right by λ). We further require that this $\mathbf{A} \otimes_{\mathbb{F}_p} \mathbf{A}$ action factors through $\mathbf{A} \otimes_{\mathbb{F}_p} \mathbf{A}/\Delta^m$ for some m.

The definition of $\mathbf{A}$-*module* is dual to the notion of motive. The reader will see that Definition 4.1 can be rephrased in terms of $(L \otimes_{\mathbb{F}_p} \mathbf{A}[\tau])$-modules in the obvious fashion. It is then an exercise to translate the results on T-motives into results on $\mathbf{A}$-motives; replacing free $L[T]$-modules with projective $\mathbf{A}$-modules and so on.

LEMMA 4.2 **[A1**, 4.2.1]. *The functor* $P \mapsto \operatorname{Hom}_{\mathbb{F}_p[T]}(P, \mathbb{F}_p[T]dT)$ *of finitely generated projective* $\mathbf{A}$*-modules is represented by* $\Gamma(\mathscr{C} - \infty, \Omega_{\mathscr{C}})$.

PROOF. We know that $\mathbf{k}$ is a separable extension of $\mathbb{F}_p(T)$. Thus the trace gives rise to a nondegenerate pairing $(x, y) \mapsto \mathrm{Tr}(xy)$. Thus the functor in question is represented by the projective $\mathbf{A}$-module

$$D := \{f \in \mathbf{k} \mid \mathrm{Tr}(fg) \in \mathbb{F}_p[T] \text{ for all } g \in \mathbf{A}\}.$$

However, one checks that

$$D = \{f \in \mathbf{k} \mid f\,dT \in \Gamma(\mathscr{C} - \infty, \Omega_{\mathscr{C}})\}.$$

The result follows. □

Lemma 4.2 is *precisely* the reason why we imposed the degree condition on our choice of α_T to begin with.

To define L-series in general, one needs to exponentiate ideals which may be nonprincipal, or even principal but nonpositively (where "positive" is the generalization of "monic" to arbitrary $\mathbf{A}$) generated. In **[Go10]** (and, to some extent, **[Go4]**) this is carried out. One sees that, by perhaps making finitely many choices, one can define I^s, for $s \in S_\infty$ and $I \subset \mathbf{A}$ an ideal, in a fashion that *agrees* with our previous definition on principal and positively generated ideals. One then proceeds as before to define and handle L-series. For more on general $\mathbf{A}$ along the lines of Theorem 3.17, see **[Th7]** and **[A6]**. See (a) at the end of the paper for a theory that allows $\mathbf{A}$ to also have *arbitrary* dimension.

5. Fermat families

Recall that the classical Fermat equation $x^n + y^n = z^n$ is equivalent (through a trivial change of variables) to the equation

$$x^n - y^n - z^n = 0.$$

Writing things in this fashion makes the connection with cyclotomic fields clear. Indeed, the nth roots of unity are the roots of the equation $u^n - 1 = 0$ and

$$x^n - y^n - z^n = y^n((x/y)^n - 1) - z^n = y^n(u^n - 1) - z^n,$$

where $u = x/y$. This is the basis of much classical work in this area.

Let $\mathbf{A} = \mathbb{F}_r[T]$. The connection with cyclotomic fields also gives a hint how to create a "Fermat equation" out of Drinfeld modules. Indeed, we have *seen how the division values of the Carlitz module give rise to geometric cyclotomic extensions of* $\mathbf{k}$. *Let* $a \in \mathbf{A}$ *have degree* d. *The division values* associated to a are just the roots of the equation $C_a(u) = 0$; so the above prescription leads us **[Go7]** to the equation

$$\mathscr{F}_0(a)(x, y, z) = 0,$$

where

$$\mathscr{F}_0(a)(x, y, z) = y^{r^d} C_a(x/y) - z^{r^d}.$$

Clearly, $\mathscr{F}_0(a)(x, y, z)$ is a homogeneous polynomial of degree r^d.

A more subtle inhomogeneous equation associated to C is

$$\mathcal{F}_1(a)(x, y, z) = 0$$

where

$$\mathcal{F}_1(a)(x, y, z) = y^{r^d} C_a(x/y) - z^p .$$

This inhomogeneous equation is more properly thought of as the correct analog of the Fermat equation. One is interested in finding nontrivial solutions to these equations; that is, a solution $(x, y, z) \in \mathbf{A}^3$ with $0 \neq xyz$. A nontrivial solution to the homogeneous equation is easily seen to give rise to a nontrivial solution of the inhomogeneous one. Note that if (x, y, z) is a nontrivial solution of $\mathcal{F}_0(a)(x, y, z) = 0$ or $\mathcal{F}_1(a)(x, y, z) = 0$ then one can find a solution with $\gcd\{x, y\} = \gcd\{y, z\} = 1$.

Let $a \in \mathbf{A}$ be a linear polynomial. One sees that the projective curve associated to $\mathcal{F}_0(a)(x, y, z) = 0$ is isomorphic to the projective line **[Go7]**. Thus we deduce the infinitude of solutions to both types of equations; this is analogous to what happens classically for $x^2 + y^2 = z^2$.

The study of these equations was recently taken up by L. Denis **[De1]**. He shows the following result:

THEOREM 5.1 (DENIS).

(1) *Let* $r \neq 2$ *and* $\deg(a) > 1$ *or* $r = 2$ *and* $\deg(a) > 2$. *Then both* $\mathcal{F}_0(a)(x, y, z) = 0$ *and* $\mathcal{F}_1(a)(x, y, z) = 0$ *have only a finite number of solutions with* $\gcd\{x, y\} = \gcd\{y, z\} = 1$.

(2) *Let* $r \geq 3$ *and* $\deg(a) \geq 2$. *Then* $\mathcal{F}_0(a)(x, y, z) = 0$ *has no nontrivial solutions.*

(3) *Let* $r \geq 3$, $p \neq 2$, *and* $\deg(a) \geq 2$. *Then* $\mathcal{F}_1(a)(x, y, z) = 0$ *has no nontrivial solutions.* □

Of course, the second and third parts of the theorem establish "Fermat's Last Theorem" in this context! In fact, for $r \neq 2$ there is a particularly simple proof.

More generally, let E be a T-module over $\mathbf{k}$, let $a \in \mathbf{A}$, and suppose that E is g-dimensional. In order to distinguish between vectors in $\mathbb{G}_a^g$ and scalars, we now use boldface to denote vectors. Thus, suppose that

$$a_E \cdot \mathbf{x} = (aI_g + N_1)\tau^0 + a_1\tau + \cdots + a_d\tau^d$$

where $\{N_1, a_i\} \subset M_g(\mathbf{A})$, N_1 is nilpotent, $a_d \neq 0$ and where we continue to let $\tau = \tau_r$, the rth power mapping. We then have the homogeneous system of equations

$$\mathcal{F}_0^E(a)(\mathbf{x}, y, \mathbf{z}) = 0 ,$$

where

$$\mathcal{F}_0^E(a)(\mathbf{x}, y, \mathbf{z}) = y^{r^d} a_E \cdot (\mathbf{x}/y) - \tau_{r^d}(\mathbf{z}) .$$

The inhomogeneous version $\mathcal{F}_1^E(a)(\mathbf{x}, y, \mathbf{z}) = 0$ is formed by replacing τ_{r^d} with τ_p.

The above families of equations involve g equations in $2g+1$ unknowns. The original equation involving the Carlitz module appears as the case $g = 1$ and $E = (C)$. In particular, if the T-module is a Drinfeld module over $\mathbf{k}$, one obtains one equation in three unknowns. These families of equations were also studied by Denis in **[De1]**. He establishes a number of *partial* results on them such as the following:

THEOREM 5.2. *Let a_d' be the derivative of a_d with respect to T. Suppose that a_d is invertible, $d \geq 1$, and $r^d > 2$. Then under these conditions $\mathcal{F}_1^E(a)(\mathbf{x}, y, \mathbf{z}) = 0$ has at most a finite number of solutions with* $\gcd\{x_1, \dots, x_g, y\} = 1$. □

In particular, some of Denis' results apply to the tensor power $E = C^{\otimes g}$ of the Carlitz module and *some* a. For instance, let $p > g$, $a = T^{gh+1}$, $h \geq 1$, and $r^h > 2$. Then the equation

$$\mathcal{F}_1^{C^{\otimes g}}(a)(\mathbf{x}, y, \mathbf{z}) = 0$$

has only finitely many solutions with $\gcd\{x_1, \dots, x_g, y\} = 1$! So there really is a great deal to be said about such general Fermat families of equations!

Denis' methods involve examining the difference between the "Weil height" and the "canonical height" of T-modules.

The modern approach to the classical Fermat equation involves the connection with elliptic curves, automorphic forms, and Galois representations **[Se2]**. One aspect of this approach is that one is also able to handle twisted Fermat equations of the form

$$x^p + y^p + Lz^p = 0$$

for p a prime number and for *certain* L **[Se2**, §4.3]. Of course, once one has a pth root of L in the base field, the above equation is isomorphic to the Fermat equation. These new equations also have their analogues in the function field theory. Indeed, for $\mathbb{F}_r[T]$, all rank-one Drinfeld modules over $\mathbf{k}$ are isomorphic over $\bar{\mathbf{k}}$. Thus, *all* the equations associated to rank-one Drinfeld modules are twists of each other.

Finally we end with a number of questions and projects:

(1) Extend Denis' results to the case of arbitrary $\mathbf{A}$.

(2) What is the geometry of the projective schemes associated to the homogeneous families? These results may very well depend on the base field that one chooses because of peculiarities of finite characteristic. (Dummit has shown that curves obtained from the homogeneous equation associated to the Carlitz module are geometrically irreducible **[Go7]**.)

(3) Are there analogue of the Fermat families associated to some classical motives? For instance, we have seen how close $C^{\otimes g}$ is to $\mathbb{Z}(g)$. See (b) at the end of the paper.

Notes added in proof. (a) In the new paper **[K2]** M. Kapranov explains how to generalize the L-series of §3 to regular rings $\mathbf{A}$ of arbitrary dimension. More precisely, let X be a smooth projective variety over $\mathbf{F}_r$. Let $X \supset X_{n-1} \supset X_{n-2} \supset \cdots \supset X_0$ be a complete flag of smooth irreducible subvarieties, $\dim X_i = i$, such that X_{i-1} is *ample* in X_i. In this set-up Kapranov uses an idea of Parshin to "localize at the flag" and obtain the completion $k_\infty = \mathbb{F}_r((t_1))\cdots((t_n))$, as well as the exponent space $S = (\overline{k}_\infty^*)^n \times \mathbb{Z}_p$. He then defines a zeta function and establishes it to be a family of *entire* n-dimensional power series in the sense of Theorem 3.12. Finally, the special values of this zeta function at negative integers are shown to be *integral* (i.e., in the affine ring $\mathbf{A}$ of $X - X_{n-1}$) and have good congruences at the closed points of $\mathbf{A}$.

(b) The theory described above for Fermat families associates equations to $C^{\otimes g}$ where $g \geq 1$. The wonderful old paper **[O1]** appears to give a method for associating an equation to $C^{\otimes -1}$. Indeed, in §2.8 of **[O1]** one learns how to associate to an additive polynomial its "adjoint" leading directly into the concept of the "adjoint of a Drinfeld module". The procedure is very simple: Let ϕ be a Drinfeld module over an $\mathbf{A}$-field L with adoint ϕ^* and let $a \in \mathbf{A}$. To obtain ϕ_a^* from ϕ_a one simply replaces the monomials $c_i^a \tau^i$ in ϕ_a with

$$(c_i^a)^{1/r^i} \tau^{-i}.$$

One thus obtains an object that actually lives over the perfection of L. (See also Remark 5.9 of **[Th5]** where this idea is used in a somewhat hidden form.) In the case of a rank-one Drinfeld module one sees that the v-adic character obtained from the v-adic Tate module of ϕ^* (obvious definition) is the *adjoint* of the character associated to ϕ. One now follows the procedure of our §5 to obtain a Fermat equation using $y^{r^{-d}}$ and C_a^*; of course one obtains an equation involving r^ith roots which are then eliminated by raising to an appropriate power of r. So, for instance, when $a = T$, we obtain the equation

$$\mathscr{F}_1^{C^{\otimes -1}}(a)(x, y, z) = T^r x^r + x y^{r-1} - z^p = 0.$$

This should be contrasted to the equation

$$\mathscr{F}_1(a)(x, y, z) = T x y^{r-1} + x^r - z^p = 0.$$

These adjoint Fermat equations have yet to be studied.

REFERENCES

[A1] G. Anderson, *t-Motives,* Duke Math. J. **53** (1986), 457–502.

[A2] ——, *Notes on t-modules,* Lectures given at the Institute for Advanced Studies.

[A3] ——, *On a question arising from complex multiplication theory,* Galois Representations and Arithmetic Algebraic Geometry, Adv. Stud. Pure Math., vol. 12, Academic Press, New York, 1987, pp. 221–234.

[A4] ——, *A two-dimensional analogue of Stickelberger's theorem*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Workshop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 51–73.

[A5] ——, *On Tate modules of formal t-modules*, Mathematical Research Notices, Duke Math. J. **2** (1993), 41–52.

[A6] ——, *Rank one elliptic A-modules, A-harmonic series and Akhiezer-Baker functions*, preprint.

[AT1] G. Anderson and D. Thakur, *Tensor powers of the Carlitz module and zeta values*, Ann. of Math. (2) **132** (1990), 159–191.

[B1] J.-F. Boutot and H. Carayol, *Uniformisation p-adique des courbes de Shimura: Les théorèmes de Čerednik et de Drinfeld*, Astérisque **196–197** (1991), 45–158.

[Br1] M. L. Brown, *Singular moduli and supersingular moduli of Drinfeld modules*, Invent. Math. **110** (1992), 419–439.

[C1] L. Carlitz, *On certain functions connected with polynomials in a Galois field*, Duke Math. J. **1** (1935), 137–168.

[C2] ——, *A class of polynomials*, Trans. Amer. Math. Soc. **43** (1938), 167–182.

[D1] P. Deligne, *Valeurs de fonctions L et périodes d'intégrales*, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 313–346.

[D2] ——, *Le groupe fondamental de la droite projective moins trois points*, Galois Groups over $\mathbb{Q}$ (Y. Ihara, K. Ribet, and J.-P. Serre, eds.), Math. Sci. Res. Inst. Publ., Springer-Verlag, Berlin, Heidelberg, and New York, 1989, pp. 79–297.

[DH1] P. Deligne and D. Husemöller, *Survey of Drinfeld modules*, Contemp. Math. **67** (1987), 25–91.

[DMOS] P. Deligne, J. Milne, A. Ogus, and K. Shih, *Hodge cycles, motives and Shimura varieties*, Lecture Notes in Math., vol. 900, Springer-Verlag, Berlin, Heidelberg, and New York, 1982.

[De1] L. Denis, *Hauteurs canoniques et modules de Drinfeld*, Math. Ann. **294** (1992), 213–223.

[De2] ——, *Le théorème de Fermat-Goss*, Trans. Amer. Math. Soc. (to appear).

[Dr1] V. G. Drinfeld, *Elliptic modules*, Mat. Sb. **94** (1974), 594–627; English transl. in Math. USSR-Sb. **23** (1976), 561–592.

[Dr2] ——, *Elliptic modules*. II, Mat. Sb. **102** (1974); English transl. in Math. USSR-Sb. **31** (1977), 159–170.

[Dr3] ——, *Commutative subrings of some noncommutative rings*, Funct. Anal. **11** (1977), 11–14.

[Dr4] ——, *Varieties of modules of F-sheaves*, Funct. Anal. **21** (1987), 23–41.

[Ge1] E.-U. Gekeler, *Quasi-periodic functions and Drinfeld modular forms*, Compositio Math. **69** (1989), 277–293.

[Ge2] ——, *De Rham cohomology for Drinfeld modules*, Sém. Théorie des Nombres, Paris 1988–1989, Birkhäuser, Basel and Boston, 1990, pp. 57–85.

[Ge3] ——, *On finite Drinfeld modules*, J. Algebra **141** (1991), 187–203.

[Ge4] ——, *On the de Rham isomorphism for Drinfeld modules*, J. Reine Angew. Math. **401** (1989), 188–208.

[Go1] D. Goss, *The arithmetic of function fields* 2: *The "cyclotomic" theory*, J. Algebra **81** (1983), 107–149.

[Go2] ——, *The theory of totally real function fields*, Contemp. Math., Part 2, vol. 55, Amer. Math. Soc., Providence RI, 1986, pp. 449–477.

[Go3] ——, *Analogies between global fields*, CMS Conf. Proc., vol. 7, Amer. Math. Soc., Providence, RI, 1987, pp. 83–114.

[Go4] ——, *L-series of Grössencharakters of type A_0 for function fields*, p-Adic Methods in Number Theory and Algebraic Geometry (A. Adolphson and M. Tretkoff, eds.), Contemp. Math., vol. 133, Amer. Math. Soc., Providence, RI, 1992, pp. 119–139.

[Go5] ——, *Report on transcendency in the theory of function fields*, Lecture Notes in Math., vol. 1383, Springer-Verlag, Berlin, Heidelberg, and New York, 1989, pp. 59–63.

[Go6] ——, *A formal Mellin transform in the theory of function fields*, Trans. Amer. Math. Soc. **327** (1991), 567–582.

[Go7] ——, *On a Fermat equation arising in the arithmetic theory of function fields*, Math. Ann. **261** (1982), 269–286.

[Go8] ——, *The Γ-function in the arithmetic of function fields*, Duke Math. J. **56** (1988), 163–191.

[Go9] ——, *Harmonic analysis and the flow of a Drinfeld module*, J. Algebra **146** (1992), 219–241.

[Go10] ——, *L-series of t-motives and Drinfeld modules*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Workshop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 313–402.

[Go11] ——, *Some integrals attached to modular forms in the theory of function fields*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Workshop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 227–251.

[H1] Y. Hamahata, *Tensor products of Drinfeld modules and v-adic representations*, Manuscripta Math. **79** (1993), 307–327.

[Ha1] D. Hayes, *Explicit class field theory in global function fields*, Studies in Algebra and Number Theory, Adv. Math., Suppl. Stud., vol. 6, Academic Press, 1980, pp. 173–217.

[Ha2] ——, *Explicit class field theory for rational function fields*, Trans. Amer. Math. Soc. **189** (1974), 77–91.

[Ha3] ——, *Hecke characters and Eisenstein reciprocity in function fields*, J. Number Theory **43** (1993), 251–292.

[He1] Y. Hellegouarch, *Modules de Drinfeld généralisés*, Approximations Diophantiennes et Nombres Transcendants (Luminy 1990), de Gruyter, Berlin and New York, 1992, pp. 123–164.

[K1] M. Kapranov, *On cuspidal divisors on the modular varieties of elliptic modules*, Math. USSR-Izv. **30** (1988), 533–547.

[K2] ——, *A higher-dimensional generalization of the Goss zeta-function*, J. Number Theory (to appear).

[Mi1] J. Milne, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, Automorphic Forms, Shimura Varieties, and L-Functions (Laurent Clozel and J. S. Milne, eds.), Vol. I, Academic Press, New York, 1990, pp. 283–414.

[Mu1] D. Mumford, *An algebro-geometric construction of commuting operators and of solutions to the Toda lattice equation, Korteweg–de Vries equation and related nonlinear equations*, Internat. Sympos. on Algebraic Geometry (Kyoto), 1977, pp. 115–153.

[O1] O. Ore, *On a special class of polynomials*, Trans. Amer. Math. Soc. **35** (1933), 559–584.

[P1] B. Poonen, *Local Height Functions and the Mordell-Weil Theorem for Drinfeld Modules* (preprint).

[S1] N. Schappacher, *Periods of Hecke characters*, Lecture Notes in Math., vol. 1301, Springer-Verlag, Berlin, Heidelberg, and New York, 1988.

[Se1] J.-P. Serre, *Facteurs locaux des fonctions zêta des variétés algébriques (définitions et conjectures)*, Sém. Delange-Pisot-Poitou, 11e année, 1969/70, pp. 19-01–19-15.

[Se2] ——, *Sur les représentations modulaires de degré* 2 *de* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Duke Math. J. **54** (1987), 179–230.

[T1] Y. Taguchi, *Semi-simplicity of the Galois representations attached to Drinfeld modules over fields of infinite characteristics*, J. Number Theory **44** (3) (1993), 292–314.

[T2] ——, *The duals of Drinfeld modules* (preprint).

[T3] ——, *A duality for finite t-modules* (preprint).

[T4] ——, *Semisimplicity of the Galois representations attached to Drinfeld modules over fields of "finite characteristics"*, Duke Math. J. **62** (3) (1991), 593–599.

[T5] ——, *On the π-adic theory–Galois cohomology*, Proc. Japan Acad. **68** (1992), 214–218.

[Ta1] T. Takahashi, *Good reduction of elliptic modules*, J. Math. Soc. Japan **34** (1982), 475–487.

[Te1] J. Teitelbaum, *The Poisson kernel for Drinfeld modular curves*, J. Amer. Math. Soc. **4** (1991), 491–511.

[Te2] ——, *Rigid analytic modular forms: An integral transform approach*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Workshop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 189–207.

[Th1] D. Thakur, *Number fields and function fields (zeta and gamma functions at all primes)*, Proc. Conf. p-Adic Analysis, Hengelhoef 1986 (N. De Grande-De Kimpe and L. Van Hamme, eds.), Univ. Brussels, Brussels, 1986, pp. 149–157 .

[Th2] ——, *Gauss sums for* $\mathbb{F}_q[T]$, Invent. Math. **94** (1988), 105–112.

[Th3] ——, *Zeta measure associated to* $\mathbb{F}_q[T]$, J. Number Theory **35** (1990), 1–17.

[Th4] ——, *Gamma functions for function fields and Drinfeld modules*, Ann. of Math. (2) **134** (1991), 25–64.

[Th5] ——, *Shtukas and Jacobi sums*, Invent. Math. **111** (1993), 557–570.

[Th6] ——, *On gamma functions for function fields*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Workshop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 75–86.

[Th7] ——, *Drinfeld modules and arithmetic in the function fields*, Mathematical Research Notices, Duke Math. J. **9** (1992), 185–197.

[Th8] ——, *Behaviour of function field Gauss sums at* ∞, Bull. London Math. Soc. (to appear).

[Thi1] A. Thiery, *Indépendance algébrique des périodes et quasi-périodes d'un module de Drinfeld*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Workshop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 265–284.

[W1] L. Wade, *Certain quantities transcendental over* $\mathrm{GF}(p^n, x)$, Duke Math. J. **8** (1941), 701–720.

[Y1] J. Yu, *Transcendence and Drinfeld modules*, Invent. Math. **83** (1986), 507–517.

[Y2] ——, *Transcendence and Drinfeld modules: Several variables*, Duke Math. J. **58** (1989), 559–575.

[Y3] ——, *Transcendence and special zeta values in characteristic* p, Ann. of Math. (2) **134** (1991), 1–23.

[Y4] ——, *On periods and quasi-periods of Drinfeld modules*, Compositio Math. **74** (1990), 235–245.

[Y5] ——, *Transcendence in finite characteristic*, The Arithmetic of Function Fields (D. Goss, D. Hayes, and M. Rosen, eds.), Proc. Worskhop Ohio State Univ., June 17–26, 1991, de Gruyter, Berlin and New York, 1992, pp. 253–264.

OHIO STATE UNIVERSITY, COLUMBUS, OHIO

E-mail address: `Internet: goss@coltrane.mps.ohio-state.edu`

Automorphic Forms and Shimura Varieties

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

The Local Langlands Correspondence: The Non-Archimedean Case

STEPHEN S. KUDLA

The goal of this purely expository article is to provide some local background for the theory of automorphic representations of $\mathrm{GL}(n)$. Specifically, we shall review some of what is known about the classification of irreducible algebraic (algebraic = smooth and admissible) representations of $\mathrm{GL}(n, F)$ where F is a non-Archimedean local field. The main goal here is the local Langlands conjecture, which gives a complete parameterization of the isomorphism classes of such representations in terms of equivalence classes of representations into $\mathrm{GL}(n, \mathbb{C})$ of the Weil-Deligne group W'_F of F. The parameterization is required to be compatible with L and ε-factors, twisting by characters, etc., cf. §4.2 below, and, in the case $n = 1$, it is essentially equivalent to local class field theory. The cases $n = 2$ (Kutzko **[39]**, Tunnell **[56]**) and $n = 3$ (Henniart **[23]**) have been known for some time.

The verification of the conjecture in general can be broken into two parts:

(i) Proof of the conjecture for supercuspidal representations of $\mathrm{GL}(n,F)$ and irreducible representations of the Weil group W_F of F.

(ii) Proof that the conjecture is "compatible with induction".

The second step was achieved by Bernstein and Zelevinski **[4]** as a consequence of their classification of all irreducible algebraic representations in terms of supercuspidal ones. A very clear exposition of their results together with a description of the proofs was given by Rodier **[48]**. We shall give a summary of this work below.

Step (i) has now been *completed* when $\mathrm{char}(F) = p > 0$ by Laumon, Rapoport and Stuhler **[43]**, extending global methods introduced by Drinfeld **[15, 16]**. Thus, the local Langlands conjecture is now a theorem in the function field case! In the case where F has characteristic zero much progress

1991 *Mathematics Subject Classification*. Primary 11S37, 11S40.
This paper is in final form and no version of it will be submitted for publication elsewhere.
Supported by National Science Foundation Grant DMS 9003109.

has been made. For example, a complete construction of the supercuspidal representations has been announced by Kutzko and Bushnell [**10**] and also by Corwin [**12**], but the actual construction of a bijection with all of the required properties has not yet been completed. On the other hand, some work of Henniart [**27**] gives a *characterization* of the local correspondence in terms of certain epsilon factors of pairs of representations.

This article is intended as an introduction to the subject for beginners. Much of it has been extracted from the excellent surveys of Rodier [**48**] and Henniart [**24**], and the reader is urged to consult these and [**28, 45**] for more detail. In particular, we have not discussed the extensive work which has been done on the explicit construction of supercuspidal representations [**30, 45, 10,** and **12**] nor have we mentioned the relations among irreducible admissible representations of $\mathrm{GL}(n)$ and those of the unit groups of other central simple algebras over F [**45, 14**].

The real situation is discussed in the companion article of Knapp [**37**]. For the relations with things global and motivic the reader should consult [**55, 42** and **11**]. The present article might be viewed as a footnote to the comments on page 21 of [**55**] and as background for [**42**] and [**11**]. Finally, the reader, who might well be misled by the simplicity of the theory for $\mathrm{GL}(n)$, especially the absence of L-packets, should consult Borel's survey in Corvallis [**7**] for a discussion of the local Langlands conjecture for arbitrary connected reductive groups.

We have made no attempt to explain the connections between the local results and conjectures described in this article and the global theory of automorphic forms. For this, the introductory articles of [**17, 3** and **1**] are highly recommended.

Our list of references is very far from complete. It is only intended as a quick guide, and the reader is refered to [**45**] and the bibliography of our other references for much more extensive information.

I would like to thank J. Cogdell, D. Goldberg, B. Gross, A. Knapp, M. Rapoport, D. Rohrlich, and F. Shahidi for helpful comments on the manuscript. I would like to thank F. Shahidi for his advice about L and ε-factors. Finally, I am endebted to B. Gross, M. Harris, and D. Ramakrishnan for their encouragement when it was most needed, and I would like to thank Uwe Jannsen for his editorial suggestions and patience.

0. Conventions

Throughout this article F will denote a complete non-Archimedean local field with residue characteristic p and with residue field $\mathbb{F}_q$, $q = p^t$. Also, $\mathcal{O}_F$ (resp., $\mathcal{P}$) will denote the ring of integers of F (resp., the maximal ideal of $\mathcal{O}_F$) and we shall fix a generator $\varpi \in \mathcal{P}$. We fix a separable closure $\overline{F}$ of F and we write W_F (resp., W_F') for the Weil group (resp., Weil-Deligne group) of F [**55**]. We write $G = G_n = \mathrm{GL}(n, F)$. A representation (π, V) of G on a complex vector space V is *smooth* if the stabilizer in G

of any vector $v \in V$ is an open subgroup of G. The representation (π, V) is *admissible* if the subspace V^K of vectors in V fixed by any compact open subgroup $K \subset G$ is finite dimensional. An *algebraic* representation is one that is both admissible and smooth (this is the terminology of **[4]**, for example). Jacquet **[31]** proved that any irreducible smooth representation is automatically admissible.

The *finite-dimensional* irreducible algebraic representations have the form $\pi(g) = \chi(\det(g))$ for a (quasi)character $\chi : F^\times \to \mathbb{C}^\times$. For any algebraic representation π and any character χ, we denote by $\pi(\chi)$ the twist of π by χ:

$$\pi(\chi)(g) = \chi(\det(g))\pi(g). \tag{0.1}$$

When $\chi(x) = |x|^s$ for a complex number s, we shall write $\pi(s) = \pi(|\ |^s)$.

Given any algebraic representation (π, V), $(\pi^\vee, V^\vee)$ will denote the representation of G on the set $V^\vee$ of smooth vectors (vectors whose stabilizer is open in G) in the full linear dual $V^* = \operatorname{Hom}(V, \mathbb{C})$. Note that if $V = \bigoplus_{\sigma \in \hat{K}} V(\sigma)$ is the decomposition of V into isotypic components for some compact open subgroup $K \subset G$, then the admissibility of V implies that the $V(\sigma)$'s are all finite dimensional (they are fixed by the kernel of σ, which has finite index in K). Thus, $V^\vee = \bigoplus_{\sigma \in \hat{K}} V(\sigma)^*$ is clearly admissible and hence, algebraic as well.

If π is an irreducible algebraic representation, then ω_π will denote its central character.

We fix the Borel subgroup B of G consisting of the upper triangular matrices and let U be its unipotent radical, i.e., the subgroup with 1's on the diagonal. We call any parabolic subgroup P of G with $P \supset B$ a *standard* parabolic subgroup.

1. Representations of $\mathrm{GL}(n, F)$; the Langlands classification

1.1. Induction, Jacquet functors. We shall be interested in describing the set $\mathscr{A}_F(n)$ of isomorphism classes of irreducible algebraic (or smooth) representations of $G_n = \mathrm{GL}(n, F)$. Here the $\mathscr{A}$ is supposed to remind us that these representations are connected with automorphic forms. It will actually be more convenient to consider the set

$$\mathscr{A}_F = \coprod_{n \geq 1} \mathscr{A}_F(n). \tag{1.1.1}$$

The main method of producing new representations from old is the process of parabolic induction. For any partition $n = n_1 + n_2 + \cdots + n_r$ of n there is a unique standard parabolic subgroup of $G = G_n$:

$$P = \left\{ \begin{pmatrix} a_1 & * & \dots & * \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & a_r \end{pmatrix} \in G \mid a_i \in \mathrm{GL}(n_i, F) \right\} = MN \tag{1.1.2}$$

where N is the unipotent radical and $M \simeq G_{n_1} \times \cdots \times G_{n_r}$ is a Levi component. If $\sigma = \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_r$ is a smooth representation of M on a vector space V, extend σ to a representation of P, trivial on N, and consider the induced representation
(1.1.3)
$$I_P^G(\sigma) = \{f : G \longrightarrow V \mid f \text{ smooth on } G \text{ and } f(nmg) = \delta_P^{\frac{1}{2}}(m)\sigma(m)f(g)\},$$
where

$$\delta_P(m) = |Ad_N m| \tag{1.1.4}$$

is the modulus character of the adjoint action of M on N, and the group G acts by right translations. The representation $I_P^G(\sigma)$ is admissible and of finite length if σ is admissible and of finite length **[4]** (i.e., if each of the σ_i's is admissible and of finite length). Moreover, the set of Jordan-Hölder constituents of $I_P^G(\sigma)$ depends only on the data (M, σ) up to conjugacy in G and not on the choice of parabolic subgroup P with M as Levi factor. The modulus character is included so that $I_P^G(\sigma)$ is unitarizable whenever σ is unitarizable. In particular, $I_P^G(\sigma)^\vee = I_P^G(\sigma^\vee)$.

The adjoint of induction is given by the *Jacquet functor*. For any smooth representation (π, W), Frobenius reciprocity gives

$$\mathrm{Hom}_G(\pi, I_P^G(\sigma)) \simeq \mathrm{Hom}_P(\pi|_P, \sigma \otimes \delta_P^{\frac{1}{2}}) \simeq \mathrm{Hom}_M(r_N(\pi), \sigma), \tag{1.1.5}$$

where

$$r_N(\pi) = W_N = W/\langle w - \pi(n)w \mid n \in N, w \in W\rangle \tag{1.1.6}$$

is the space of N co-invariants of π with the action of M given by

$$r_N(m) = (\text{natural action}) \cdot \delta_P^{-\frac{1}{2}}(m). \tag{1.1.7}$$

A fundamental fact **[4]** is that r_N is exact and that it carries algebraic representations of G of finite length to algebraic representations of M of finite length.

DEFINITION 1.1.1. An irreducible algebraic representation π is **supercuspidal** if $r_N(\pi) = 0$ for all proper standard parabolic subgroups $P = MN$ of G.

We let $\mathscr{A}_F^{\mathrm{sc}}(n)$ denote the set of isomorphism classes of irreducible admissible supercuspidal representations of $\mathrm{GL}(n, F)$.

Observe that π is supercuspidal precisely when $\mathrm{Hom}_G(\pi, I_P^G(\sigma)) = 0$ for all representations $I_P^G(\sigma)$ induced from proper parabolics. Since every Jordan-Hölder constituent of an $I_P^G(\sigma)$ actually occurs as a submodule of some $I_{P'}^G(\sigma')$ **[4,** Proposition 3.19], we see that the supercuspidal representations are precisely those that do not occur as constituents of any $I_P^G(\sigma)$, and thus, are not accessible by the process of parabolic induction. The supercuspidals can also be characterized by the fact that their matrix coefficients

$\phi_{v,v^\vee}(g) = \langle \pi(g)v, v^\vee \rangle$, $v \in \pi$, $v^\vee \in \pi^\vee$ are compactly supported modulo the center Z of G [**4**]. We shall return to this point below.

Finally, it should be noted that all of this discussion may be extended to the algebraic representations of the F points G of any (connected) reductive algebraic group over F.

1.2. The Langlands classification. We now discuss the Langlands classification following Rodier [**48**] and the methods of Bernstein and Zelevinski. The first result detects the reducibility of the induced representation $I_P^G(\sigma)$ when σ is supercuspidal.

THEOREM 1.2.1 ([**5, 59**]). *Let $\sigma = \sigma_1 \otimes \cdots \otimes \sigma_r$ be an irreducible algebraic representation of M with σ_i supercuspidal for all i. Then the representation $I_P^G(\sigma)$ is reducible if and only if for some pair of indices i, j with $i \neq j$, $n_i = n_j$ and $\sigma_i = \sigma_j(1)$.*

EXAMPLE. For the partition $n = \underbrace{m + m + \cdots + m}_{r \text{ times}}$ and for an irreducible supercuspidal representation σ of G_m, we call

$$(\sigma, \sigma(1), \sigma(2), \ldots, \sigma(r-1)) = [\sigma, \sigma(r-1)] = \Delta \tag{1.2.1}$$

a **segment**. The induced representation $I_P^G(\Delta)$ is reducible.

THEOREM 1.2.2. (i) *For any segment Δ the induced representation $I_P^G(\Delta)$ has a unique irreducible quotient $Q(\Delta)$ and a unique irreducible subrepresentation $Z(\Delta)$.*

(ii) *Moreover, $Q(\Delta)$ is essentially square integrable* [**5**], *i.e., there exists a character $\nu : F^\times \to \mathbb{R}_+^\times$ such that the integral*

$$\int_{Z\backslash G} |\phi_{v,\check{v}}(g)|^2 \nu(g)\, dg$$

is finite, for any matrix coefficient $\phi_{v,\check{v}}$ of $Q(\Delta)$.

The unique irreducible quotient $Q(\Delta)$ is called the **Langlands quotient**. The unique subrepresentation $Z(\Delta)$ can also be taken as a basis for the classification; this is nicely discussed in Rodier [**48**]. If ν can be taken to be 1, then $Q(\Delta)$ is *square integrable* and is said to lie in the *discrete series* for G. Moreover,

PROPOSITION 1.2.3 (I. N. Bernstein, [**48,** Proposition 11]). *Every square integrable representation π of G has the form $Q(\Delta)$, where*

$$\Delta = (\sigma, \sigma(1), \sigma(2), \ldots, \sigma(r-1))$$

with $\sigma(\frac{r-1}{2})$ unitary.

We shall write $\mathscr{A}_F^{\mathrm{ds}}(n)$ (resp., $\mathscr{A}_F^{\mathrm{ds}}$) for the subset of $\mathscr{A}_F(n)$ (resp., $\mathscr{A}_F$) consisting of the classes of the discrete series.

Recall that for any irreducible supercuspidal representation (σ, V) of G_m the matrix coefficients of σ are compactly supported modulo the center Z_m of G_m. If the central character ω_σ of σ is unitary, then the matrix coefficients $\phi_{v,\check{v}}$ lie in $L^2(Z_m\backslash G_m)$. In particular, for any fixed $\check{v} \in \check{V}$ with $\check{v} \neq 0$, the map $\sigma \hookrightarrow L^2(Z_m\backslash G_m)$ given by $v \mapsto \phi_{v,\check{v}}$ provides a unitary structure on σ. Thus, every supercuspidal with unitary central character lies in the discrete series. Moreover, for *any* supercuspidal σ there is some twist $\sigma(\nu)$ by a character $\nu : F^\times \to \mathbb{R}_+^\times$ whose central character is unitary. Thus, Proposition 1.2.3 implies that every essentially square-integrable representation π of G has the form $Q(\Delta)$, i.e., that the $Q(\Delta)$'s are precisely the essentially square-integrable representations of G.

EXAMPLE. If $n = 1 + \cdots + 1$ and $\sigma = |\ |^{\frac{1-n}{2}}$, then $P = B$ is the standard Borel subgroup, and

$$\Delta = (|\ |^{\frac{1-n}{2}}, |\ |^{\frac{3-n}{2}}, \ldots, |\ |^{\frac{n-1}{2}}). \tag{1.2.2}$$

Then, viewed as a character of B, $\Delta = \delta_P^{-\frac{1}{2}}$, and so

$$I_B^G(\Delta) = \{\text{smooth functions on } B\backslash G\} \tag{1.2.3}$$

with the action of G induced by the natural action of G on the flag variety $B\backslash G$. Thus $Z(\Delta) = \mathbb{1}$, the trivial representation of G on the constant functions, and $Q(\Delta)$ is the **Steinberg** representation or **special** representation. If $\Delta = [\sigma, \sigma(r-1)]$ for some irreducible supercuspidal representation σ of G_m, $n = rm$, then $Q(\Delta)$ is called a **generalized Steinberg** representation. The generalized Steinberg representations (for $r > 1$) are (essentially) square integrable but *not* supercuspidal.

Note that if $n > 1$ then the trivial representation $\mathbb{1}$ of G_n is not essentially square integrable. To obtain such a representation as a Langlands quotient we must induce again.

DEFINITION 1.2.4. Two segments

$$\Delta_1 = [\sigma_1, \sigma_1(r_1 - 1)] \quad \text{and} \quad \Delta_2 = [\sigma_2, \sigma_2(r_2 - 1)]$$

are said to be **linked** if $\Delta_1 \not\subset \Delta_2$, $\Delta_2 \not\subset \Delta_1$, and $\Delta_1 \cup \Delta_2$ is a segment. We say that Δ_1 **precedes** Δ_2 if Δ_1 and Δ_2 are linked and if $\sigma_2 = \sigma_1(k)$ for some positive integer k.

For example, the segments $(\sigma, \sigma(1), \sigma(2))$ and $(\sigma(1), \sigma(2), \sigma(3))$ are linked and the first precedes the second. On the other hand, the segments $(\sigma, \sigma(1), \sigma(2))$, and $(\sigma(1), \sigma(2))$ are not linked since the first covers the second. Similarly, no segment is linked with itself.

THEOREM 1.2.5 (Langlands classification). *Given segments $\Delta_1, \ldots, \Delta_r$, assume that for $i < j$, Δ_i does not precede Δ_j. Then* (a) *the induced representation $I_P^G(Q(\Delta_1) \otimes \cdots \otimes Q(\Delta_r))$ admits a unique irreducible quotient $Q(\Delta_1, \ldots, \Delta_r)$.*

(b) *If* $\Delta'_1, \ldots, \Delta'_s$ *is another collection of segments satisfying the "does not precede" condition, then* $Q(\Delta_1, \ldots, \Delta_r) \simeq Q(\Delta'_1, \ldots, \Delta'_s)$ *if and only if* $r = s$ *and* $\Delta'_i = \Delta_{\tau(i)}$ *for some permutation* τ.

(c) *Every irreducible admissible representation* π *of* G_n *is isomorphic to some* $Q(\Delta_1, \ldots, \Delta_r)$.

(d) *The induced representation* $I_P^G(Q(\Delta_1) \otimes \cdots \otimes Q(\Delta_r))$ *is irreducible if and only if no two of the segments* Δ_i *and* Δ_j *are linked.*

EXAMPLE.

$$\mathbb{1} \simeq Q(|\ |^{\frac{n-1}{2}}, |\ |^{\frac{n-3}{2}}, \ldots, |\ |^{\frac{1-n}{2}}). \tag{1.2.4}$$

Here the segments have length 1. If, on the other hand, $\pi = Q(\Delta)$ is an essentially square-integrable representation, then we take $P = G$ and the "second induction" is trivial.

2. Various examples

We now discuss how various important types of representations can be located in this classification.

2.1. Principal series representations, Satake transform. If we take the partition $n = 1 + \cdots + 1$ and characters $\chi_1, \ldots, \chi_n$ of $G_1 = F^\times$, such that $\chi_i\chi_j^{-1} \neq |\ |$ for all i and j (i.e., viewed as segments of length 1, the χ_i's are not linked), then $Q(\chi) := Q(\chi_1, \ldots, \chi_n) = I_B^G(\chi_1 \otimes \cdots \otimes \chi_n) = I(\chi)$ is an irreducible principal series representation. Its isomorphism class depends only on the set of $\chi_1, \ldots, \chi_n$ up to permutation (action of the Weyl group). If the χ_i's are all unramified, then the unramified principal series representation $I(\chi)$ contains a nonzero K-invariant vector, unique up to scalars, where $K = \mathrm{GL}(n, \mathcal{O}_F)$.

If we drop the condition that the χ_i's not be linked, then the induced representation still contains a unique K-fixed vector, and thus, $I(\chi)$ contains a unique "unramified" constituent, which we again denote by $Q(\chi)$. In fact, we may reorder the χ_i's so that the "does not precede" condition of Theorem 1.2.5 is satisfied. Then $Q(\chi) = Q(\chi_1, \ldots, \chi_n)$, i.e., the spherical constituent is precisely the Langlands quotient for this ordering. It is a fundamental fact that every representation with a K-fixed vector is isomorphic to some $Q(\chi)$. In fact, if $I = \{k \in K \mid \bar{k} \in \bar{B}\}$ where $\bar{k}$ denotes the image of k in $\mathrm{GL}(n, \mathbb{F}_q)$ and $\bar{B} \subset \mathrm{GL}(n, \mathbb{F}_q)$ denotes the subgroup of upper triangular matrices—the Iwahori subgroup—then every constituent of $I_B^G(\chi)$ has a nonzero I-fixed vector and every irreducible admissible representation of G with a nonzero I-fixed vector is a constituent of some $I_B^G(\chi)$ **[6]**. Note that almost all local components of an irreducible admissible representation of $\mathrm{GL}(n)$ over the adèles of a global field are unramified. More precisely, if $\mathbb{A}_L$ is the ring of adèles of a global field L, then every irreducible admissible representation π of $\mathrm{GL}(n, \mathbb{A}_L)$ can be written as a restricted tensor product

$\pi = \bigotimes'_v \pi_v$, where v runs through the places of L and π_v is an irreducible admissible representation of $\mathrm{GL}(n, L_v)$, for L_v the completion of L at v. Here "restricted" means that for almost all (non-Archimedean) places v, π_v is unramified, some choice of nonzero $K_v = \mathrm{GL}(n, \mathscr{O}_{L_v})$ fixed vector $x_v^0 \in \pi_v$ has been made, and every vector in the tensor product is a finite linear combination of vectors of the form $x = \bigotimes_v x_v$, where $x_v = x_v^0$ for almost all places v. Details may be found in Flath's article in the Corvallis Proceedings (*Decomposition of representations into tensor products*, Automorphic Forms, Representations and L-Functions, part 1, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979, pp. 179–183.

Note that an unramified representation is determined by the n-tuple of nonzero complex numbers $\chi_1(\varpi), \dots, \chi_n(\varpi)$, up to permutation. Thus, the set of such representations may be identified with the set of orbits under the Weyl group $W \simeq S_n$ in the maximal torus

$$\hat{T} = \{\mathrm{diag}(t_1, \dots, t_n) \mid t_i \in \mathbb{C}^\times\} \subset \mathrm{GL}(n, \mathbb{C}) =: \hat{G}, \tag{2.1.1}$$

via

$$Q(\chi) \leftrightarrow t_\chi = \mathrm{diag}(\chi_1(\varpi), \dots, \chi_n(\varpi)). \tag{2.1.2}$$

Here ${}^LG = \hat{G} \times W_F$ is the L**-group** of $G = \mathrm{GL}(n)$. The element t_χ is called the Satake parameter of $Q(\chi)$, cf. [**7**, §7].

Let $\mathscr{H}(G, K)$ be the Hecke algebra of (G, K), i.e., the algebra under convolution of bi-K-invariant functions on G of compact support. This algebra acts on any irreducible admissible representation π via

$$\pi(\phi)v = \int_G \phi(g)\pi(g)v\, dg. \tag{2.1.3}$$

If $v \in I(\chi)^K$ is a nonzero K-fixed vector in the unramified principal series $I(\chi)$, then v is an eigenvector for the action of $\mathscr{H}(G, K)$ with eigenfunctional $h \mapsto \lambda(h, \chi)$. If we let t_j, $j = 1, \dots, n$ denote the standard coordinate functions on the torus $\hat{T}$, then the map that sends an element $h \in \mathscr{H}(G, K)$ to is eigenvalue $\lambda(h, \cdot)$ induces the **Satake isomorphism**

$$\mathscr{H}(G, K) \xrightarrow{\sim} \mathbb{C}[t_1, t_1^{-1}, \dots, t_n, t_n^{-1}]^W \tag{2.1.4}$$

of the Hecke algebra with the algebra of Weyl invariant rational functions on $\hat{T}$.

2.2. Tempered representations. An irreducible algebraic representation π of G is *tempered* if it has a unitary central character and its matrix coefficients $\phi_{v,\check{v}}$ lie in $L^{2+\varepsilon}(Z\backslash G)$ for all $\varepsilon > 0$. The *essentially tempered* representations are those that are obtained from a tempered representation by twisting by a character.

PROPOSITION 2.2.1 (Jacquet [32]). *The tempered representations are precisely the*

$$Q(\Delta_1, \dots, \Delta_r) \simeq I_P^G(Q(\Delta_1) \otimes \cdots \otimes Q(\Delta_r)),$$

where $Q(\Delta_i)$ is square integrable (i.e., in $\mathscr{A}_F^{\mathrm{ds}}$) for all i.

Note that, by Proposition 1.2.3 above, if $Q(\Delta)$ is square integrable, then $\Delta = [\sigma(\frac{1-r}{2}), \sigma(\frac{r-1}{2})]$ for some supercuspidal σ with unitary central character. The set of central characters of such a segment looks, for example, like

$$(2.2.1) \qquad \begin{array}{ccccc} \times & \times & | & \times & \times \\ \omega(\sigma(-2)) & \omega(\sigma(-1)) & \omega(\sigma) & \omega(\sigma(1)) & \omega(\sigma(2)) \end{array}$$

(here $r = 5$). Two such segments cannot be linked. Thus, (d) of Theorem 1.2.5 applies and the tempered representations are irreducibly induced from discrete series representations. Moreover, if $\pi_1, \dots, \pi_m$ are tempered representations with $\pi_j = I_{P_j}^{G_{n_j}}(Q(\Delta_1^j), \dots, Q(\Delta_{r_j}^j))$, then by transitivity of induction we have that

(2.2.2)

$$I_P^G(\pi_1 \otimes \cdots \otimes \pi_m) \simeq I_{P_0}^G(Q(\Delta_1^1) \otimes \cdots \otimes Q(\Delta_{r_1}^1) \otimes \cdots \otimes Q(\Delta_1^m) \otimes \cdots \otimes Q(\Delta_{r_m}^m)),$$

where $P_0 \subset P$ is the standard parabolic subgroup determined by the π_i's. This representation is, in turn, irreducible and tempered.

More generally, if $x_1 \geq x_2 \geq \cdots \geq x_m$ are real numbers, we can consider

$$(2.2.3) \qquad \begin{aligned} \xi &= I_P^G(\pi_1(x_1) \otimes \cdots \otimes \pi_m(x_m)) \\ &\simeq I_{P_0}^G(Q(\Delta_1^1)(x_1) \otimes \cdots \otimes Q(\Delta_{r_1}^1)(x_1) \\ &\quad \otimes \cdots \otimes Q(\Delta_1^m)(x_m) \otimes \cdots \otimes Q(\Delta_{r_m}^m)(x_m)). \end{aligned}$$

Since the $Q(\Delta_j^i)$'s are all square integrable, none of the Δ_i^j's can be linked. After the twist, the $\Delta_i^k(x_k)$'s and $\Delta_j^k(x_k)$'s are still not linked, but the $\Delta_i^k(x_k)$'s and $\Delta_j^\ell(x_\ell)$ can become linked. However, the condition $x_1 \geq x_2 \geq \cdots \geq x_m$ forces the "does not precede" condition to be satisfied. Thus, ξ has a unique irreducible quotient, via (a) of Theorem 1.2.5. Also, observe that if $x_i = x_{i+1}$ for some i, then we can replace the pair of tempered representations $\pi_i(x_i)$, $\pi_{i+1}(x_{i+1})$ by the tempered representation $Q(\pi_i, \pi_{i+1})$ (= the full induced representation).

THEOREM 2.2.2.

(a) *Every irreducible admissible representation π can be written as a quotient of an induced representation of the form*

$$\xi = I_P^G(\pi_1(x_1) \otimes \cdots \otimes \pi_m(x_m))$$

where the π_i's are tempered representations and $x_1 > x_2 > \cdots > x_m$ are real numbers.

(b) *For a given* π, *the standard parabolic* P, *the tempered representations* $\pi_1, \dots, \pi_m$ *(up to isomorphism) and the real numbers* $x_1 > x_2 > \cdots > x_m$ *are unique.*

This version of the Langlands classification extends to arbitrary p-adic reductive groups **[8, 53]**. Unfortunately, for such groups we lack a nice and explicit description of the set of (isomorphism classes of) discrete series (square-integrable) representations. Specifically, we would like to have a recipe, like that involving the segments, etc., above, which begins with the supercuspidal representations as the basic building blocks (not necessarily explicitly known), as in Proposition 1.2.3., and "cooks up" all of the square-integrable representations. Note that the supercuspidal representations for $\mathrm{GL}(n)$ have now been completely constructed **[10, 12]**.

2.3. Generic representations. An irreducible algebraic representation (π, V) of G is **generic** if the following condition is satisfied: Let U be the unipotent radical of the Borel subgroup B, and, for some choice of a nontrivial additive character $\psi : F \to \mathbb{C}$, let $\theta : U \to \mathbb{C}$ be the character of U given by

$$\theta(u) = \psi(u_{1,2} + u_{2,3} + \cdots + u_{n-1,n}), \qquad u = (u_{ij}) \in \mathrm{GL}(n, F). \tag{2.3.1}$$

Then π is generic if there exists a nonzero linear functional $\lambda : V \to \mathbb{C}$ such that $\lambda(\pi(u)v) = \theta(u)\lambda(v)$ for all $u \in U$ and $v \in V$. The definition is independent of the choice of ψ, via the action of B by conjugation; this independence can fail for more general reductive groups.

THEOREM 2.3.1 (Zelevinski **[59**, Theorem 9.7**]**). *The representation* $\pi = Q(\Delta_1, \dots, \Delta_r)$ *is generic if and only if no two of the segments* Δ_i *are linked. In particular,*

$$\pi \simeq I_P^G(Q(\Delta_1) \otimes \cdots \otimes Q(\Delta_r))$$

for essentially square-integrable representations $Q(\Delta_i)$.

For example, the essentially square-integrable representations $Q(\Delta)$ are always generic (since $r = 1$, i.e., there is no second induction). More generally, every tempered representation is generic. At the other extreme are the representations $Q(\Delta(\frac{r-1}{2}), \dots, \Delta(\frac{1-r}{2}))$ like $\mathbb{1}$, which are never generic.

Generic representations play a fundamental role in the global theory, particularly in the theory of L-functions **[33, 50]**. For example, every local component of an irreducible admissible *cuspidal* representation $\pi = \bigotimes_v \pi_v$ of the adèle group $\mathrm{GL}(n, \mathbb{A}_L)$ of a global field L (see §2.1, above) is generic (cf. J. A. Shalika, *The multiplicity one theorem for* $\mathrm{GL}(n)$, Ann. of Math. **100** (1974), 171–193). This result fails for other groups, e.g., for the rank-two symplectic group $\mathrm{Sp}(2)$. On the other hand, the generalized Ramanujan conjecture claims that every local component of such a π is *tempered.* Once again, this is known to fail for $\mathrm{Sp}(2)$, cf. R. Howe and I. I. Piatetski-Shapiro, *A counterexample to the "Generalized Ramanujan Conjecture" for (quasi)-split* *groups*, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979, pp. 315–322. The present hope is that every local component of a *generic* irreducible admissible cuspidal representation is tempered [**33**].

2.4. Unitary representations. Next we note that since the local components of automorphic cuspidal representations will be unitary, it is important to identify the unitary representations in the set $\mathscr{A}_F$. For a general reductive group, this is a very difficult open problem. The case of $\mathrm{GL}(n)$ has been solved by Tadič [**54**] in the non-Archimedean case and by Vogan [**57**] in the Archimedean case.

Let $\mathscr{A}_F^u \subset \mathscr{A}_F$ denote the set of classes of irreducible unitary representations. Of course $\mathscr{A}_F^{\mathrm{ds}} \subset \mathscr{A}_F^u$. For any $\tau \in \mathscr{A}_F^{\mathrm{ds}}$ and any integer r, we can form the Langlands quotient

$$u(\tau, r) := Q\left(\tau\left(\frac{r-1}{2}\right), \tau\left(\frac{r-3}{2}\right), \dots, \tau\left(\frac{1-r}{2}\right)\right), \tag{2.4.1}$$

where we write τ instead of the segment Δ defining it. Also, for any $\alpha \in (0, \frac{1}{2})$, the induced representation

$$uc(\tau, r; \alpha) := I_P^G(u(\tau, r)(\alpha), u(\tau, r)(-\alpha)) \tag{2.4.2}$$

is irreducible and is called a **complementary series** representation. Note that this representation can be written as the Langlands quotient

$$\begin{aligned} uc(\tau, r; \alpha) &\\ = Q\Bigg(&\tau\left(\frac{r-1}{2}+\alpha\right), \dots, \tau\left(\frac{1-r}{2}+\alpha\right), \\ &\tau\left(\frac{r-1}{2}-\alpha\right), \dots, \tau\left(\frac{1-r}{2}-\alpha\right)\Bigg) \end{aligned} \tag{2.4.3}$$

since the condition $\alpha \in (0, \frac{1}{2})$ insures that the "does not precede" condition is satisfied.

THEOREM 2.4.1 (Tadič [**54**, Theorem D, p. 338)]. *Let $\mathscr{B} \subset \mathscr{A}_F$ denote the subset consisting of (the classes of) the $u(\tau, r)$'s and $uc(\tau, r; \alpha)$'s as τ runs over $\mathscr{A}_F^{\mathrm{ds}}$, $r \geq 1$, and $\alpha \in (0, \frac{1}{2})$. Then*

(a) *For any $\pi_1, \dots, \pi_m \in \mathscr{B}$, $I_P^G(\pi_1 \otimes \dots \otimes \pi_m) \in \mathscr{A}_F^u$. In particular, the $u(\tau, r)$'s and $uc(\tau, r; \alpha)$'s are unitary.*

(b) *For every $\pi \in \mathscr{A}_F^u$, there exist $\pi_1, \dots, \pi_m \in \mathscr{B}$ such that $\pi \simeq I_P^G(\pi_1 \otimes \dots \otimes \pi_m)$. Moreover, the collection $\pi_1, \dots, \pi_m$ is unique up to permutation.*

2.5. Example: $\mathrm{GL}(2, F)$

The essentially square-integrable representations will be the supercuspidals together with the $Q(\Delta)$'s where $\Delta = (\chi|\,|^{-\frac{1}{2}}, \chi|\,|^{\frac{1}{2}})$ for a character χ of $F^\times$. Note that $Q(\Delta) = Q((|\,|^{-\frac{1}{2}}, |\,|^{\frac{1}{2}}))(\chi)$, i.e., all such representations are twists of the Steinberg representation. The remaining representations are the $Q(\chi_1, \chi_2)$'s where the case $\chi_2 = \chi_1|\ |$ is excluded. By (c) of Theorem 1.2.5, these are irreducible if and only if $\chi_1 \neq \chi_2|\ |$, and in this case are the irreducible principal series representations. Finally, when $\chi_1 = \chi_2|\ |$, $Q(\chi_1, \chi_2)$ is the character $\chi_2|\ |^{\frac{1}{2}}$. By Theorem 2.3.1, these characters are the only nongeneric representations of $\mathrm{GL}(2, F)$. The tempered representations are the discrete series (supercuspidals and unitary twists of the Steinberg) and the irreducible unitary principal series $Q(\chi_1, \chi_2) = I_B^G(\chi_1, \chi_2)$, where χ_1 and χ_2 are unitary. The remaining nontempered unitary representations are the complementary series representations

$$uc(\chi, 2; \alpha) = Q(\chi|\ |^{\alpha}, \chi|\ |^{-\alpha}) = I_B^G(\chi|\ |^{\alpha}, \chi|\ |^{-\alpha}), \tag{2.5.1}$$

for a unitary character χ of $F^\times$ and $\alpha \in (0, \frac{1}{2})$, and the unitary characters

$$u(\chi, 2) = Q(\chi|\ |^{\frac{1}{2}}, \chi|\ |^{-\frac{1}{2}}) = \chi \circ \det, \tag{2.5.2}$$

which occur as quotients at the ends of the complementary series.

We leave the cases $n = 3$ and $n = 4$ as exercises for the interested reader.

3. L and ε-factors

Before discussing the local Langlands conjecture, we summarize a few facts about the local L and ε-factors attached to the elements of the set $\mathscr{A}_F$. Here we more or less follow the first section of **[25]**. We fix a nontrivial additive character $\psi : F \to \mathbb{C}^\times$.

If $\pi \in \mathscr{A}_F$, Jacquet and Godement **[21, 33]** define both an L-factor $L(s, \pi) = P_\pi(q^{-s})^{-1}$, for a certain polynomial $P_\pi(X) \in \mathbb{C}[X]$ with $P_\pi(0) = 1$, and an epsilon factor $\varepsilon(s, \pi, \psi)$, which is a monomial in q^{-s}. If $\pi = Q(\chi)$ is an unramified principal series representation with Satake parameter t_χ, then $L(s, \pi) = \det(1 - q^{-s}t_\chi)^{-1}$ and $\varepsilon(s, \pi, \psi) = 1$, provided we assume, as we now do, that ψ is trivial on $\mathscr{O}_F$ and nontrivial on $\varpi^{-1}\mathscr{O}_F$. Moreover, if π and $\pi' \in \mathscr{A}_F$, then factors $L(s, \pi \times \pi')$ and $\varepsilon(s, \pi \times \pi', \psi)$ were defined by Jacquet, Shalika, and Piatetski-Shapiro **[35]**. If $\pi = Q(\chi)$ and $\pi' = Q(\chi')$ are both unramified principal series, then $L(s, \pi \times \pi') = \det(1 - q^{-s}t_\chi \otimes t_{\chi'})^{-1}$ and $\varepsilon(s, \pi \times \pi', \psi) = 1$. In both cases these factors for arbitrary representations are defined directly in terms of certain families of integrals. The definition of Rankin-Selberg L and ε-factors can be deduced from the general theory developed in **[50]**. We mention in passing that for a global irreducible admissible representation $\pi = \bigotimes'_v \pi_v$ of $\mathrm{GL}(n, \mathbb{A}_L)$ (cf. §2.1), the global L-function and ε-factor are defined as products

$$L(s, \pi) = \prod_v L(s, \pi_v) \quad \text{and} \quad \varepsilon(s, \pi) = \prod_v \varepsilon(s, \pi_v, \psi_v),$$

where the latter does not depend on the choice of the family $\{\psi_v\}$ provided they arise as restrictions of a global additive character $\psi : \mathbb{A}_L/L \to \mathbb{C}^\times$. If π

is automorphic, then $L(s, \pi)$ converges in a half plane, has a meromorphic analytic continuation, and satisfies a functional equation

$$L(s, \pi) = \varepsilon(s, \pi)L(1 - s, \pi^\vee),$$

where $\pi^\vee$ is the contragredient of π [**33**]. Similar results hold for the global L-function and ε-factor for pairs [**35**].

We now summarize various results about the inductive relations satisfied by these factors.

3.1. $L(s, \pi)$ **and** $\varepsilon(s, \pi, \psi)$. If $n = 1$ then π is just a character of $F^\times$, and $L(s, \pi)$ and $\varepsilon(s, \pi, \psi)$ are the local L and ε-factors defined by Tate [**55**]. For general n, the L and ε-factors satisfy certain "multiplicative relations" with respect to induction. First, if $n > 1$ and if π is a supercuspidal representation, then $L(s, \pi) = 1$, while $\varepsilon(s, \pi, \psi)$ is given by a generalized Gauss sum [**9**]. Next, if $\pi = Q(\Delta)$, with $\Delta = [\sigma, \sigma(r-1)]$, is essentially square integrable, then [**25**, p. 153],

$$L(s, \pi) = L(s, \sigma(r-1)) = L(s + r - 1, \sigma), \tag{3.1.1}$$

and

$$\begin{aligned} \varepsilon(s, \pi, \psi) &= \prod_{i=0}^{r-1} \varepsilon(s, \sigma(i), \psi) \prod_{i=1}^{r-2} \frac{L(-s, \sigma^\vee(-i))}{L(s, \sigma(i))} \\ &= \prod_{i=0}^{r-1} \varepsilon(s + i, \sigma, \psi) \prod_{i=1}^{r-2} \frac{L(-s - i, \sigma^\vee)}{L(s + i, \sigma)}. \end{aligned} \tag{3.1.2}$$

Note that only the "top" twist in the segment contributes to the L-factor, while all of the twists in the segment contribute to the epsilon factor. Also, note that the ratios of L-factors occurring here are just monomials.

Next, if $\pi = Q(\tau_1, \dots, \tau_r)$ is an arbitrary Langlands quotient with essentially square-integrable τ_i's, we have

$$L(s, \pi) = \prod_{i=1}^{r} L(s, \tau_i) \tag{3.1.3}$$

and

$$\varepsilon(s, \pi, \psi) = \prod_{i=1}^{r} \varepsilon(s, \tau_i, \psi). \tag{3.1.4}$$

In passing, we note that the **conductor** of an irreducible admissible representation π can be defined as follows: write $\varepsilon(s, \pi, \psi) = Cq^{-ms}$ for some integer m. Here we take ψ as before. Then it is proved in [**34**] that (i) $m \geq 0$, (ii) $m = 0$ if and only if π is an unramified $Q(\chi)$, and hence, has a nonzero K-fixed vector, and (iii) when $m > 0$, the dimension of the space of vectors on V which are fixed by

(3.1.5)

$$K_n(m) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K \mid c \in M_{1,n-1}(\varpi^m \mathcal{O}_F),\ d \equiv 1 \bmod \varpi^m \mathcal{O}_F \right\}$$

is one. Moreover, m is the smallest integer t for which nonzero fixed vectors of $K_n(t)$ exist.

3.2. $L(s, \pi \times \pi')$ **and** $\varepsilon(s, \pi \times \pi', \psi)$. For pairs, we have the following "inductive" relations. First of all, $L(s, \pi \times \pi') = L(s, \pi' \times \pi)$ and similarly for ε-factors. If $\pi = Q(\tau_1, \dots, \tau_r)$ and $\pi' \in \mathscr{A}_F$ is arbitrary, then

$$L(s, \pi \times \pi') = \prod_{i=0}^{r} L(s, \tau_i \times \pi'), \tag{3.2.1}$$

and

$$\varepsilon(s, \pi \times \pi', \psi) = \prod_{i=0}^{r} \varepsilon(s, \tau_i \times \pi', \psi). \tag{3.2.2}$$

Next, if $\tau = Q(\Delta)$, $\Delta = [\sigma, \sigma(r-1)]$, and $\tau' = Q(\Delta')$, $\Delta' = [\sigma', \sigma'(r'-1)]$ are essentially square integrable with $r' \geq r$, then

$$L(s, \tau \times \tau') = \prod_{i=1}^{r} L(s + r + r' - 1 - i, \sigma \times \sigma') \tag{3.2.3}$$

and
(3.2.4)

$$\varepsilon(s, \tau \times \tau', \psi) = \prod_{i=1}^{r} \left(\prod_{j=0}^{r+r'-2i} \varepsilon(s + i + j - 1, \sigma \times \sigma', \psi) \right) \times \left(\prod_{j=0}^{r+r'-2i-1} \frac{L(-s - i - j + 1, \sigma^\vee \times (\sigma')^\vee)}{L(s + i + j - 1, \sigma \times \sigma')} \right).$$

Finally, if σ and σ' are supercuspidals, then

$$L(s, \sigma \times \sigma') = \prod_{\chi} L(s, \chi), \tag{3.2.5}$$

where χ runs over characters of $F^\times$ such that $(\sigma')^\vee(\chi) = \sigma$. Here we should not expect a simple formula for $\varepsilon(s, \sigma \times \sigma', \psi)$, see Remark 4.2.5 below.

4. The local Langlands conjecture

The Weil group W_F comes equipped with an isomorphism $r_F : F^\times \xrightarrow{\sim} W_F^{\mathrm{ab}}$ and with a continuous homomorphism $\varphi : W_F \to \mathrm{Gal}(\overline{F}/F)$ with dense image. The composition

$$F^\times \xrightarrow{r_F} W_F^{\mathrm{ab}} \xrightarrow{\varphi} \mathrm{Gal}(\overline{F}/F)^{\mathrm{ab}} \tag{4.0.1}$$

is required to be the reciprocity homomorphism of local class field theory. Here we follow the convention of Deligne [13] and send a uniformizer $\varpi \in F^\times$ to a geometric Frobenius in $\mathrm{Gal}(\overline{F}/F)^{\mathrm{ab}}$. Recall that a geometric Frobenius in $\mathrm{Gal}(\overline{F}/F)^{\mathrm{ab}}$ is an element Φ that induces the *inverse* of the map $x \mapsto x^q$ on the residue field $\mathbb{F}_q$ of F. In particular, the continuous complex characters of W_F, i.e., the continuous representations of W_F into $\mathrm{GL}(1, \mathbb{C})$, may be identified via r_F with the irreducible admissible representations of $F^\times = \mathrm{GL}(1, F)$. The local Langlands conjecture extends this correspondence to one between $\mathscr{A}_F(n)$ and the set of isomorphism classes of certain representations of degree n of the Weil-Deligne group W_F'.

4.1. Representations of the Weil-Deligne group. As explained in [**55**], a complex representation of the Weil-Deligne group W_F' of F is equivalent to a collection (ρ, N, V) consisting of a continuous representation $\rho : W_F \to \mathrm{GL}(V)$, $\dim_{\mathbb{C}}(V) = n$, together with a nilpotent endomorphism $N \in \mathrm{End}(V)$ such that $\rho(w)N\rho(w)^{-1} = \|w\|N$. A representation $\rho' = (\rho, N, V)$ will be called **admissible** (or Φ-semisimple) if the representation ρ of W_F is semisimple (cf. [**55**, (4.1.3)]). For any $n \geq 1$ the representation $\mathrm{sp}(n)$ is defined by

$$V = \mathbb{C}^n = \mathbb{C}e_0 + \mathbb{C}e_1 + \cdots + \mathbb{C}e_{n-1} \tag{4.1.1}$$

with

$$\rho(w)e_i = \|w\|^i e_i \quad \text{and} \quad Ne_i = e_{i+1} \tag{4.1.2}$$

for $0 \leq i < n-1$ and $Ne_{n-1} = 0$. A representation (ρ, N, V) is irreducible if and only if $N = 0$ and ρ is irreducible. An admissible representation (ρ, N, V) is indecomposable if and only if it has the form $\rho_0 \otimes \mathrm{sp}(n)$ with ρ_0 irreducible [**55**, p. 20]. To any admissible representation $\rho' = (\rho, N, V)$, there are associated L and ε-factors defined by

$$L(s, \rho') = \det(1 - q^{-s}\Phi|V_N^I)^{-1} \tag{4.1.3}$$

and

$$\varepsilon(s, \rho', \psi) = \varepsilon(s, \rho, \psi)\det(-\Phi|V^I/V_N^I), \tag{4.1.4}$$

where $V_N = \ker(N)$ and V_N^I is the space of invariants in V_N for the action of the inertia group I. Here Φ is a geometric Frobenius, as in [**55**, p. 19] and $\varepsilon(s, \rho, \psi)$ is the epsilon factor attached to the representation ρ of W_F. Note that $\varepsilon(s, \rho, \psi)$ is the factor $\varepsilon(s, \rho, \psi, dx)$ for Deligne and Tate, where dx is the additive Haar measure on F which is taken to be self-dual with respect to the Fourier transform defined by the pairing $\langle x, y\rangle = \psi(xy)$. The reader is referred to Tate's survey and [**13**] for more details.

It may be useful to note that, since we are only considering complex representations, we could just as well replace the group W_F' by the group $W_F \times \mathrm{SL}(2, \mathbb{C})$ and its continuous complex semisimple representations. If η is such a representation, we associate the representation $\rho' = (\rho, N)$ with

$$\rho(w) = \eta\left(w, \begin{pmatrix} \|w\|^{\frac{1}{2}} & 0 \\ 0 & \|w\|^{-\frac{1}{2}} \end{pmatrix}\right), \tag{4.1.5}$$

and N defined by

(4.1.6) $$\exp(N) = \eta\left(1, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right).$$

The fact that every possible nilpotent endomorphism N arises in this way and that the equivalence classes of complex representations of W_F' and of $W_F \times \mathrm{SL}(2, \mathbb{C})$ are in bijection is a consequence of the Jacobson-Morozov theorem **[38]**.

The situation for ℓ-adic representations is more subtle and is discussed in **[13,** §8] and **[55,** (4.2)].

4.2. The conjecture. Let $\mathcal{G}_F(n)$ denote the set of isomorphism classes of admissible representations of W_F' of degree n. Here two representations into $\mathrm{GL}(n, \mathbb{C})$ are said to be equivalent if they are conjugate by an element of $\mathrm{GL}(n, \mathbb{C})$. We identify the element of the set $\mathcal{G}_F(1)$, the characters of W_F, with characters of $F^\times$ via the reciprocity isomorphism $r_F : W_F^{\mathrm{ab}} \xrightarrow{\sim} F^\times$.

THE LOCAL LANGLANDS CONJECTURE. *For each $n \geq 1$ there is a canonical bijection*

$$\pi_F : \mathcal{G}_F(n) \to \mathcal{A}_F(n)$$
$$\rho' \mapsto \pi(\rho')$$

such that

(i) $\pi_F(\rho'(\chi)) = \pi_F(\rho')(\chi)$ *for any character χ of $F^\times$.*
(ii) $\det(\rho')$ *corresponds to $\omega_{\pi_F(\rho')}$, the central character of $\pi_F(\rho')$ via the isomorphism of local class field theory.*
(iii) $(\pi_F(\rho'))^\vee = \pi_F((\rho')^\vee)$.
(iv) $L(s, \rho_1' \otimes \rho_2') = L(s, \pi_F(\rho_1') \times \pi_F(\rho_2'))$.
(v) $\varepsilon(s, \rho_1' \otimes \rho_2', \psi) = \varepsilon(s, \pi_F(\rho_1') \times \pi_F(\rho_2'), \psi)$.
(vi) π_F *preserves conductors.*
(vii) *If F is a finite Galois extension of a field F_0, then π_F is compatible with the natural actions of* $\mathrm{Gal}(F/F_0)$ *on $\mathcal{G}_F$ and $\mathcal{A}_F$.*

Of course, when $n = 1$, condition (ii) forces the conjectured bijection to be just the one given by the isomorphism of local class field theory. In (vi), as explained in **[24,** p. 117], the conductors are the Artin conductor of ρ **[55]** and the conductor defined by the ε-factor of π in **[34]**. We will explain the word "canonical" in a moment.

The most basic case is that of unramified representations. If $\rho \in \mathcal{G}_F(n)$ is an unramified representation of W_F, i.e., trivial on the inertia group I_F and with $N = 0$, then ρ is determined by $\rho(\Phi)$, the image of the geometric Frobenius. Up to conjugacy, we may assume that $\rho(\Phi) \in {}^L T$, and we define an n-tuple of unramified characters $\chi = (\chi_1, \dots, \chi_n)$ and an unramified principal series representation $Q(\chi)$ by setting $\rho(\Phi) = t_\chi$. Then $\pi_F(\rho) = Q(\chi)$.

The results of Bernstein and Zelevinski and the facts about local factors discussed above imply that the local Langlands conjecture is compatible with induction. More precisely, assume that a bijection $\rho \mapsto \pi_F(\rho)$ from the set $\mathcal{G}_F^0$ of isomorphism classes of *irreducible* representations of W_F to the set of irreducible admissible *supercuspidal* representations $\mathcal{A}_F^{sc}$ is given. To any *indecomposable* representation $\rho \otimes \mathrm{sp}(r)$ of W_F' associate the essentially square-integrable representation $Q(\Delta)$ with $\Delta = [\pi_F(\rho), \pi_F(\rho)(r-1)]$. Finally, to any admissible representation $\rho' = (\rho_1 \otimes \mathrm{sp}(r_1)) \oplus \cdots \oplus (\rho_m \otimes \mathrm{sp}(r_m))$ of W_F' associate the Langlands quotient $Q(\Delta_1, \dots, \Delta_m)$ where $\Delta_i = [\pi_F(\rho_i), \pi_F(\rho_i)(r_i - 1)]$. Then Theorem 1.2.5 and the results about L and ε-factors described above yield

THEOREM 4.2.1 (Bernstein-Zelevinski, Rodier [**48,** Théorème 4]). *If the bijection $\rho \mapsto \pi_F(\rho)$ from $\mathcal{G}_F^0$ to $\mathcal{A}_F^{sc}$ satisfies conditions* (i), (ii), (iii), (iv), *and* (v) *of the local Langlands conjecture, then its extension to $\mathcal{G}_F \to \mathcal{A}_F$ gives a bijection satisfying* (i), (ii), (iii), (iv), *and* (v).

The compatibility with the full range of conditions (i)–(vii) will be discussed in a forthcoming paper of Henniart [**22**].

Actually, the conjectured bijection is characterized by its compatibility with the epsilon factors for pairs. This has recently been proved by Henniart [**27**]. The gamma factor for a pair of representations π and $\pi' \in \mathcal{A}_F$ is defined by

$$\gamma(s, \pi \times \pi', \psi) = \varepsilon(s, \pi \times \pi', \psi)\frac{L(1-s, \pi^\vee \times (\pi')^\vee)}{L(s, \pi \times \pi')}. \tag{4.2.1}$$

This factor occurs in the local functional equation of a family of zeta integrals of Rankin-Selberg type [**35**] and also in [**49**]. Here recall that we have fixed the Haar measure on F that is self-dual with respect to ψ.

THEOREM 4.2.2 (Henniart [**27**]). *For an integer $n \geq 2$, suppose that π and π' are generic irreducible admissible representations of* $\mathrm{GL}(n, F)$. *Suppose that, for all irreducible, admissible generic representations τ of* $\mathrm{GL}(n-1, F)$,

$$\gamma(s, \pi \times \tau, \psi) = \gamma(s, \pi' \times \tau, \psi).$$

Then $\pi \simeq \pi'$.

Recall that every supercuspidal representation is generic.

COROLLARY 4.2.3 (Henniart [**27**]). *For an integer $n \geq 2$, suppose that π and π' are supercuspidal irreducible admissible representations of* $\mathrm{GL}(n, F)$. *Suppose that, for all integers r with $1 \leq r \leq n-1$ and all irreducible admissible supercuspidal representations τ of* $\mathrm{GL}(r, F)$,

$$\varepsilon(s, \pi \times \tau, \psi) = \varepsilon(s, \pi' \times \tau, \psi).$$

Then $\pi \simeq \pi'$.

Jacquet has suggested that it should be possible to replace $n-1$ by $[\frac{n}{2}]$, the integer part of $\frac{n}{2}$.

COROLLARY 4.2.4 (Henniart [27]). *Let $n \geq 2$ be a given integer. Suppose that for each integer r with $1 \leq r \leq n-1$ there is a given surjective map θ_r from the set of isomorphism classes of irreducible complex continuous representations of W_F of degree r onto the set of isomorphism classes of irreducible admissible supercuspidal representations of* $\mathrm{GL}(r, F)$. *Suppose that σ is an irreducible complex continuous representation of W_F of degree n. Then there exists (up to isomorphism) at most one irreducible admissible supercuspidal representation π of* $\mathrm{GL}(n, F)$ *such that*

$$\varepsilon(\pi \times \theta_r(\rho), s, \psi) = \varepsilon(\sigma \otimes \rho, s, \psi)$$

for all r with $1 \leq r \leq n-1$ and all irreducible complex continuous representations of W_F of degree r.

Thus the local Langlands correspondence for supercuspidal representations is completely characterized by surjectivity, property (v), and its agreement with class field theory for $n = 1$. This is the sense in which it is "canonical".

REMARK 4.2.5. According to (v) of the local Langlands conjecture, the epsilon factor $\varepsilon(s, \sigma \times \sigma', \psi)$ attached to a pair of supercuspidal representations $\sigma = \pi_F(\rho)$ and $\sigma' = \pi_F(\rho')$ should be $\varepsilon(s, \rho \otimes \rho', \psi)$, the epsilon factor associated to the tensor product of the irreducible representations ρ and ρ' of W_F. This factor will thus depend on the decomposition of the tensor product $\rho \otimes \rho'$ into irreducibles, and hence, we do not expect a simple formula. The L-factor $L(s, \sigma \times \sigma')$, on the other hand, should be equal to $L(s, \rho \otimes \rho')$ by (iv), and hence, should depend only on the one-dimensional irreducible components of $\rho \otimes \rho'$. This is compatible with (3.2.5).

5. Examples

We now return to the various examples discussed above and discuss their Langlands parameters, i.e., the corresponding representations $\rho : W_F' \longrightarrow \mathrm{GL}(n, \mathbb{C})$.

5.1. Principal series. Suppose that $\chi_1, \ldots, \chi_n$ are (quasi-)characters of $F^\times$, and let $\rho = \chi_1 \oplus \cdots \oplus \chi_n$. Then, as mentioned above, we can assume that the χ_i's are ordered so that the "does not precede" condition is satisfied and we have $\pi_F(\rho) = Q(\chi_1, \ldots, \chi_n)$. If $\chi_i\chi_j^{-1} \neq |\ |$ for all i, j, then $\pi_F(\rho) = I(\chi)$ is an irreducible principal series. Otherwise, $I(\chi)$ has other constituents. For example, if our collection of characters is $\chi(n-1)$, $\chi(n-2), \ldots, \chi$, then the constituents of $I(\chi)$ occur with multiplicity one and have Langlands parameters obtained by introducing N's in all possible ways. For example, for GL(4) the representation $I_B^G(\chi(3), \chi(2), \chi(1), \chi)$

has 8 constituents and their Langlands parameters and L-factors are

$\chi(3)\oplus\chi(2)\oplus\chi(1)\oplus\chi$	$\prod_{i=0}^{3} L(s+i,\chi)$
$(\chi(2)\otimes \mathrm{sp}(2))\oplus\chi(1)\oplus\chi$	$L(s+3,\chi)L(s+1,\chi)L(s,\chi)$
$\chi(3)\oplus(\chi(1)\otimes \mathrm{sp}(2))\oplus\chi$	$L(s+3,\chi)L(s+2,\chi)L(s,\chi)$
$\chi(3)\oplus\chi(2)\oplus(\chi\otimes \mathrm{sp}(2))$	$L(s+3,\chi)L(s+2,\chi)L(s+1,\chi)$
$(\chi(2)\otimes \mathrm{sp}(2))\oplus(\chi\otimes \mathrm{sp}(2))$	$L(s+3,\chi)L(s+1,\chi)$
$(\chi(1)\otimes \mathrm{sp}(3))\oplus\chi$	$L(s+3,\chi)L(s,\chi)$
$\chi(3)\oplus(\chi\otimes \mathrm{sp}(3))$	$L(s+3,\chi)L(s+2,\chi)$

and

$\chi\otimes \mathrm{sp}(4)$	$L(s+3,\chi)$.

The last of these is essentially square integrable, hence tempered and generic, and is the only such constituent. If χ is unramified, then the first parameter on the list corresponds to the unique spherical (unramified) constituent. This result has been extended to arbitrary split groups by Rodier **[46]**.

In general the constituents of $I_P^G(Q(\Delta_1)\otimes\cdots\otimes Q(\Delta_r))$ have been described by Zelevinski **[59,** Theorem 7.1], **[48,** Théorème 5 and Remark at the end of §5.3].

THEOREM 5.1.1. *The Langlands quotient* $Q(\Delta_1',\ldots,\Delta_r')$ *occurs as a constituent of the induced representation* $I_P^G(Q(\Delta_1)\otimes\cdots\otimes Q(\Delta_t))$ *if and only if the collection of segments* $\{\Delta_1',\ldots,\Delta_r'\}$ *(unordered but counted with multiplicity) can be obtained from the collection* $\{\Delta_1,\ldots,\Delta_t\}$ *by the elementary operations of replacing linked segments* Δ *and* Δ' *by the segments* $\Delta\cup\Delta'$ *and* $\Delta\cap\Delta'$.

For example, if $\sigma\in\mathscr{A}_F^{\mathrm{sc}}(m)$, then the induced repesentation $I_P^G(\sigma(2)\otimes\sigma(1)\otimes\sigma)$ of $\mathrm{GL}(3m)$ has constituents

$Q([\sigma(2)],[\sigma(1)],[\sigma])$	(the Langlands quotient)
$Q([\sigma(1),\sigma(2)],[\sigma])$	
$Q([\sigma(2)],[\sigma(1),\sigma])$	
$Q([\sigma,\sigma(1),\sigma(2)])$	(essentially square integrable and generic).

Zelevinski has also given a conjecture, analogous to the Kazhdan-Lusztig conjecture, which describes the multiplicities of the constituents **[60, 48,** §5.4].

5.2. Tempered and generic representations. The supercuspidal representations in $\mathscr{A}_F(n)$ correspond (under the local Langlands conjecture) to the irreducible representations of W_F (and W_F'), while the essentially square-integrable representations correspond to the indecomposables $\rho\otimes\mathrm{sp}(r)$ for ρ

irreducible. These will be square integrable, and hence, unitary if the character $\det(\rho)|\ |^{\frac{r-1}{2}}$ is unitary. The tempered representations correspond to direct sums of such square-integrable parameters. In fact, one checks that

LEMMA 5.2.1. (i) *$\pi = \pi_F(\rho)$ is essentially square integrable if and only if the image $\rho(W_F')$ is not contained in any proper Levi subgroup of* $\mathrm{GL}(n, \mathbb{C})$.

(ii) *Suppose that $\pi = \pi_F(\rho)$, and let η be the representation of $W_F \times$* $\mathrm{SL}(2, \mathbb{C})$ *associated to ρ as explained at the end of* §4.1 *above. Then π is tempered if and only if the image $\eta(W_F)$ is bounded.*

These conditions are expected to hold for general reductive groups, cf. [**7**, 10.3(3) and (4)]. Note that in stating (ii), Borel considers only representations with $N = 0$, so that $\rho = \eta$. As an example, observe that the Steinberg representation $\pi = \pi_F(\mathrm{sp}(n))(\frac{1-n}{2})$ (the twist here yields a unitary central character) has $\eta(W_F) = 1$.

Finally, the generic representations have Langlands parameters of the form $(\rho_1 \otimes \mathrm{sp}(r_1)) \oplus \cdots \oplus (\rho_t \otimes \mathrm{sp}(r_t))$, where no two "segments" are "linked". In fact, one has the following result which has been observed by many people.

PROPOSITION 5.2.2 (Gross, D. Prasad, Rallis). *Assume the LLC. Then $\pi = \pi_F(\rho)$ is generic if and only if $L(s, \pi, Ad) = L(s, Ad \circ \rho)$ has no pole at $s = 1$.*

Here $Ad\colon \mathrm{GL}(n, \mathbb{C}) \to \mathrm{GL}(n^2, \mathbb{C})$ is the adjoint representation of $\mathrm{GL}(n, \mathbb{C})$. The proof is an amusing exercise using the classification of irreducible admissible representations and the information about L and ε-factors for pairs given above. Note that $L(s, Ad \circ \rho) = L(s, \rho \otimes \rho^\vee)$.

5.3. Functoriality. The local Langlands conjecture (LLC) implies that various natural operations on the set $\mathcal{G}_F = \coprod \mathcal{G}_F(n)$ of (equivalence classes of) representations of W_F' have corresponding operations on the set (of isomorphism classes of) irreducible admissible representations $\mathcal{A}_F = \coprod \mathcal{A}_F(n)$. We now discuss these operations in an ideal world, i.e., assuming the LLC. The actual proof of the LLC, e.g., in characteristic 0, may require that many of these operations be constructed and characterized independently of the parameterization!

First of all, if $r : \mathrm{GL}(n, \mathbb{C}) \to \mathrm{GL}(m, \mathbb{C})$ is any analytic representation, there is a natural map:

$$\mathcal{G}_F(n) \to \mathcal{G}_F(m), \qquad \rho \mapsto r \circ \rho, \tag{5.3.1}$$

hence, by LLC, a corresponding map

$$\mathcal{A}_F(n) \to \mathcal{A}_F(m), \qquad \pi = \pi_F(\rho) \mapsto \Pi = \pi_F(r \circ \rho). \tag{5.3.2}$$

For example, the map $\mathcal{A}_F(n) \to \mathcal{A}_F(1)$ given by $\pi \mapsto \omega_\pi$, the central character of π, is associated to $r = \det$. The case of $r : \mathrm{GL}(2, \mathbb{C}) \to \mathrm{GL}(3, \mathbb{C})$ given by the adjoint representation of $\mathrm{GL}(2, \mathbb{C})$ on the Lie algebra of $\mathrm{SL}(2, \mathbb{C})$ was considered by Gelbart and Jacquet [**18**]. Note that for each such r one can attach to $\pi = \pi_F(\rho)$ the L and ε-factors $L(s, \pi, r) := L(s, r\circ\rho)$ and $\varepsilon(s, \pi, r, \psi) := \varepsilon(s, r\circ\rho, \psi)$. For the unramified principal series representation $\pi = Q(\chi)$ with Satake parameter t_χ, we then have

$$L(s, \pi, r) = \det(1 - q^{-s}r(t_\chi))^{-1}. \tag{5.3.3}$$

It is a fundamental problem to define $L(s, \pi, r)$ and $\varepsilon(s, \pi, r, \psi)$ *directly* for all π, cf. **[47, 50]**.

Given any finite extension E/F ($E \subset \overline{F}$) of degree $|E : F| = d$, we have $W_E \subset W_F$ and natural operations

$$\mathrm{ind}_{W'_E}^{W'_F} : \mathscr{G}_E(m) \longrightarrow \mathscr{G}_F(dm) \tag{5.3.4}$$

and

$$\mathrm{res}_{W'_E}^{W'_F} : \mathscr{G}_F(n) \longrightarrow \mathscr{G}_E(n) \tag{5.3.5}$$

of induction and restriction. Here some comment about N might be useful. The restriction of a representation (ρ, N) of W'_F to W'_E is just $(\mathrm{res}_E^F(\rho), N)$. On the other hand, if (ρ, N) is a representation of W'_E on a vector space V, then we define $\widetilde{N}$ on the space

(5.3.6)

$$\mathrm{ind}_{W_E}^{W_F}(\rho) = \{f : W_F \longrightarrow V \mid f(hg) = \rho(h)f(g) \text{ for } g \in W_F \text{ and } h \in W_E\}$$

by

$$(\widetilde{N}f)(g) = ||g||N(f(g)). \tag{5.3.7}$$

We then have $\mathrm{ind}_{W'_E}^{W'_F}((\rho, N)) = (\mathrm{ind}_{W_E}^{W_F}(\rho), \widetilde{N})$.

5.4. Induction. From (5.3.4) and the LLC we should have a map

$$\mathscr{A}_E(m) \longrightarrow \mathscr{A}_F(dm), \qquad \pi = \pi_E(\eta) \mapsto \Pi = \pi_F(\mathrm{ind}_{W'_E}^{W'_F}(\eta)). \tag{5.4.1}$$

For example, if η is a character of $E^\times$, then $\rho = \mathrm{ind}_{W_E}^{W_F}(\eta)$ is *monomial* (cf. **[55,** p. 10]) of degree d. In fact, as explained in **[55,** pp. 10–11], every irreducible representation ρ of W_F of degree n with $(n, p) = 1$ is monomial. Thus, the supercuspidals of $\mathrm{GL}(n, F)$ for such n should all be associated to characters of field extensions of degree n.

Suppose, for example, that E/F is cyclic and that σ is a generator of $\mathrm{Gal}(E/F)$. The character η is said to be *regular* if $\eta \not\simeq \eta^\sigma$. When η is regular the induced representation ρ is irreducible and the associated representation $\pi_F(\rho)$ of $\mathrm{GL}(n, F)$ (under LLC) is supercuspidal. The set $\mathscr{G}_F^0(n)^\omega$ of irreducible representations that arise in this way is characterized by the condition $\rho \otimes \omega \simeq \rho$, where $\omega = \omega_{E/F}$ is the character of $F^\times$ corresponding to E via local class field theory. Moreover, assuming the LLC, the set of supercuspidal representations that arise in this way can be characterized by the condition that $\pi \otimes \omega_{E/F} \simeq \pi$. Let $\mathscr{A}_F^{\mathrm{sc}}(n)^{\omega}$ be the set of such supercuspidals. The existence of such a map $\mathscr{G}_E(1)^{\mathrm{reg}} \to \mathscr{A}_F^{\mathrm{sc}}(n)^{\omega}$ was proved by Kazhdan [36] via the trace formula. Let $\mathscr{G}_F^0(n)^{\mathrm{cyclic}} = \bigcup_{\omega} \mathscr{G}_F^0(n)^{\omega}$ and $\mathscr{A}_F^{\mathrm{sc}}(n)^{\mathrm{cyclic}} = \bigcup_{\omega} \mathscr{A}_F^{\mathrm{sc}}(n)^{\omega}$ as ω runs over the set of characters of $F^{\times}$ of order exactly n. Refining Kazhdan's argument slightly, Henniart proved the following

THEOREM 5.4.1 [25]. *For each* $n \geq 1$ *there is a canonical bijection*

$$\pi_F^{\mathrm{cyclic}} : \mathscr{G}_F^0(n)^{\mathrm{cyclic}} \to \mathscr{A}_F^{\mathrm{sc}}(n)^{\mathrm{cyclic}}, \qquad \sigma \mapsto \pi_F^{\mathrm{cyclic}}(\sigma)$$

satisfying conditions (i)–(v) *and* (vii) *of the LLC.*

Here in conditions (iv) and (v) only elements of $\mathscr{A}_F^{\mathrm{sc,\,cyclic}}$ and $\mathscr{G}_F^{0,\,\mathrm{cyclic}}$ are considered.

An explicit local construction of a supercuspidal representation associated to each regular $\eta \in \mathscr{G}_E(1)^{\mathrm{reg}}$ was given by Howe [30] and Gerardin [19], and the relation between this explicitly constructed representation and that arising from global considerations has been recently determined by Henniart [29] (they agree if n is odd and differ by a quadratic twist if n is even); cf. also [44] and [24].

As mentioned above, when $(n, p) = 1$, all of the irreducible representations of W_F are monomial, and hence, are associated to characters of field extensions of degree n. In [44] Moy proved that Howe's construction [30] gives all supercuspidals in this case. Moy also constructed a bijective map from $\mathscr{G}_F^0(n)$ to $\mathscr{A}_F^{\mathrm{sc}}(n)$ that satisfies (i) and (ii) of LLC. H. Reimann [61] constructed a bijective map π_F^R from $\mathscr{G}_F^0(n)$ to $\mathscr{A}_F^{\mathrm{sc}}(n)$ satisfying (i), (ii), and (iii) of LLC, and such that $L(s, \rho) = L(s, \pi_F^R(\rho))$ and $\varepsilon(s, p, \psi) = \varepsilon(s, \pi_F^R(\rho), \psi)$. Moy observed that such a map is not unique. An analogous bijection in the case $n = p$ was obtained in [40] and independently by Henniart.

Finally, for any n, a bijection from $\mathscr{G}_F^0(n)$ to $\mathscr{A}_F^{\mathrm{sc}}(n)$ preserving conductors and compatible with twisting by unramified characters was constructed by Henniart [26].

5.5. Base change. From (5.3.5) and the LLC we should have a "base change" map

$$(5.5.1) \quad BC = BC_F^E : \mathscr{A}_F(n) \to \mathscr{A}_E(n), \qquad \pi = \pi_F(\rho) \mapsto \Pi = \pi_E(\mathrm{res}_{W_E}^{W_F} \rho)$$

whose properties can be read off from the conditions (i)–(vi) of the LLC together with the corresponding properties of the restriction map on representations of W_F'. For example, the restriction map on characters of $\mathrm{Gal}(\overline{F}/F)^{\mathrm{ab}}$ just corresponds, under the reciprocity isomorphism, to composition with the norm: $BC(\eta) = BC_F^E(\eta) = \eta \circ N_{E/F}$.

For simplicity assume that E/F is cyclic of degree n, with associated σ and $\omega_{E/F}$ as above. Then the base change $BC(\pi)$ is invariant under σ. Also, we have

(i) $BC(\pi(\chi)) = BC(\pi)(\chi \circ N_{E/F})$;
(ii) $\omega_{BC(\pi)} = \omega_\pi \circ N_{E/F} = BC(\omega_\pi)$;
(iii) $BC(\pi^\vee) = BC(\pi)^\vee$;
(iv) if $\pi \in \mathscr{A}_F^{\mathrm{sc}}(m)$ and $\pi' \in \mathscr{A}_F^{\mathrm{sc}}(m')$, then

$$L(s, BC(\pi) \times BC(\pi')) = \prod_{i=0}^{n-1} L(s, \pi \times \pi'(\omega^i)),$$

where $\omega = \omega_{E/F}$ is the character of order n associated to E/F;

(v) for π and π' as in (iv),

$$\begin{aligned}&\epsilon(s, BC(\pi) \times BC(\pi'), \psi)\\ &\quad = \varepsilon(s, BC(\mathbb{1}), \psi \circ \mathrm{tr}_{E/F})^{mm'} \prod_{i=0}^{n-1} \frac{\varepsilon(s, \pi \times \pi'(\omega^i), \psi)}{\varepsilon(s, \omega^i, \psi)^{mm'}},\end{aligned}$$

where $BC(\mathbb{1})$ is the trivial character of $E^\times$ (= the base change of the trivial character of $F^\times$);

(vi) BC is compatible with conductors;
(vii) (transport of structure) if F is a finite extension of a field F_0 and σ is an embedding of F into $\overline{F}$ over F_0, then

$$BC_F^E(\pi)^\sigma = BC_{F^\sigma}^{E^\sigma}(\pi^\sigma).$$

Note that formula (iv) follows from the fact that the L-factor for W_F is inductive—this is why we have assumed that π and π' are supercuspidal. If $\pi = \pi_F(\rho)$ and $\pi' = \pi_F(\rho')$, then

$$\begin{aligned} L(s, BC(\pi) \times BC(\pi')) &= L(s, \mathrm{res}_E^F(\rho) \otimes \mathrm{res}_E^F(\rho')) \\ &= L(s, \mathrm{res}_E^F(\rho \otimes \rho')) \\ &= L(s, \mathrm{ind}_E^F \circ \mathrm{res}_E^F(\rho \otimes \rho')) \quad (L \text{ inductive for } W_F) \\ &= \prod_{i=0}^{n-1} L(s, \rho \otimes \rho'(\omega^i)) \\ &= \prod_{i=0}^{n-1} L(s, \pi \times \pi'(\omega^i)). \end{aligned} \tag{5.5.2}$$

Here we use the fact that $\mathrm{ind}_E^F \circ \mathrm{res}_E^F(\tau) = \bigoplus_{i=0}^{n-1} \tau(\omega^i)$. Formula (v) is obtained in a similar fashion, taking into account the fact that ε is *only inductive in degree zero*, so that a suitable multiple of the trivial representation must be carried along.

Again for E/F cyclic, the image BC is precisely the set of all $\Pi \in \mathscr{A}_E$ such that $\Pi \circ \sigma \simeq \Pi$, and two representations π and π' have the same base change if and only if $\pi' = \pi(\omega^i)$ for some i.

One can read off from the LLC the behavior under BC of various types of representations. For example, suppose that $\pi \in \mathscr{A}_F^{\mathrm{sc}}(n)$ is a supercuspidal representation with $\pi = \pi_F(\rho)$ for some irreducible representation ρ of W_F. The nature of $BC(\pi)$ then depends on the nature of $\mathrm{res}_E^F(\rho)$ as a representation of W_E. Suppose, for example, that E/F is cyclic of degree n and that ρ is monomial, induced from a regular character η of $E^\times$. Then $BC(\pi) = \pi(\eta \oplus \eta^\sigma \oplus \cdots \oplus \eta^{\sigma^{n-1}}) = I_B^G(\eta \otimes \eta^\sigma \otimes \cdots \otimes \eta^{\sigma^{n-1}})$. Note that the conjugates of η cannot be linked, so this induced representation must be irreducible. Note that such supercuspidals are characterized by $\pi(\omega) = \pi \otimes \omega_{E/F} \simeq \pi$, as discussed above. At the other extreme, suppose that $\pi \not\simeq \pi(\omega^i)$ for $1 \le i \le n-1$. Then we must have $\rho \not\simeq \rho(\omega^i)$ for $1 \le i \le n-1$ as well, and hence, $\mathrm{res}_E^F(\rho)$ is irreducible via

$$
\begin{aligned}
\mathrm{Hom}_{W_E}(\mathrm{res}_E^F(\rho), \mathrm{res}_E^F(\rho)) &\simeq \mathrm{Hom}_{W_F}(\rho, \mathrm{ind}_{W_E}^{W_F}(\mathrm{res}_E^F(\rho))) \\
&\simeq \mathrm{Hom}_{W_F}(\rho, \mathrm{ind}_{W_E}^{W_F}(\mathbb{1}) \otimes \rho) \\
&\simeq \mathrm{Hom}_{W_F}\left(\rho, \bigoplus_{i=0}^{n-1} \rho(\omega^i)\right).
\end{aligned}
\tag{5.5.3}
$$

Of course, *all of this discussion assumes the LLC!* Shintani made the fundamental observation that, in the cyclic case, one could give a good definition of the relation $\Pi = BC(\pi)$ without a definition of the correspondence $\pi = \pi(\rho)$. This is done in terms of characters and a "norm" map from $\mathrm{GL}(n, E)$ to $\mathrm{GL}(n, F)$. The reader should consult [**52, 20**] and especially [**2**, Chapter 1, §6]. In this last reference, Arthur and Clozel prove rather definitive results on the existence of the base change BC for cyclic extensions. For example, the relations (iv) and (v) predicted by the LLC are valid [**2**, Proposition 6.9].

5.6. ***L*-indistinguishability.** As remarked in the introduction, the representation theory of the group $\mathrm{GL}(n)$ is deceptively simple. For the F-rational points G of a general connected reductive group over F, there is a general version of the LLC that relates irreducible admissible representations of G to admissible representations of W_F' into the L-group ${}^LG$; c.f. Borel's article [**7**]. A representation $\varphi : W_F' \to {}^LG$ should now parameterize a finite set Π_φ of isomorphism classes of irreducible admissible representations of G. For any finite-dimensional continuous representation $r : {}^LG \to \mathrm{GL}(n, \mathbb{C})$, the representations in the set Π_φ are assigned the L and ε-factors $L(s, r \circ \varphi)$ and $\varepsilon(s, r \circ \varphi, \psi)$ associated to the complex representation $r \circ \varphi$ of W_F'. Hence, these representations are *L-indistinguishable.* Note that, by the analogue of the discussion about functoriality above, the whole L-packet Π_φ would be "sent to" a representation $\pi_F(r \circ \varphi)$ of $\mathrm{GL}(n, F)$.

In fact, the size of the L-packet Π_φ should be controlled by the finite group $\pi_0(C_\varphi)$, where C_φ is the centralizer in the complex group ${}^LG^0$ of the image of φ and $\pi_0(C_\varphi)$ is its group of components. In the case of $\mathrm{GL}(n)$, ${}^LG^0 = \mathrm{GL}(n, \mathbb{C})$, and so these centralizers are always connected, and the L-packets consist of single representations.

The simplest example for which nontrivial L-packets arise is the group $G = \mathrm{SL}(2)$. This case, in which ${}^LG^0 = \mathrm{PGL}(2, \mathbb{C})$ and the centralizers can have order 2, is discussed in detail in Shelstad's Corvallis article **[51]** following the work of Labesse and Langlands **[41]**. The reader should also consult the article of Arthur and Gelbart **[3]**. A more geometric version of the local Langlands conjecture has recently been proposed by Vogan **[58]**.

References

1. J. Arthur, *Automorphic representations and number theory*, Canad. Math. Soc. Conf. Proc. **1** (1981), 3–51.
2. J. Arthur and L. Clozel, *Simple algebras, base change, and the advanced theory of the trace formula*, Ann. of Math. Stud., vol. 120, Princeton Univ. Press, Princeton, NJ, 1989.
3. J. Arthur and S. Gelbart, *Lectures on automorphic L-functions*, L-Functions and Arithmetic (J. Coates and M. J. Taylor, eds.), London Math Soc. Lecture Notes, vol. 153, Cambridge Univ. Press, London and New York, 1991, pp. 1–59.
4. J. Bernstein and A. V. Zelevinski, *Representations of the group* $\mathrm{GL}(n, F)$ *where* F *is a local non-Archimedean field*, Russian Math. Surveys **31** (1976), 1–68.
5. ——, *Induced representations of reductive p-adic groups*. I., Ann. Sci. École Norm. Sup. **4** (1977), 441–472.
6. A. Borel, *Admissible representations of a semi-simple group over a local field with vectors fixed under an Iwahori subgroup*, Invent. Math. **35** (1976), 233–259.
7. ——, *Automorphic L-functions*, Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 27–61.
8. A. Borel and N. Wallach, *Continuous cohomology discrete subgroups, and representations of reductive groups*, Ann. of Math. Stud., vol. 94, Princeton Univ. Press, Princeton, NJ, 1980.
9. C. Bushnell, *Gauss sums and local constants for* $\mathrm{GL}(N)$, L-Functions and Arithmetic (J. Coates and M. J. Taylor, eds.), London Math Soc. Lecture Notes, vol. 153, Cambridge Univ. Press, London and New York, 1991, pp. 61–73.
10. C. Bushnell and P. Kutzko, *The admissible dual of* $\mathrm{GL}(N)$ *via compact open subgroups*, Ann. of Math. Stud., vol. 129, Princeton Univ. Press, Princeton, NJ, 1993.
11. L. Clozel, *Motifs et formes automorphes: Applications du principe de fonctorialité*, Automorphic Forms, Shimura Varieties, and L-Functions (L. Clozel and J. Milne, eds.), Academic Press, New York, 1990, pp. 77–159.
12. L. Corwin, *A construction of the supercuspidal representations of* $\mathrm{GL}_n(F)$, F *p-adic*, Trans. Amer. Math. Soc., to appear.
13. P. Deligne, *Les constantes des équations fonctionnelles des fonctions* L, Modular Functions of One Variable II, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973.
14. P. Deligne, D. Kazhdan, and M.-F. Vignéras, *Représentations des algèbres centrales simples p-adiques*, Représentations des Groups Réductifs sur un Corp Locaux, Hermann, Paris, 1984, pp. 33–118.
15. V. G. Drinfeld, *Elliptic modules*, Math. USSR-Sb. **23** (1974), 561–592.

16. ——, *Elliptic modules.* II, Math. USSR-Sb. **31** (1977), 159–170.
17. S. Gelbart, *An elementary introduction to the Langlands program*, Bull. Amer. Math. Soc. **10** (N. S.) (1980), 177–219.
18. S. Gelbart and H. Jacquet, *A relation between automorphic representations of* GL(2) *and* GL(3), Ann. Sci. École Norm. Sup. **11** (1978), 471–542.
19. P. Gerardin, *Cuspidal unramified series for central simple algebras over local fields*, Automorphic Forms, Representations and L-Functions, part 1, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 157–170.
20. P. Gerardin and J.-P. Labesse, *The solution of a base change problem for* GL(2), Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 115–133.
21. R. Godement and H. Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Math., vol. 260, Springer-Verlag, Berlin and New York, 1972.
22. G. Henniart, *La correspondance de Langlands locale: Caractérisation et propriétés fonctorielles*, in preparation.
23. ——, *La conjecture de Langlands locale pour* GL(3), Mem. Soc. Math. France **11–12** (1984), 1–186.
24. ——, *Le point sur la conjecture de Langlands pour* GL(n) *sur un corp local*, Sém. Théor. Nombres, vol. 59, Birkhäuser, Boston, 1985, pp. 115–131.
25. ——, *On the local Langlands conjecture for* GL(n): *The cyclic case*, Ann. of Math. **123** (1986), 143–203.
26. ——, *La conjecture de Langlands locale numérique pour* GL(n), Ann. Sci. École Norm. Sup. **21** (1988), 497–544.
27. ——, *Caractérisation de la correspondance de Langlands locale par les facteurs* ε *de paires*, preprint, 1991.
28. ——, *Représentations des groupes réductifs* p*-adiques*, Astérisque, to appear.
29. ——, *Correspondance de Langlands-Kazhdan explicite dans le cas non ramifié*, preprint, 1992.
30. R. Howe, *Tamely ramified supercuspidal representations of* $GL_n(F)$, Pacific J. Math. **73** (1977), 437–460.
31. H. Jacquet, *Sur les représentations des groupes réductifs* p*-adiques*, C. R. Acad. Sci. Paris **280** (1975), 1271–1272.
32. ——, *Generic representations*, in Non-Commutative Harmonic Analysis, Lecture Notes in Math., vol. 587, Springer-Verlag, Berlin and New York, 1977, pp. 91–101.
33. ——, *Principal* L*-functions of the linear group*, Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 63–86.
34. H. Jacquet, I. I. Piatetski-Shapiro, and J. Shalika, *Conducteur des représentations du groupe linéaire*, Math. Ann. **256** (1981), 199–214.
35. ——, *Rankin-Selberg convolutions*, Amer. J. Math. **105** (1983), 367–464.
36. D. Kazhdan, *On lifting*, Lie Group Representations II, Lecture Notes in Math., vol. 1041, Springer-Verlag, Berlin and New York, 1983, pp. 207–249.
37. A. W. Knapp, *Local Langlands correspondence: The archimedean case*, these Proceedings, vol. 2, pp. 393–410.
38. B. Kostant, *The principal three-dimensional subgroups and the Betti numbers of a complex simple Lie group*, Amer. J. Math. **81** (1959), 973–1032.
39. P. Kutzko, *The local Langlands conjecture for* GL(2) *of a local field*, Ann. of Math. **112** (1980), 381–412.
40. P. Kutzko and A. Moy, *On the local Langlands conjecture in prime dimension*, Ann. of Math. **121** (1985), 495–517.
41. J.-P. Labesse and R. P. Langlands, *L-indistinguishability for* $SL(2)$, Canad. J. Math. **31** (1979), 726–785.
42. R. P. Langlands, *Automorphic representations, Shimura varieties, and motives. Ein Märchen*, Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 205–246.
43. G. Laumon, M. Rapoport, and U. Stuhler, $\mathcal{D}$*-elliptic sheaves and the Langlands correspondence*, preprint, 1991.

44. A. Moy, *Local constants and the tame Langlands correspondence*, Amer. J. Math. **108** (1986), 863–929.
45. J. Ritter, ed., *Representation theory and number theory in connection with the local Langlands conjecture*, Contemp. Math., vol. 86, Amer. Math. Soc., Providence, RI, 1989.
46. F. Rodier, *Décomposition de la série principale des groupes réductifs p-adiques*, Noncommutative Harmonic Analysis and Lie Groups (Marseille, 1980), Lecture Notes in Math., no. 880, Springer-Verlag, Berlin and New York, 1981, pp. 408–424.
47. ——, *Sur les facteurs eulériens associés aux sous-quotients des séries principales des groupes réductifs p-adiques*, Publ. Math. Univ. Paris VII **15** (1981), 107–133.
48. ——, *Représentations de* $\mathrm{GL}(n, k)$ *où* k *est un corps p-adique*, Astérisque **92-93** (1982), 201–218, Sém. Bourbaki 34ème année, 1981–82, exposé n° 583.
49. F. Shahidi, *Fourier transforms of intertwining operators and Plancherel measures for* $\mathrm{GL}(n)$, Amer. J. Math. **106** (1984), 67–111.
50. ——, *On the Ramanujan conjecture and finiteness of poles for certain L-functions*, Ann. of Math. **127** (1988), 547–584.
51. D. Shelstad, *L-indistinguishability*, Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 193–203.
52. T. Shintani, *On liftings of holomorphic cusp forms*, Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 97–110.
53. A. Silberger, *The Langlands quotient theorem for p-adic groups*, Math. Ann. **236** (1978), 95–104.
54. M. Tadic, *Classification of unitary representations in irreducible representations of general linear group (non-Archimedean case)*, Ann. Sci. École Norm. Sup. **19** (1986), 335–382.
55. J. Tate, *Number theoretic background*, Automorphic Forms, Representations, and L-Functions, part 2, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 3–22.
56. J. Tunnell, *On the local Langlands conjecture for* $\mathrm{GL}(2)$, Invent. Math. **46** (1978), 179–200.
57. D. Vogan, *The unitary dual of* $\mathrm{GL}(n)$ *over an archimedean field*, Invent. Math. **83** (1986), 449–505.
58. ——, *The local Langlands conjecture*, Representation Theory of Groups and Algebras (J. Adams, et al., eds.), Contemporary Math., vol. 145, Amer. Math. Soc., Providence, RI, 1993, pp. 305–379.
59. A. Zelevinski, *Induced representations of p-adic reductive groups.* II; *Irreducible representations of* $\mathrm{GL}(n)$, Ann. Sci. École Norm. Sup. **3** (1980), 165–210.
60. ——, *A p-adic analogue of the Kazhdan-Lusztig conjecture*, Funktsional Anal. i Prilozhen. **15** (1981), 9–21. (Russian)

Added in proof.

61. H. Reimann, *Representations of tamely ramified p-adic division and matrix algebras*, J. Number Theory **38** (1991), 58–105.

UNIVERSITY OF MARYLAND, COLLEGE PARK

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Local Langlands Correspondence: The Archimedean Case

A. W. KNAPP

The theory of group representations provides a rich supply of automorphic L functions that are candidates to be the L functions of motives. A motivic L function encodes arithmetic information in an analytic function defined as an Euler product and convergent in a right half-plane. Exploiting this information requires deriving conjectural properties of this analytic function, such as its analytic continuation and functional equation. In practice these properties are obtained only as a consequence of identifying the given motivic L function with an automorphic L function.

Actually the calculus of L functions is a rather small manifestation of a rather large enterprise known as the Langlands program. The general Langlands program works with a reductive group G over a global field F and with the representations that occur (in a suitable sense) in $L^2(G(F)\backslash G(\mathbb{A}))$, where $\mathbb{A}$ denotes the adèles of F. To each such representation the program associates an L function, and it is hoped that these L functions have the same kinds of nice analytic properties as the L functions of Hecke. Some original papers of Langlands on this program are **[19, 20, 21]**. Gelbart **[7]** has given an exposition of the scope of the theory.

The point of the present paper is to give an account of the relevant parts of representation theory that are occurring at the Archimedean places. Largely what we shall discuss is the Langlands treatment of representation theory of GL_n over $\mathbb{R}$ and $\mathbb{C}$. For $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{C})$, the Langlands classification theorem for irreducible admissible representations can be stated in a framework that sounds plausible for all local fields. We call this framework the "local Langlands correspondence." Within this framework, the contribution

1991 *Mathematics Subject Classification.* 20G20, 22E45; Secondary 20G30, 22E55.
This paper is in final form and no version of it will be submitted for publication elsewhere.
Supported by National Science Foundation Grant DMS 91-00367.

of the Archimedean places to L functions is a rather simple topic.[1] A companion paper by Kudla [18] treats the non-Archimedean places, where the analogous statements are partly theorems and partly conjectures.

Discussion of the classification theorem and the local Langlands correspondence for GL_n will occupy §§2–4 of this paper. In §5 we shall mention the extent to which the Langlands theory generalizes from GL_n to connected reductive groups. The special case of symplectic groups is relevant for the theory of Shimura varieties.

It is important for understanding the Langlands program to know some features of the historical transition from classical automorphic forms to automorphic representation theory. The theory separated into two directions at one time in the early 1970s and then came together several years later. The interplay between the two successful theories accounts for the relatively advanced state of knowledge for GL_n in comparison with other groups, and it is where we begin.

I am happy to acknowledge helpful discussions with E. Bifet, S. Kudla, H. Matumoto, and C.-H. Sah in connection with writing this paper.

1. Historical transition

Classical automorphic L functions in the work of Hecke arise as Mellin transforms of certain automorphic forms—particularly as transforms of a kind of θ function (in the case of Hecke's theory of grossencharacters) or as transforms of modular forms of the subgroup

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) \;\middle|\; c \equiv 0 \bmod N \right\}.$$

Hecke's theory of grossencharacters establishes analytic continuations and functional equations for L functions that generalize Dirichlet L functions. Tate's thesis [28] recast this theory of Hecke's in a representation-theoretic setting that corresponds to the regular representation of $GL_1(\mathbb{A})$ on $GL_1(F)\backslash GL_1(\mathbb{A})$, $\mathbb{A}$ again being the adèles of a global field F. For a fixed grossencharacter, Tate's method in effect attaches to each place a local L factor given by an integral, as well as a local ε factor that contributes to a local functional equation. Also in effect, the method constructs a global L function as the product of the local L factors, and then it proves directly a global functional equation. We shall amplify this discussion shortly.

In the 1950s Gelfand and Fomin realized that modular forms are connected with representations of GL_2. Some expositions of this connection are in parts of Gelfand, Graev, and Pyatetskiĭ-Shapiro [9], Weil [35], Deligne [2], Gelbart [6], and Piatetski-Shapiro [23]. In part, the connection is that one can identify cusp forms for $\Gamma_0(N)$ in two stages with functions on groups. In the

[1] In classical terminology, it is customary to include only the factors from the non-Archimedean places in L and to give another name to the product of L with various gamma factors that come from the Archimedean places. But we shall follow the convention used in representation theory of including factors from all places in the definition of L.

first stage the identification is with certain functions on $SL_2(\mathbb{R})$ transforming on the left side by $\Gamma_0(N)$ and on the right side by the rotation subgroup. In the second stage it is with functions on $GL_2(\mathbb{A})$ whose integrals of a certain kind vanish, where $\mathbb{A}$ is now the ring of adèles of $\mathbb{Q}$. The cusp forms are then intimately connected with the decomposition into irreducible representations of the representation of $GL_2(\mathbb{A})$ on the space of functions on $GL_2(\mathbb{Q})\backslash GL_2(\mathbb{A})$ that transform by a character of the center $Z_{\mathbb{A}}$ of $GL_2(\mathbb{A})$ and are square integrable on $GL_2(\mathbb{Q})Z_{\mathbb{A}}\backslash GL_2(\mathbb{A})$.

The Hecke operators turn out to have a nice interpretation in the representation-theoretic setting. As a result, each cusp form with an Euler product expansion corresponds, under the above two-stage identification, with a function that lies in a subspace irreducible under $GL_2(\mathbb{A})$. This infinite-dimensional representation is like a grossencharacter in that it is the "tensor product" of representations for the local groups, one for $GL_2(\mathbb{R})$ and one for each $GL_2(\mathbb{Q}_p)$, and Jacquet and Langlands **[13]** were able to develop a theory for GL_2 parallel to the Tate theory for GL_1. They made critical use of a classification of the irreducible admissible representations of GL_2 of each local field. Expositions are in Robert **[24]** and Gelbart **[6]**.

In the early 1970s, generalization to GL_n took two different paths, because of the absence of a classification for $n > 2$. Godement and Jacquet **[10]** used the method of Tate to develop a theory that led to the analytic continuation and functional equation of L functions, without an explicit identification of all the L functions. Langlands **[21]** set up a conjectural framework for classification, with explicit L functions, but his theory did not account for the analytic continuation and functional equation. The final theory for GL_n requires both parts, and it was Jacquet **[12]** who showed the two parts are compatible.

To make clear the distinction between the two parts, let us amplify the discussion of Tate's method, using notation appropriate to GL_n. Proofs of the various steps may be found in **[28, 13, 10, 12]**. We shall sketch the theory at the Archimedean places, indicate how it can be adjusted for the non-Archimedean places, and say briefly what happens globally.

First let k be $\mathbb{R}$ or $\mathbb{C}$, and let $M_n(k)$ be n-by-n matrix space over k. Let

$$K = \begin{cases} O(n) & \text{if } k = \mathbb{R}, \\ U(n) & \text{if } k = \mathbb{C}. \end{cases} \tag{1.1}$$

Let (ρ, V) be an admissible representation of $GL_n(k)$, "admissible" being defined in §2, and let $(\tilde{\rho}, \widetilde{V})$ be the admissible dual. A K **finite matrix coefficient** of ρ is a function

$$c(x) = \langle \rho(x)u, \tilde{u} \rangle$$

with $u \in V$ and $\tilde{u} \in \widetilde{V}$ both transforming in finite-dimensional spaces under K. The function $\check{c}(x) = c(x^{-1})$ is another K finite matrix coefficient, because it is given by

$$\check{c}(x) = \langle u, \tilde{\rho}(x)\tilde{u}\rangle.$$

The subspace $\mathscr{S}_0$ of the Schwartz space $\mathscr{S}(M_n(k))$ is to consist of all functions of the form

$$P(x_{ij}) \exp\left(-\pi \sum x_{ij}^2\right) \qquad \text{if } k = \mathbb{R},$$
$$P(z_{ij}\bar{z}_{ij}) \exp\left(-2\pi \sum z_{ij}\bar{z}_{ij}\right) \qquad \text{if } k = \mathbb{C},$$

where P is an arbitrary polynomial. For any K finite matrix coefficient c of ρ and any function f in the space $\mathscr{S}_0$, we define

$$\zeta(f, c, s) = \int_{M_n(k)} f(x)c(x)|\det x|_k^s \, d^\times x \tag{1.2}$$

for s complex. Here $|z|_\mathbb{R} = |z|$ and $|z|_\mathbb{C} = |z|^2$. The measure is $d^\times x = |\det x|_k^{-n}\, dx$, where dx is a fixed invariant measure for $M_n(k)$; use of $d^\times x$ in the notation may be regarded as an additive normalization of the parameter s.

Assume ρ is irreducible. Then all the integrals (1.2) converge for s in a common right half-plane and extend to be meromorphic functions for s in $\mathbb{C}$. Moreover, there exist finitely many choices of (c, f), say (c_i, f_i), such that

$$L(s, \rho) = \sum_i \zeta(f_i, c_i, s) \tag{1.3}$$

has the following property: For any (c, f),

$$\zeta(f, c, s + \tfrac{1}{2}(n-1)) = P(f, c, s)L(s, \rho) \tag{1.4}$$

for a polynomial P in s. The function $L(s, \rho)$ is uniquely determined by these properties, up to a scalar factor, and is called a **local L factor**. (In Tate's original work, in which $n = 1$, the matrix coefficient c is essentially unique, and the sum on the right side of (1.3) collapses to a single term.)

Let ψ be the additive character of k given by

$$\begin{aligned} \psi(x) &= \exp(2\pi i x) && \text{if } k = \mathbb{R}, \\ \psi(z) &= \exp(2\pi i(z + \bar{z})) && \text{if } k = \mathbb{C}, \end{aligned} \tag{1.5}$$

and define the Fourier transform $\widehat{f}$ of a member f of $\mathscr{S}_0$ by

$$\widehat{f}(x) = \int_{M_n(k)} f(y)\psi(\operatorname{Tr}(xy))\, dy, \tag{1.6}$$

where dy is the self-dual Haar measure on $M_n(k)$. Then $\widehat{f}$ is again in $\mathscr{S}_0$.

With ρ still irreducible, there exists a meromorphic function $\gamma(s, \rho, \psi)$ independent of f and c such that

$$\zeta(\widehat{f}, \check{c}, 1 - s + \tfrac{1}{2}(n-1)) = \gamma(s, \rho, \psi)\zeta(f, c, s + \tfrac{1}{2}(n-1)) \tag{1.7}$$

for all f in $\mathscr{S}_0$ and all K finite matrix coefficients c of ρ. In terms of

$$\varepsilon(s, \rho, \psi) = \gamma(s, \rho, \psi)L(s, \rho)/L(1-s, \tilde{\rho}), \tag{1.8}$$

the local functional equation reads

$$\frac{\zeta(\widehat{f}, \check{c}, 1-s+\frac{1}{2}(n-1))}{L(1-s, \tilde{\rho})} = \varepsilon(s, \rho, \psi)\frac{\zeta(f, c, s+\frac{1}{2}(n-1))}{L(s, \rho)}. \tag{1.9}$$

From (1.3) and (1.4), it follows that $\varepsilon(s, \rho, \psi)$ is a polynomial. Let ω_ρ be the quasicharacter of $k^\times$ such that $\rho(a \cdot 1) = \omega_\rho(a)1$. Since $(\widehat{f})^\wedge(x) = f(-x)$, the change of variables $x \to -x$ in (1.1) and iteration of (1.9) gives

$$\varepsilon(s, \rho, \psi)\varepsilon(1-s, \tilde{\rho}, \psi) = \omega_\rho(-1).$$

Therefore $\varepsilon(s, \rho, \psi)$ is a constant. Apart from evaluation of $L(s, \rho)$ and $\varepsilon(s, \rho, \psi)$, this completes the local part of Tate's method at Archimedean places.

Next let k be a non-Archimedean local field. In the definition of $\zeta(f, c, s)$, we make the following adjustments: Admissibility of (ρ, V) is defined in **[18]**, mention of K finiteness of the matrix coefficient $c(x)$ may be omitted, and $\mathscr{S}_0$ is taken as the whole Schwartz-Bruhat space of locally constant functions of compact support on $M_n(k)$.

For this new k, assume ρ is irreducible. It is still true that the integrals (1.2) converge in a common right half-plane and extend to be meromorphic functions for s in $\mathbb{C}$. Moreover, the definition of $L(s, \rho)$ in (1.3) and (1.4) still applies, with the following modifications: P in (1.4) is a polynomial in q^s and q^{-s}, where q is the number of elements in the residue field, and $L(s, \rho)$ can be normalized so as to be $1/Q(q^{-s})$, where Q is a polynomial with constant term 1.

Let ψ be a nontrivial additive character of k, and define the Fourier transform by (1.6). Then there exists a meromorphic function $\gamma(s, \rho, \psi)$ independent of f and c such that (1.7) holds for all f in $\mathscr{S}_0$ and all matrix coefficients c of ρ. With $\varepsilon(s, \rho, \psi)$ as in (1.8), the local functional equation reads as in (1.9). One sees that $\varepsilon(s, \rho, \psi)$ is a nonzero multiple of a power of q^{-s}.

Finally let F be a global field, let $\mathbb{A}$ be the adèles of F, and let $\mathbb{I}$ be the idèles of $F^\times$. Some care is required in identifying what irreducible admissible representations ρ of $\mathrm{GL}_n(\mathbb{A})$ and what matrix coefficients are to be allowed; these details are in §§10–12 of **[10]**. The representation ρ is to occur in the regular representation on $\mathrm{GL}_n(F)\backslash \mathrm{GL}_n(\mathbb{A})$. Moreover, it is to reduce to a quasicharacter ω_ρ on members of $\mathbb{I}$ (scalar matrices) in the sense that $\rho(a \cdot 1_n) = \omega_\rho(a)1_V$, and ω_ρ is to be trivial on $F^\times$. In addition, ρ is to incorporate conditions of vanishing integrals of the kind satisfied by embedded cusp forms when $n = 2$ and $F = \mathbb{Q}$.

Such an irreducible representation ρ is the adelic product of irreducible representations ρ_v of each local field k_v. This result is proved for $n = 2$ in **[13]**, and a general argument may be found in Flath **[5]**. One defines $\zeta(f, c, s)$ by (1.2), with k replaced by $\mathbb{A}$, and with f suitably restricted. The main step is to prove that $\zeta(f, c, s)$ extends to be an entire function and satisfies

$$\zeta(\widehat{f}, \check{c}, n-s) = \zeta(f, c, s). \tag{1.10}$$

We define

$$\begin{aligned} L(s, \rho) &= \prod_v L(s, \rho_v), \\ \varepsilon(s, \rho, \psi) &= \prod_v \varepsilon(s, \rho_v, \psi_v), \end{aligned} \tag{1.11}$$

where ψ is a nontrivial additive character of $\mathbb{A}/F$ with local components ψ_v. The **global L function** $L(s, \rho)$ initially is convergent in a right half-plane, and $\varepsilon(s, \rho, \psi)$ has almost all factors 1 and hence is entire. At each place, $L(s, \rho_v)$ is a sum of functions $\zeta(f_v, c_v, s)$. Assembling these f_v's and c_v's into global f and c, we find that $L(s, \rho)$ extends to be entire. Taking the product of (1.9) over all places and substituting from (1.10) and (1.11), we obtain

$$L(s, \rho) = \varepsilon(s, \rho, \psi) L(1-s, \tilde{\rho}). \tag{1.12}$$

This is the functional equation for $L(s, \rho)$ that we have sought, and our discussion of Tate's method is complete.

We have not discussed the evaluation of $L(s, \rho)$ and $\varepsilon(s, \rho, \psi)$ in the local case. Godement and Jacquet **[10]** showed that it is possible to compute these expressions for most places and most ρ's with rather little information about classification of irreducible admissible representations. Later Jacquet **[12]** showed that one can compute these functions in all cases with just a little more information about classification.

In the second part of the theory of GL_n, Langlands **[21]** took the known values of L and ε, went a long way toward classification of irreducible admissible representations in the Archimedean case, and organized the information about classification and L factors into a framework that showed promise for being valid for all local fields, Archimedean or not. This framework is called the **local Langlands correspondence**. One of its features is that it makes sense for general reductive groups, not just GL_n. We shall describe the classification for GL_n at the Archimedean places and then the local Langlands correspondence for organizing this information. We work with $\mathbb{R}$ in §§2–3 and with $\mathbb{C}$ in §4. The part of Jacquet **[12]** that deals with Archimedean places effectively shows that the local Langlands correspondence for GL_n attaches the same L and ε factors for $\mathbb{R}$ and $\mathbb{C}$ as in **[10]**.

2. Langlands classification for $\mathrm{GL}_n(\mathbb{R})$

The Langlands classification for $G = \mathrm{GL}_n(\mathbb{R})$ describes all irreducible admissible representations[2] of G up to infinitesimal equivalence.

Let $K = O(n)$ be the maximal compact subgroup of G given in (1.1). A representation (ρ, V) of G on a Hilbert space V will be said to be **admissible** if, in the restriction of ρ to K, each irreducible representation of K occurs with at most finite multiplicity. It is **irreducible** if V has no nontrivial closed invariant subspace. If ρ is admissible, let V_K be the space of K **finite** vectors, those transforming in a finite-dimensional space under K. Each member of V_K is a C^∞ vector, and V_K is a representation space for both K and the Lie algebra $\mathfrak{g}$ of G. (See [**14**, Chapter III].)

Moreover, these representations of K and $\mathfrak{g}$ are compatible, and (ρ, V_K) is a $(\mathfrak{g}, K)$ module, in the sense of [**33**]. From Harish-Chandra [**11**], it is known that the closed G invariant subspaces U of V are in one-one correspondence with the arbitrary $(\mathfrak{g}, K)$ invariant subspaces U_K of V_K, the correspondence being $U = \overline{U}_K$ and $U_K = U \cap V_K$. In particular, (ρ, V) is irreducible if and only if (ρ, V_K) is algebraically irreducible. (See [**14**, Chapter VIII].)

Two admissible representations are said to be **infinitesimally equivalent** if their underlying $(\mathfrak{g}, K)$ modules are isomorphic. Infinitesimally equivalent admissible representations have the same K finite matrix coefficients.

Any irreducible admissible representation (ρ, V) has a **central character** given as a quasicharacter $\omega_\rho : \mathbb{R}^\times \to \mathbb{C}^\times$ by $\rho(a\cdot 1_n) = \omega_\rho(a)1_V$ for $a \in \mathbb{R}^\times$.

Let $\mathrm{SL}_m^\pm(\mathbb{R})$ be the subgroup of elements g of $\mathrm{GL}_m(\mathbb{R})$ with $|\det g| = 1$. The unimodular subgroup $\mathrm{SL}_m(\mathbb{R})$ has index 2 in $\mathrm{SL}_m^\pm(\mathbb{R})$. We shall specify certain irreducible representations of $\mathrm{SL}_m^\pm(\mathbb{R})$ for the cases $m = 1$ and $m = 2$. For $m = 1$, there are only two representations, and they are both one dimensional; we write 1 for the trivial one and sgn for the nontrivial one. For $m = 2$ the representations of interest are the ones in the "discrete series," denoted D_l for integers $l \geq 1$. These representations are induced from $\mathrm{SL}_2(\mathbb{R})$ as

$$D_l = \mathrm{ind}_{\mathrm{SL}_2(\mathbb{R})}^{\mathrm{SL}_2^\pm(\mathbb{R})}(D_l^+). \tag{2.1a}$$

Here D_l^+ acts in the space of analytic functions f in the upper half-plane with

$$\|f\|^2 = \iint |f(z)|^2 y^{l-1}\, dx\, dy$$

finite, the action by $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ being

$$D_l^+(g)f(z) = (bz+d)^{-(l+1)} f\left(\tfrac{az+c}{bz+d}\right). \tag{2.1b}$$

[2] By convention a representation of a group always takes place on a complex vector space, and the group action is assumed to be jointly continuous.

The representations D_l of $SL_2^{\pm}(\mathbb{R})$ are irreducible unitary, and their matrix coefficients are square integrable. (See [**14**, p. 35].)

The building blocks for irreducible admissible representations of $GL_n(\mathbb{R})$ are the representations of $GL_1(\mathbb{R})$ and $GL_2(\mathbb{R})$ obtained by tensoring the above representations on $SL^{\pm}$ with a quasicharacter $a \to |\det a|_{\mathbb{R}}^t$ on the positive scalar matrices of size 1 or 2. Here, as earlier, $|\cdot|_{\mathbb{R}}$ denotes ordinary absolute value, and t is in $\mathbb{C}$. Thus the building blocks will be

$$\text{(2.2a)} \qquad \left.\begin{array}{c} 1 \otimes |\cdot|_{\mathbb{R}}^t \\ \operatorname{sgn} \otimes |\cdot|_{\mathbb{R}}^t \end{array}\right\} \quad \text{for } GL_1(\mathbb{R}),$$

$$\text{(2.2b)} \qquad D_l \otimes |\det(\cdot)|_{\mathbb{R}}^t \} \quad \text{for } GL_2(\mathbb{R}).$$

To any partition of n into 1's and 2's, say $(n_1, \ldots, n_r)$ with each n_j equal to 1 or 2 and with $\sum n_j = n$, we associate the block diagonal subgroup

$$\text{(2.3)} \qquad D = GL_{n_1}(\mathbb{R}) \times \cdots \times GL_{n_r}(\mathbb{R}).$$

For each j with $1 \le j \le r$, let σ_j be a representation of $GL_{n_j}(\mathbb{R})$ of the form (2.2), and write t_j for t. Then $(\sigma_1, \ldots, \sigma_r)$ defines by tensor product a representation of the block diagonal subgroup (2.3), and we extend this representation to the corresponding block upper triangular subgroup $Q = DU$ by making it be the identity on the block strictly upper triangular subgroup U. We set

$$\text{(2.4)} \qquad I(\sigma_1, \ldots, \sigma_r) = \operatorname{ind}_Q^G(\sigma_1, \ldots, \sigma_r),$$

using unitary induction as in [**14**, Chapter VII]. (That is, in the transformation law under Q, $(\sigma_1, \ldots, \sigma_r)$ is tensored with a one-dimensional representation so that when $\sigma_1, \ldots, \sigma_r$ are unitary, $I(\sigma_1, \ldots, \sigma_r)$ is automatically unitary.)

THEOREM 1. *For* $G = GL_n(\mathbb{R})$,

(a) *if the parameters* $n_j^{-1} t_j$ *of* $(\sigma_1, \ldots, \sigma_r)$ *satisfy*

$$\text{(2.5)} \qquad n_1^{-1} \operatorname{Re} t_1 \ge n_2^{-1} \operatorname{Re} t_2 \ge \cdots \ge n_r^{-1} \operatorname{Re} t_r,$$

then $I(\sigma_1, \ldots, \sigma_r)$ *has a unique irreducible quotient* $J(\sigma_1, \ldots, \sigma_r)$,

(b) *the representations* $J(\sigma_1, \ldots, \sigma_r)$ *exhaust the irreducible admissible representations of* G, *up to infinitesimal equivalence,*

(c) *two such representations* $J(\sigma_1, \ldots, \sigma_r)$ *and* $J(\sigma_1', \ldots, \sigma_{r'}')$ *are infinitesimally equivalent if and only if* $r' = r$ *and there exists a permutation* $j(i)$ *of* $\{1, \ldots, r\}$ *such that* $\sigma_i' = \sigma_{j(i)}$ *for* $1 \le i \le r$.

Two ways are known for picking out the constituent $J(\sigma_1, \ldots, \sigma_r)$ of $I(\sigma_1, \ldots, \sigma_r)$. One way, following Langlands [**21**], is as the image of a certain standard intertwining operator on $I(\sigma_1, \ldots, \sigma_r)$; see pp. 198–200 of [**14**] for an exposition. If any of the inequalities in (2.5) is an equality, some normalization of the operator may be necessary in order to eliminate poles.

Techniques for this normalization are explained in Chapter XIV of **[14]**. The other way, following Vogan **[30]**, is by the theory of minimal K types. For a fixed induced representation, the Vogan theory singles out finitely many irreducible representations of K (one such in this case) that occur in the induced representation with multiplicity one, and a vector transforming by any one of these representations of K generates $J(\sigma_1, \dots, \sigma_r)$ as a subquotient. (See **[14,** Chapter XV].) Since G is disconnected, the original Vogan theory is not quite general enough to handle this case, and one must appeal to Vogan **[32]** instead.

The original paper **[21]** of Langlands reduced the classification of irreducible admissible representations of fairly general reductive groups to the classification of the irreducible tempered representations. Here "tempered" is a term referring to the asymptotic behavior of matrix coefficients and need not be explained in this exposition. (See **[14,** p. 198], for a precise definition.) Langlands proved also that the irreducible tempered representations are exactly the irreducible constituents of the representations that in this case have $t_1, \dots, t_r$ purely imaginary. (See **[14,** Chapter VIII].) The irreducible tempered representations were classified in 1976 by Knapp and Zuckerman, and the result was reformulated in **[16]**. Detailed proofs are in **[17]**; for an exposition, see **[14,** Chapter XIV].

The papers **[16]** and **[17]** treat connected semisimple groups and are not literally applicable to $\mathrm{GL}_n(\mathbb{R})$. However, the results are still valid for $\mathrm{GL}_n(\mathbb{R})$, and no new ideas are needed for their proofs. (For general reductive groups with some disconnectedness, a new idea *is* needed. This extension of the classification was carried out by Mirković **[22]**.)

With account taken of the remarks in the previous paragraph, parts (a) and (b) of Theorem 1 are a special case of Theorem 5 of **[16]**. The latter theorem has four hypotheses, and the first three are checked by inspection. For the fourth, it is enough to check that $I(\sigma_1, \dots, \sigma_r)$ is irreducible in the tempered case, i.e., when all t_j are purely imaginary. This is easily done; see §7 of **[15]**. Part (c) is not explicitly written down in these sources, but §3 of **[16]** tells what one has to do to classify the infinitesimal equivalences; Example 3 in that section helps illustrate the technique.

We mention that the irreducible unitary representations of $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{C})$ have been classified by Vogan **[31]**. The original Langlands program makes no predictions about this kind of result.

3. Local Langlands correspondence for $\mathrm{GL}_n(\mathbb{R})$

The Weil group of $\mathbb{R}$, denoted $W_{\mathbb{R}}$, is the nonsplit extension of $\mathbb{C}^\times$ by $\mathbb{Z}/2\mathbb{Z}$ given by

$$W_{\mathbb{R}} = \mathbb{C}^\times \cup j\mathbb{C}^\times,$$

where $j^2 = -1$ and $jcj^{-1} = \bar{c}$. Here bar denotes complex conjugation. We shall be interested in the set of equivalence classes of n-dimensional complex representations of $W_{\mathbb{R}}$ whose images consist of semisimple elements.

The one-dimensional representations of $\mathbb{C}^\times$ are of the form

$$z \to z^\mu \bar{z}^\nu \quad \text{with } \mu \text{ and } \nu \text{ in } \mathbb{C} \text{ and } \mu - \nu \text{ in } \mathbb{Z}. \tag{3.1}$$

In fact, if we write $z = re^{i\theta}$, we obtain $re^{i\theta} \to r^{\mu+\nu} e^{i(\mu-\nu)\theta}$. Hence $\mu - \nu$ is an integer and $\mu + \nu$ is unrestricted.

Let us classify the one-dimensional representations φ of $W_{\mathbb{R}}$. On $\mathbb{C}^\times$, we have $\varphi(z) = z^\mu \bar{z}^\nu$ as in (3.1). Let $\varphi(j) = w$. Then

$$\varphi(\bar{z}) = \varphi(jzj^{-1}) = w\varphi(z)w^{-1} = \varphi(z).$$

Hence $\varphi(z) = z^\nu \bar{z}^\mu$ and $\mu = \nu$. In other words, $\varphi(re^{i\theta}) = r^{2\mu}$. Now

$$1 = \varphi(-1) = \varphi(j^2) = w^2$$

says $w = \pm 1$. Thus the one-dimensional representations are parametrized by a sign and a complex parameter $t = 2\mu$ as follows:

$$\begin{aligned} (+, t):& \quad \varphi(z) = |z|_{\mathbb{R}}^t \text{ and } \varphi(j) = +1, \\ (-, t):& \quad \varphi(z) = |z|_{\mathbb{R}}^t \text{ and } \varphi(j) = -1. \end{aligned} \tag{3.2}$$

Next let us classify the irreducible two-dimensional semisimple representations φ of $W_{\mathbb{R}}$, up to equivalence. The set $\varphi(\mathbb{C}^\times)$ consists of commuting diagonable transformations. Let u, v be a basis in which $\varphi(\mathbb{C}^\times)$ is diagonal. Say $\varphi(z)u = z^\mu \bar{z}^\nu u$ and $\varphi(z)v = z^{\mu'} \bar{z}^{\nu'} v$. If $\mu = \mu'$ and $\nu = \nu'$, then any one-dimensional invariant subspace for $\varphi(j)$ will exhibit φ as reducible. Hence we may assume that $\mu \neq \mu'$ or $\nu \neq \nu'$. Put $u' = \varphi(j)u$. Then

$$\varphi(z)u' = \varphi(j\bar{z}j^{-1})u' = \varphi(j)\varphi(\bar{z})u = z^\nu \bar{z}^\mu \varphi(j)u = z^\nu \bar{z}^\mu u'.$$

If $\mu = \nu$, then u' is in $\mathbb{C}u$, and $\mathbb{C}u$ is an invariant subspace, contradiction. Thus u' must be in $\mathbb{C}v$ with $\nu = \mu'$ and $\mu = \nu'$. We shall replace the basis u, v by the basis u, u'. Since $\varphi(j)^{-1} = \varphi(j)\varphi(-1) = (-1)^{\mu-\nu}\varphi(j)$ on the span of u and u', we can write the result as

$$\begin{aligned} \varphi(z)u &= z^\mu \bar{z}^\nu u, & \varphi(j)u &= u', \\ \varphi(z)u' &= z^\nu \bar{z}^\mu u', & \varphi(j)u' &= (-1)^{\mu-\nu} u. \end{aligned}$$

In terms of the basis u', $(-1)^{\mu-\nu}u$, these formulas become

$$\begin{aligned} \varphi(z)u' &= z^\nu \bar{z}^\mu u', & \varphi(j)u' &= (-1)^{\mu-\nu} u, \\ \varphi(z)((-1)^{\mu-\nu}u) &= z^\mu \bar{z}^\nu ((-1)^{\mu-\nu}u), & \varphi(j)((-1)^{\mu-\nu}u) &= (-1)^{\mu-\nu}u'. \end{aligned}$$

In view of the symmetry here, we may assume that the nonzero integer $\mu - \nu = l$ is positive. We conclude that the equivalence class of φ is classified by a pair (l, t) with $l = \mu - \nu$ an integer ≥ 1 and with $2t = \mu + \nu$ in $\mathbb{C}$. For the pair (l, t) there exists a basis u, u' such that

$$\begin{aligned} (l, t):& \quad \varphi(re^{i\theta})u = r^{2t} e^{il\theta} u, & \varphi(j)u &= u', \\ & \quad \varphi(re^{i\theta})u' = r^{2t} e^{-il\theta} u', & \varphi(j)u' &= (-1)^l u. \end{aligned} \tag{3.3}$$

LEMMA. *Every finite-dimensional semisimple representation φ of $W_{\mathbb{R}}$ is fully reducible, and each irreducible representation has dimension one or two.*

PROOF. Let φ act on the vector space V. Since $\varphi(\mathbb{C}^\times)$ consists of commuting diagonable transformations, V is the direct sum of spaces $V_{\mu,\nu}$ where all $\varphi(z)$ act by $z^\mu \bar{z}^\nu$. As above, we have $\varphi(j)V_{\mu,\nu} = V_{\nu,\mu}$. If $\mu = \nu$, then we can choose a basis of eigenvectors for $\varphi(j)$ in $V_{\mu,\mu}$, and the span of each eigenvector is a one-dimensional invariant subspace under $\varphi(W_{\mathbb{R}})$. If $\mu \neq \nu$, choose a basis $u_1, \dots, u_r$ of $V_{\mu,\nu}$, and put $u_i' = \varphi(j)u_i$ for $1 \le i \le r$. Then $\mathbb{C}u_i \oplus \mathbb{C}u_i'$ is a two-dimensional invariant subspace under $\varphi(W_{\mathbb{R}})$, and the direct sum of these subspaces as i varies is $V_{\mu,\nu} \oplus V_{\nu,\mu}$. This proves the lemma.

Now let φ be an n-dimensional semisimple complex representation of $W_{\mathbb{R}}$. By the lemma, φ is fully reducible. If we list the dimensions of the irreducible constituents in any order, we can regard the result as a partition of n into 1's and 2's, say $(n_1, \dots, n_r)$ with each n_j equal to 1 or 2 and with $\sum n_j = n$. Fix attention on n_j, and let φ_j be the corresponding irreducible constituent of φ. To φ_j we associate a representation σ_j from (2.2) as follows:

$$\begin{array}{lrcll} (3.4) & (+, t) \text{ in } (3.2) & \longrightarrow & 1 \otimes |\cdot|_{\mathbb{R}}^t & \text{in } (2.2\text{a}), \\ & (-, t) \text{ in } (3.2) & \longrightarrow & \operatorname{sgn} \otimes |\cdot|_{\mathbb{R}}^t & \text{in } (2.2\text{a}), \\ & (l, t) \text{ in } (3.3) & \longrightarrow & D_l \otimes |\det(\cdot)|_{\mathbb{R}}^t & \text{in } (2.2\text{b}). \end{array}$$

In this way, we associate a tuple $(\sigma_1, \dots, \sigma_r)$ of representations to φ. If the complex numbers $t_1, \dots, t_r$ do not satisfy (2.5), we permute $(\sigma_1, \dots, \sigma_r)$ so that (2.5) ends up being satisfied. Using Theorem 1, we can then make the association

$$\varphi \longrightarrow \rho_{\mathbb{R}}(\varphi) = J(\sigma_1, \dots, \sigma_r) \tag{3.5}$$

and come to the following conclusion.

THEOREM 2 (Local Langlands Correspondence for $\mathrm{GL}_n(\mathbb{R})$). *The association* (3.5) *is a well-defined bijection between the set of all equivalence classes of n-dimensional semisimple complex representations of $W_{\mathbb{R}}$ and the set of all equivalence classes of irreducible admissible representations of* $\mathrm{GL}_n(\mathbb{R})$.

To each finite-dimensional semisimple complex representation φ of the Weil group of a local field, Weil [**34**] has associated a local L factor with certain nice properties. The results are summarized in Tate [**29**]; see also Shahidi [**25**, especially p. 990]. Some of Tate's results are taken from Deligne

[3]. In the case of $W_{\mathbb{R}}$, when φ is irreducible, the formula is

(3.6)
$$L(s, \varphi) = \begin{cases} \pi^{-(s+t)/2}\Gamma(\frac{s+t}{2}) & \text{if } \varphi \text{ is given by } (+, t) \text{ in (3.2)}, \\ \pi^{-(s+t+1)/2}\Gamma(\frac{s+t+1}{2}) & \text{if } \varphi \text{ is given by } (-, t) \text{ in (3.2)}, \\ 2(2\pi)^{-(s+t+\frac{l}{2})}\Gamma(s+t+\frac{l}{2}) & \text{if } \varphi \text{ is given by } (l, t) \text{ in (3.3)}. \end{cases}$$

(Recall our convention that $l \geq 1$.) For φ reducible, $L(s, \varphi)$ is the product of the L factors of the irreducible constituents of φ.

Fix the additive character ψ of $\mathbb{R}$ in (1.5), so that ordinary Lebesgue measure is self-dual Haar measure on $\mathbb{R}$. The ε factors are given for φ irreducible by

$$\varepsilon(s, \varphi, \psi) = \begin{cases} 1 & \text{if } \varphi \text{ is given by } (+, t) \text{ in (3.2)}, \\ i & \text{if } \varphi \text{ is given by } (-, t) \text{ in (3.2)}, \\ i^{l+1} & \text{if } \varphi \text{ is given by } (l, t) \text{ in (3.3)}. \end{cases} \tag{3.7}$$

For φ reducible, $\varepsilon(s, \varphi, \psi)$ is the product of the ε factors of the irreducible constituents of φ. Observe that the ε factors are constant in s.

In the terminology of **[29]**, both L and ε are "additive" in their behavior with respect to short exact sequences. Also L is "inductive" with respect to change of field, and its formula over $\mathbb{R}$ should really be considered together with its formula (4.6) over $\mathbb{C}$. The existence of ε for all local fields is a theorem of Langlands. See Theorem 3.4.1 of **[29]**. The ε factors satisfy a weaker property than "inductive"; they are "inductive in degree 0" with respect to change of field. The formula (3.7) over $\mathbb{R}$ should be considered together with the formula (4.7) over $\mathbb{C}$. The rule for how ε depends on ψ is given in (3.2.3) of **[29]**.

We can now define local factors $L(s, \rho)$ and $\varepsilon(s, \rho, \psi)$ for each irreducible admissible representation of $\mathrm{GL}_n(\mathbb{R})$ by the rule

$$\left.\begin{aligned} L(s, \rho) &= L(s, \varphi) \\ \varepsilon(s, \rho, \psi) &= \varepsilon(s, \varphi, \psi) \end{aligned}\right\} \quad \text{if } \rho = \rho_{\mathbb{R}}(\varphi) \text{ in (3.5) and Theorem 2.} \tag{3.8}$$

These formulas are consistent with Jacquet and Langlands **[13]**, especially pp. 177–195. Jacquet **[12]** proved the following result.

THEOREM 3. *The definition of* $L(s, \rho)$ *in* (3.8) *satisfies the defining conditions* (1.3) *and* (1.4) *for* $L(s, \rho)$ *over* $\mathbb{R}$ *in* §1, *and the two definitions of* $\varepsilon(s, \rho, \psi)$ *in* (3.8) *and* (1.8) *coincide.*

4. Classification and correspondence for $\mathrm{GL}_n(\mathbb{C})$

There is a corresponding theory for $G = \mathrm{GL}_n(\mathbb{C})$, quite a bit less complicated. **Admissible representations** are defined for this G as in §2, but with $K = U(n)$ playing the role of maximal compact subgroup. The notions of **irreducible** representation, K **finite** vector, **infinitesimally equivalent**, and **central character** are defined as in §2.

For z in $\mathbb{C}$, let $[z] = z/|z|$. Also recall that $|z|_{\mathbb{C}} = |z|^2$. The building blocks for irreducible admissible representations of $\mathrm{GL}_n(\mathbb{C})$ are the representations of $\mathrm{GL}_1(\mathbb{C})$ given by

$$(4.1) \qquad z \to [z]^l |z|_{\mathbb{C}}^t \quad \text{with } l \in \mathbb{Z} \text{ and } t \in \mathbb{C},$$

which we write as $[\cdot]^l \otimes |\cdot|_{\mathbb{C}}^t$. For each j with $1 \le j \le n$, let σ_j be the representation $[\cdot]^{l_j} \otimes |\cdot|_{\mathbb{C}}^{t_j}$ of $\mathrm{GL}_1(\mathbb{C})$. Then $(\sigma_1, \dots, \sigma_n)$ defines a one-dimensional representation of the diagonal subgroup of $\mathrm{GL}_n(\mathbb{C})$, and we extend this to a one-dimensional representation of the upper triangular subgroup B. We set

$$(4.2) \qquad I(\sigma_1, \dots, \sigma_n) = \mathrm{ind}_B^G(\sigma_1, \dots, \sigma_n),$$

using unitary induction.

THEOREM 4. *For* $G = \mathrm{GL}_n(\mathbb{C})$,
(a) *if the parameters* t_j *of* $(\sigma_1, \dots, \sigma_n)$ *satisfy*

$$(4.3) \qquad \mathrm{Re}\, t_1 \ge \mathrm{Re}\, t_2 \ge \cdots \ge \mathrm{Re}\, t_n,$$

then $I(\sigma_1, \dots, \sigma_n)$ *has a unique irreducible quotient* $J(\sigma_1, \dots, \sigma_n)$,
(b) *the representations* $J(\sigma_1, \dots, \sigma_n)$ *exhaust the irreducible admissible representations of* G, *up to infinitesimal equivalence,*
(c) *two such representations* $J(\sigma_1, \dots, \sigma_n)$ *and* $J(\sigma_1', \dots, \sigma_n')$ *are infinitesimally equivalent if and only if there exists a permutation* $j(i)$ *of* $\{1, \dots, n\}$ *such that* $\sigma_i' = \sigma_{j(i)}$ *for* $1 \le i \le n$.

This theorem predates the Langlands classification and is due to Želobenko and Naĭmark **[36, 37]**. For an exposition, see Duflo **[4]**. The quotient $J(\sigma_1, \dots, \sigma_n)$ may be described within $I(\sigma_1, \dots, \sigma_n)$ in the same two ways as in the case of $\mathrm{GL}_n(\mathbb{R})$.

The Weil group of $\mathbb{C}$, denoted $W_{\mathbb{C}}$, is given by $W_{\mathbb{C}} = \mathbb{C}^\times$. As in the case of $\mathbb{R}$, we shall be interested in the set of equivalence classes of n-dimensional complex representations of $W_{\mathbb{C}}$ whose images consist of semisimple elements.

Since $\mathbb{C}^\times$ is abelian, such a representation φ is diagonable and hence is the direct sum of one-dimensional representations. A one-dimensional representation is necessarily of the form (3.1). But now that conjugation no longer plays any role, it will be more convenient to write it as

$$(4.4) \qquad (l, t)\colon z \to [z]^l |z|_{\mathbb{C}}^t \quad \text{with } l \in \mathbb{Z} \text{ and } t \in \mathbb{C}.$$

Let φ be an n-dimensional semisimple complex representation of $W_{\mathbb{C}}$. We can write φ as a direct sum of one-dimensional representations φ_j with

$\varphi_j(z) = [z]^{l}|z|_{\mathbb{C}}^{t_j}$ in the notation of (4.4). To φ_j we associate the representation $\sigma_j = [\cdot]^{l_j}|\cdot|_{\mathbb{C}}^{t_j}$ of $\mathrm{GL}_1(\mathbb{C})$. In this way, we associate a tuple $(\sigma_1, \ldots, \sigma_n)$ of representations to φ. If the complex numbers $t_1, \ldots, t_n$ do not satisfy (4.3), we permute $(\sigma_1, \ldots, \sigma_n)$ so that (4.3) ends up being satisfied. Using Theorem 4, we can then make the association

$$\varphi \longrightarrow \rho_{\mathbb{C}}(\varphi) = J(\sigma_1, \ldots, \sigma_r) \tag{4.5}$$

and come to the following conclusion.

THEOREM 5 (Local Langlands Correspondence for $\mathrm{GL}_n(\mathbb{C})$). *The association* (4.5) *is a well-defined bijection between the set of all equivalence classes of n-dimensional semisimple complex representations of $W_{\mathbb{C}}$ and the set of all equivalence classes of irreducible admissible representations of* $\mathrm{GL}_n(\mathbb{C})$.

The local L factor corresponding to a one-dimensional representation φ of $W_{\mathbb{C}}$ is

$$L(s, \varphi) = 2(2\pi)^{-(s+t+\frac{|l|}{2})}\Gamma(s+t+\tfrac{|l|}{2}) \quad \text{if } \varphi \text{ is given by } (l, t) \text{ in (4.4).} \tag{4.6}$$

For φ reducible, $L(s, \varphi)$ is the product of the L factors of the irreducible constituents of φ.

Fix the additive character ψ of $\mathbb{C}$ in (1.5), so that twice the ordinary Lebesgue measure is the self-dual Haar measure on $\mathbb{C}$. The ε factors are given for φ one-dimensional by

$$\varepsilon(s, \varphi, \psi) = i^{|l|} \qquad \text{if } \varphi \text{ is given by } (l, t) \text{ in (4.4).} \tag{4.7}$$

For φ reducible, $\varepsilon(s, \varphi, \psi)$ is the product of the L factors of the irreducible constituents of φ. As was noted in §3, formula (4.7) and the third part of (3.7) are connected by the fact that ε is "inductive in degree 0."

We can now define local factors $L(s, \rho)$ and $\varepsilon(s, \rho, \psi)$ for each irreducible admissible representation of $\mathrm{GL}_n(\mathbb{C})$ by the rule

$$\left.\begin{aligned} L(s, \rho) &= L(s, \varphi) \\ \varepsilon(s, \rho, \psi) &= \varepsilon(s, \varphi, \psi) \end{aligned}\right\} \quad \text{if } \rho = \rho_{\mathbb{C}}(\varphi) \text{ in (4.5) and Theorem 5.} \tag{4.8}$$

Jacquet **[12]** proved the following result.

THEOREM 6. *The definition of $L(s, \rho)$ in* (4.8) *satisfies the defining conditions* (1.3) *and* (1.4) *for $L(s, \rho)$ over $\mathbb{C}$ in* §1, *and the two definitions of $\varepsilon(s, \rho, \psi)$ in* (4.8) *and* (1.8) *coincide.*

5. Results for other reductive groups

Some of the constructions and results in §§2–3 extend from GL_n to arbitrary connected reductive groups G defined over $\mathbb{R}$. For such a group G, we shall work in the context of the representation theory of $G(\mathbb{R})$. Langlands **[21]** reduced the classification of irreducible admissible representations of $G(\mathbb{R})$ to classification of the subset of irreducible tempered representations, which in turn were classified in **[16]**. Langlands **[21]** showed also that the classification for $G(\mathbb{R})$ fits into a framework that is consistent with what was described for $\mathrm{GL}_n(\mathbb{R})$ in §3. In this section we shall discuss aspects of that framework. For full details one can consult the excellent exposition by Borel **[1]**.

The first step is to introduce the L group of $G/\mathbb{R}$. The identity component ${}^LG^0$ is a certain complex reductive group depending only on G (not $G/\mathbb{R}$), having the same rank as G, and having root system dual to that of G. The exact definition requires some care, but it has the following features:

(a) if $G = \mathrm{GL}_n$, then ${}^LG^0 = \mathrm{GL}_n(\mathbb{C})$;
(b) if G is simply connected, then ${}^LG^0$ is an adjoint group;
(c) if G is an adjoint group, then ${}^LG^0$ is simply connected.

Now we bring in $\mathbb{R}$, letting $\Gamma = \mathrm{Gal}(\mathbb{C}/\mathbb{R}) \cong \mathbb{Z}/2\mathbb{Z}$. The definition of ${}^LG^0$ is such that the action of Γ on G yields an action of Γ on ${}^LG^0$. Then ${}^LG$ is defined as the corresponding semidirect product of ${}^LG^0$ with Γ. This semidirect product is a direct product if $G(\mathbb{R})$ is a split group, as is the case for $G(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R})$.

By a **representation** of ${}^LG$ is meant any holomorphic homomorphism of ${}^LG$ into some $\mathrm{GL}_m(\mathbb{C})$. An element of ${}^LG$ is **semisimple** if its image is semisimple under every representation of ${}^LG$. Certain parabolic subgroups described in **[1]** are defined to be **relevant**; all parabolic subgroups of ${}^LG$ are relevant if $G(\mathbb{R})$ is split or quasisplit.

A continuous homomorphism $\varphi : W_{\mathbb{R}} \to {}^LG$ is said to be **admissible** if

(a) $\varphi(\mathbb{C}^\times) \subseteq {}^LG^0$, and $\varphi(j)$ is contained in the nontrivial coset of ${}^LG^0$ in ${}^LG$;
(b) $\varphi(W_{\mathbb{R}})$ is contained in the set of semisimple elements of ${}^LG$;
(c) whenever $\varphi(W_{\mathbb{R}})$ is contained in the Levi subgroup of a parabolic subgroup P of ${}^LG$, then P is relevant.

The set of such φ's, modulo the equivalence relation defined by conjugacy within ${}^LG^0$, is denoted $\Phi(G)$. For $G(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R})$, Theorem 2 says that $\Phi(G(\mathbb{R}))$ parametrizes the set $\Pi(G(\mathbb{R}))$ of equivalence classes of irreducible admissible representations of $G(\mathbb{R})$. This statement needs adjustment for more general $G(\mathbb{R})$. What happens is that one associates to each admissible φ a finite subset Π_φ of $\Pi(G(\mathbb{R}))$ called an L **packet**. Inequivalent φ's lead to disjoint subsets, and the union of all Π_φ is all of $\Pi(G(\mathbb{R}))$.

Representations in the same Π_φ are called L-**indistinguishable**. Two phenomena contribute to this notion: discrete series with the same infinitesimal character, and reducibility of standard induced tempered representations. Shelstad **[27]** has quantified the statement that these are the only contributing factors, and she has explored the consequences of L-indistinguishability.

Local L and ε factors are associated not to irreducible admissible representations ρ of $G(\mathbb{R})$ but to pairs (ρ, r) in which r is a representation of ${}^LG$ in some $\mathrm{GL}_m(\mathbb{C})$. The definition is as follows: Given ρ, we find φ such that the class of ρ lies in Π_φ. Then $r \circ \varphi$ is an m-dimensional semisimple representation of $W_{\mathbb{R}}$, and we can put

$$L(s, \rho, r) = L(s, \varphi),$$
$$\varepsilon(s, \rho, r, \psi) = \varepsilon(s, \varphi, \psi),$$

in the notation of (3.6) and (3.7). When $G(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R})$, these definitions reduce to (3.8) if r is the standard representation of $\mathrm{GL}_n(\mathbb{C})$.

In the Langlands program, one expects that the above local correspondence for $G(\mathbb{R})$ will be valid in some form over non-Achimedean local fields as well, that there will be a global theory, and that the local and global theories will mesh in the same way that they appear to mesh for GL_n. For accounts of progress in these matters, see Gelbart and Shahidi [**8**] and Shahidi [**26**].

The full power of the Langlands program comes into play only when the functoriality of these constructions is considered. Almost all aspects of functoriality are still conjectural. A general statement of functoriality, accompanied by an overview of its consequences, is in Gelbart [**7**]. More detail may be found in Chapter V of Borel [**1**].

REFERENCES

1. A. Borel, *Automorphic L-functions*, Automorphic Forms, Representations, and L-functions, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 27–61.
2. P. Deligne, *Formes modulaires et représentations de* GL(2), Modular Functions of One Variable. II, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973, pp. 55–105.
3. ——, *Les constantes des équations fonctionnelles des fonctions* L, Modular Functions of One Variable. II, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973, pp. 501–597.
4. M. Duflo, *Représentations irréductibles des groupes semi-simples complexes*, Analyse Harmonique sur les Groupes de Lie, Lecture Notes in Math., vol. 497, Springer-Verlag, Berlin and New York, 1975, pp. 26–88.
5. D. Flath, *Decomposition of representations into tensor products*, Automorphic Forms, Representations, and L-functions, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 179–183.
6. S. S. Gelbart, *Automorphic forms on adèle groups*, Ann. of Math. Stud., vol. 83, Princeton Univ. Press, Princeton, NJ, 1975.
7. ——, *An elementary introduction to the Langlands program*, Bull. Amer. Math. Soc. **10** (N. S.) (1984), 177–219.
8. S. Gelbart and F. Shahidi, *Analytic properties of automorphic L-functions*, Perspect. Math., vol. 6, Academic Press, San Diego, CA, 1988.
9. I. M. Gelfand, M. I. Graev, and I. I. Pyatetskiĭ-Shapiro, *Representation theory and automorphic functions*, Saunders, Philadelphia, 1969.
10. R. Godement and H. Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Math., vol. 260, Springer-Verlag, Berlin and New York, 1972.
11. Harish-Chandra, *Representations of a semisimple Lie group on a Banach space*. I, Trans. Amer. Math. Soc. **75** (1953), 185–243.

12. H. Jacquet, *Principal L-functions of the linear group*, Automorphic Forms, Representations, and L-functions, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 63–86.
13. H. Jacquet and R. P. Langlands, *Automorphic forms on* GL(2), Lecture Notes in Math., vol. 114, Springer-Verlag, Berlin and New York, 1970.
14. A. W. Knapp, *Representation theory of semisimple groups: An overview based on examples*, Princeton Univ. Press, Princeton, NJ, 1986.
15. A. W. Knapp and E. M. Stein, *Intertwining operators for* $\mathrm{SL}(n, \mathbb{R})$, Studies in Mathematical Physics: Essays in Honor of Valentine Bargmann, Princeton Univ. Press, Princeton, NJ, 1976, pp. 239–267.
16. A. W. Knapp and G. Zuckerman, *Classification theorems for representations of semisimple Lie groups*, Non-Commutative Harmonic Analysis, Lecture Notes in Math., vol. 587, Springer-Verlag, Berlin and New York, 1977, pp. 138–159.
17. ——, *Classification of irreducible tempered representations of semisimple groups*, Ann. of Math. (2) **116** (1982), 389–501. (See also **119** (1984), 639.)
18. S. Kudla, *Local Langlands correspondence: The non-Archimedean case*, these Proceedings, vol. 2, pp. 365–391.
19. R. P. Langlands, *Euler products*, James K. Whittemore Lectures, bound notes, Yale Univ., 1967.
20. ——, *Problems in the theory of automorphic forms*, Lectures in Modern Analysis and Applications III, Lecture Notes in Math., vol. 170, Springer-Verlag, Berlin and New York, 1970, pp. 18–61.
21. ——, *On the classification of irreducible representations of real algebraic groups*, Representation Theory and Harmonic Analysis on Semisimple Lie Groups (P. J. Sally and D. A. Vogan, eds.), Math. Surveys Monographs, vol. 31, Amer. Math. Soc., Providence, RI, 1989, pp. 101–170.
22. I. Mirković, *Classification of irreducible tempered representations of semisimple groups*, Ph.D. dissertation, Univ. Utah, Salt Lake City, 1986.
23. I. Piatetski-Shapiro, *Classical and adelic automorphic forms. An introduction*, Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 185–188.
24. A. Robert, *Formes automorphes sur* GL_2, Sém. Bourbaki, vol. 1971/72, Exp. 400–417, Lecture Notes in Math., vol. 317, Springer-Verlag, Berlin and New York, 1973, pp. 295–318.
25. F. Shahidi, *Local coefficients as Artin factors for real groups*, Duke Math. J. **52** (1985), 973–1007.
26. ——, *Automorphic L-functions: A survey*, Automorphic Forms, Shimura Varieties, and L-functions. I (Proc. Conf. Univ. Michigan, Ann Arbor, 1988) Perspect. Math., vol. 10, Academic Press, San Diego, 1990, pp. 415–437.
27. D. Shelstad, *L-indistinguishability for real groups*, Math. Ann. **259** (1982), 385–430.
28. J. T. Tate, *Fourier analysis in number fields and Hecke's zeta functions*, Ph.D. dissertation, Princeton Univ., Princeton, NJ, 1950; Algebraic Number Theory (J. W. S. Cassels and A. Fröhlich, eds.), Academic Press, London, 1967, pp. 305–347.
29. J. Tate, *Number theoretic background*, Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 3–26.
30. D. A. Vogan, *The algebraic structure of the representation of semisimple Lie groups*. I, Ann. of Math. (2) **109** (1979), 1–60.
31. ——, *The unitary dual of* GL(n) *over an Archimedean field*, Invent. Math. **83** (1986), 449–505.
32. ——, *Unitary representations of reductive Lie groups*, Ann. of Math. Stud., vol. 118, Princeton Univ. Press, Princeton, NJ, 1987.
33. N. R. Wallach, *Representations of reductive Lie groups*, Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 71–86.
34. A. Weil, *Sur la théorie du corps de classes*, J. Math. Soc. Japan **3** (1951), 1–35.

35. ——, *Dirichlet series and automorphic forms*, Lecture Notes in Math., vol. 189, Springer-Verlag, Berlin and New York, 1971.
36. D. P. Želobenko, *Harmonic analysis on semisimple complex Lie groups*, "Nauka", Moscow, 1975. (Russian)
37. D. P. Želobenko and M. A. Naĭmark, *A characterization of completely irreducible representations of a semisimple complex Lie group*, Soviet Math. Dokl. 7 (1966), 1403–1406.

STATE UNIVERSITY OF NEW YORK, STONY BROOK, NEW YORK

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Pure Motives and Automorphic Forms

DINAKAR RAMAKRISHNAN

Contents

0. Introduction
1. Tensor categories; finite groups
2. Automorphic forms on GL_n
3. L-functions of $\mathrm{GL}(n)$: Properties and conjectures
Bibliography

0. Introduction

Let M be a pure motive over a number field F of rank n, weight w and with coefficients in $E \subseteq \mathbb{C}$. To fix ideas let us think of M as a direct factor of the cohomology of a smooth projective variety X over F cut out by an algebraic (or absolutely Hodge) correspondence. A powerful principle, propounded in general by Langlands **[La1, La2]**, implies that M should be associated to an automorphic representation π (or an automorphic form f in a more traditional language) of $\mathrm{GL}(n, \mathbb{A}_F)$, where $\mathbb{A}_F = F_\infty \times \mathbb{A}_{f,F}$ denotes the ring of adèles of F. ($\mathbb{A}_{f,F} = \mathbb{Q} \otimes \hat{o}_F$ is the ring of finite adèles of F.) The correspondence $M \to \pi$ should moreover satisfy:

$$L(M, s) = L(\pi, s)$$

with $\pi \otimes |\det(\cdot)|^{-w/2}$ unitary. (See also **[De4]** and **[Se**, vol. II, no. 78].)

Here $L(\pi, s)$ denotes the standard L-function of π. When M is absolutely irreducible, π should be cuspidal. Conversely, every cuspidal π that is algebraic **[Cl1]** is expected to be attached to a motive M. This bijection should match Artin motives M (of weight zero) with cuspidal π of *Galois type* (see §2).

When $n = 1$, this principle is, for Artin motives, a consequence of (abelian) class field theory, as automorphic forms on $\mathrm{GL}(1, \mathbb{A}_F)$ of Galois

1991 *Mathematics Subject Classification.* Primary 11F, 11G, 22E.
This paper is in final form and no version of it will be submitted for publication elsewhere.

type are none other than characters χ of the idèle class group $\mathbb{A}_F^*/F^*$ which are trivial on the connected component of F_∞^*.

It is natural then to hope for the existence of a (Langlands) group $\mathscr{L}_F$, whose n-dimensional representations parametrize "nice" automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$. Fixing an algebraic closure $\overline{F}$ of F, we would like to have a surjective homomorphism $\varphi\colon \mathscr{L}_F \to \mathrm{Gal}(\overline{F}/F)$, whose kernel $\mathscr{L}_F^0$ is a connected pro-reductive group $/\mathbb{C}$, and $\mathscr{L}_F$ should be uniquely defined up to inner conjugacies by $\mathscr{L}_F^0$. Moreover, one wishes for a surjective homomorphism $\mu\colon \mathscr{L}_F \to \mathfrak{G}_F$, where $\mathfrak{G}_F$ is the motivic Galois group of (pure) motives over F with coefficients *extended* to $\mathbb{C}$ relative to absolutely Hodge correspondences [**De-M,** §6]. The conjectural map $M \mapsto \pi$ should then correspond to pulling back the representations of $\mathfrak{G}_F$ via μ. It is a big open problem to prove the existence of $\mathscr{L}_F$ and it is not evident if it should come out of a Tannakian formalism. In any case it is now clear, thanks to [**La2**] and [**J-S1**], that the right notion of niceness for an automorphic representation π is for it to be *isobaric* (see §2.1 below). Perhaps the most important thing is that the formalism of $\mathscr{L}_F$ leads to many predictions about automorphic forms and about their relationship to Galois representations. A striking and well-known consequence is that Artin's conjecture will follow if one establishes automorphic induction corresponding to $\chi \mapsto \mathrm{Ind}(\mathscr{L}_F, \mathscr{L}_E; \chi)$ for *arbitrary* finite extensions E/F and one-dimensional representations χ of $\mathscr{L}_E$ (see §3.2 below). We have tried to give in §3 other instances of such predictive power as well. We do not however touch upon Arthur's extension of $\mathscr{L}_F$ by $\mathrm{SU}(2)$ whose representations have led him to give precise conjectures on the discretely occurring automorphic forms on any reductive group G/F ([**A**]; see also [**Pic, B-Ro2**]).

The major obstruction to having a Tannakian formalism for $\mathscr{L}_F$ is the lack of a tensor structure on a suitable category of automorphic forms on GL. Given cuspidal automorphic representations π, π' of $\mathrm{GL}(n, \mathbb{A}_F)$ and $\mathrm{GL}(n', \mathbb{A}_F)$ respectively, one knows how to construct an admissible representation $\pi \boxtimes \pi'$, at least outside the ramified places, of $\mathrm{GL}(nn', \mathbb{A}_F)$. The first main problem is to prove that it is *automorphic*. Even admitting that, one needs further information to have a tensor category, and it is simply not sufficient to work with isomorphism classes of automorphic representations. As suggested by the referee, we begin §1 by discussing the concrete problems one encounters in defining a tensor category. It is then followed by a short account of the additional structures, namely rigidity, and identity 1 with $\mathrm{End}(1) \simeq \mathbb{C}$ and a fiber functor, needed to associate a group to this category. We end §1 with a brief discussion of the additional invariants one needs besides the character table to determine a finite group G.

In §2 after introducing the basic structures of automorphic forms we employ a formal device to define a $\mathbb{C}$-linear abelian semisimple category $\underline{A}(F)$, built up out of cuspidals. This is done simply to sidestep some inherent ambiguities in the definition of isobaric categories in **[La2]**, which is addressed in a different way in **[Cl1]**. It should be stressed however that there is *no* construction of a tensor category here. After a discussion of some of the basic questions, we consider S-versions of $\underline{A}(F)$ and construct, for each $v \notin S$, a $\mathbb{C}$-linear faithful and exact functor η_v on $\underline{A}^S(F)$ into a category $\underline{A}^{\mathrm{ur}}(F_v)$ of local unramified representations. This is done by a simple application of the existence of "new" subspaces, and again *nothing* major is claimed here. It is tempting however to wonder if a finer version of η_v could (ever) lead to a fiber functor, assuming of course a (highly hypothetical) tensor structure on $\underline{A}^S(F)$. We end §2 with a statement of the Ramanujan (or purity) conjecture for $\mathrm{GL}(n)$.

In §3 we discuss automorphic L-functions for $\mathrm{GL}(n)$ and first expose some results of Jacquet, Shalika, Piatetski-Shapiro, et al. This leads to some comments and questions from the point of view of $\mathscr{L}_F$. In particular, we describe in §3.3 the converse theorem approach to proving the automorphy of $\pi_1 \boxtimes \pi_2$ for cuspidals π_1, π_2. We conclude by posing (in §3.4) a conjecture, with some evidence, giving a refinement of the strong multiplicity-one theorem. To be precise, we suggest that if two cuspidal automorphic representations π, π' of $\mathrm{GL}_n(\mathbb{A}_F)$ agree at all the places outside a set S of primes of density $< 1/2n^2$, then π and π' are globally isomorphic.[1] There is an evident Galois analog, which is easy for representations with finite image.

Before beginning the main subject matter, it may be useful for us to indicate how, even if one forgot about the Tannakian problems and the Langlands group, the basic problem of identifying a motivic L-function with an automorphic one is incredibly hard, already for $n = 2$ and $F = \mathbb{Q}$.[2] (Indeed, most of the positive results so far have been in the other direction, namely that of associating to a class of algebraic automorphic forms the corresponding ℓ-adic and Hodge realizations of motives.) To be specific, take M to be the motive defined by H^1 of an elliptic curve E over $\mathbb{Q}$ of conductor N. (The conjecture in this motivating historic case is associated with the names of Taniyama, Shimura, and Weil.) Let us look at this instructive case a bit more closely. We have

$$L(M, s) = 2(2\pi)^{-2s}\Gamma(s)\prod_p P_p(M, p^{-s})^{-1}$$

where

$$P_p(M, T) = (1 - a_pT + \delta(p)pT^2)$$

with

$$a_p = 1 + p - \nu_p, \qquad \nu_p = \text{no. of points of } E \bmod p,$$

[1] We have come to know recently that this question has been considered earlier by J.-P. Serre **[Se**, vol. III, pp. 323–324].

[2] As this article goes into press, Andrew Wiles has announced the following spectacular theorem: *Every semistable elliptic curve over* $\mathbb{Q}$ *is modular.*

and

$$\delta(p) = 0 \quad \text{if } p|N \text{ and } 1 \text{ otherwise}.$$

When N is not divisible by p, the reduction of E modulo p is nonsingular and, by the Riemann hypothesis for curves over finite fields, one has: $a_p = \alpha_p + \overline{\alpha}_p$, with $|\alpha_p| = |\overline{\alpha}_p| = \sqrt{p}$. When $p|N$, a_p is 1 (resp. -1, resp. 0) if E has split multiplicative reduction (resp. nonsplit multiplicative reduction, resp. additive reduction) at p. (For $p \geq 3$, E is semistable at p iff E does not have additive reduction at p.) Consequently, $L(M, s)$ converges absolutely in the right half-plane $\{\text{Re}(s) > 3/2\}$. The conjecture associates to M (in the classical language) a normalized new form f of level N and weight 2 on the upper half-plane $\mathfrak{H}$, having eigenvalue a_p relative to the pth Hecke operator T_p for every good p. (Here, being a new form of level N signifies that f is a cusp form relative to the group $\Gamma_0(N)$ of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\text{SL}(2, \mathbb{Z})$ such that $N|c$, and that it does not come from any level smaller than N; being normalized means that the q-expansion of f has its first coefficient equal to 1.) It is not hard to see that f defines, and is determined by, a smooth function φ on $\text{GL}(2, \mathbb{A})$, left-invariant by $\text{GL}(2, \mathbb{Q})Z(\mathbb{A})$, where Z denotes the center of $\text{GL}(2)$ and right-invariant by a compact open subgroup $K_0(N)$ of $\text{GL}(2, \mathbb{A}_f)$ (such that $\Gamma_0(N) = K_0(N) \cap \text{GL}(2, \mathbb{Q})^+$). The right $\text{GL}(2, \mathbb{A})$-span of φ in $L^2(\text{GL}(2, \mathbb{Q})Z(\mathbb{A})\backslash \text{GL}(2, \mathbb{A}))$ defines an irreducible, invariant subspace π. One knows that, for any Dirichlet character χ, the twisted L-series $L(f, \chi, s)$ $(= L(\pi \otimes \chi, s))$ defines an entire function of $\mathbb{C}$. So the conjecture implies the same for $L(M \otimes \chi, s)$, which, if known, will have many consequences. It should be noted that if one knew directly somehow that these twisted L-functions of M are all entire and bounded in vertical strips with functional equations, then the converse theorem of Weil and Jacquet-Langlands **[J-L]** would tell us that $L(M, s)$ is the Euler product of a weight 2 new form f of the right level.

The situation is much better going the other way. In fact, Shimura has shown **[Sh1]** how to associate a motive M_f over $\mathbb{Q}$ of weight 1 starting with a new form f of weight 2 and level N. Indeed, the associated holomorphic differential $2\pi i f(z)\,dz$ on $\mathfrak{H}$ descends to the quotient Riemann surface $\Gamma_0(N)\backslash\mathfrak{H}$. It also extends to the compactification, which has a model $X_0(N)$ over $\mathbb{Q}$. It is a fact that the q-expansion of f is rational over a (totally real) number field T, and one gets an algebraic 1-form $\omega(f)$ over T on $X_0(N)$. The span of the set of conjugates $\omega(f^\tau)$ of $\omega(f)$ (under conjugation by τ in $\text{Aut}(\mathbb{C})$) allows one to obtain a corresponding abelian variety factor A_f over $\mathbb{Q}$ of dimension $[T : \mathbb{Q}]$, up to $\mathbb{Q}$-isogeny, of the Jacobian of $X_0(N)$. Shimura showed that $L(A_f, s) = \prod_\tau L(f^\tau, s)$ at least up to a finite number of Euler factors. Building on the works of Igusa, Ihara, Deligne, Langlands, and Piatetski-Shapiro, Carayol **[Ca]** then showed that the equality holds everywhere. M_f is $H^1(A_f)$. When $T = \mathbb{Q}$, A_f is evidently an elliptic curve that has rank two over the coefficient field T.

When the conjectural arrow $E \to \pi$ (or f) exists, E and A_f will be isogenous over $\mathbb{Q}$ by Faltings **[F]** since they have the same L-functions (and ℓ-adic representations). Such a modular geometric realization of E is very useful in studying the arithmetic of E. In fact, according to a striking recent theorem of Ribet **[Ri1]** (building on ideas of G. Frey and J.-P. Serre), its existence for a class of E implies Fermat's last theorem.[3]

Deligne has associated **[De1]** to any new form f of weight $k \geq 2$, and ℓ-adic representation σ_ℓ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ (of rank 2 over $T \otimes \mathbb{Q}_\ell$), whose L-function agrees outside the conductor with that of f. Moreover, this representation occurs in H^{k-1} of a Kuga-Sato variety over the modular curve, and the traces of Frobenius obey the Weil estimate proved (in general) by Deligne. One of the ingredients of this work is the Eichler-Shimura isomorphism, which identifies the "parabolic" H^1 of the modular curve with coefficients in the local system defined by the symmetric $(k-2)$th power of the standard representation of $\mathrm{SL}(2)/\mathbb{Q}$ with the space of cusp forms of weight k plus its complex conjugate. (Parabolic cohomology is the simplest instance of the middle intersection cohomology.) The irreducibility of σ_ℓ was then established by Ribet **[Ri2]**. (A totally different construction of σ_ℓ for forms of weight ≥ 2 was given, under an ordinarity hypothesis, by Hida **[Hi]** by specializing suitable Galois representations into $\mathrm{GL}_2(\mathbb{Z}_\ell[[X]])$.) When f is a new form of weight 1, a theorem of Deligne and Serre **[De-Se]** attaches an irreducible two-dimensional Galois representation over $\mathbb{C}$ with finite image. Their method consisted in using congruences with higher-weight modular forms, analyzing the images under the representations $\bmod p$, and lifting to characteristic zero.

An important consequence of the above theorems is the proof of the conjecture of Ramanujan-Petersson for f, namely: $|a_p| \leq 2p^{(k-1)/2}$. For $k \geq 3$, that σ_ℓ is the ℓ-adic realization of a Grothendieck motif M_f was later deduced by Scholl **[Sch]**.

When F is totally real of *odd* degree over $\mathbb{Q}$, the situation is essentially the same, and the A_f's (in the weight 2 Hilbert modular case) occur as factors of Jacobians of suitable Shimura curves. (For weight 1, the analog of Deligne-Serre was done by Rogawski-Tunnel **[Ro-T]**; the relevant result on points mod p of the Shimura varieties defined by quaternion algebras over F has recently been established by Reimann-Zink **[R-Z]**. See also **[Oh]**.) When $[F : \mathbb{Q}]$ is even, completely new ideas are required. Now one has the Galois representation σ_ℓ by Wiles **[Wi]** (weight 1) and Taylor **[T1]** (weight ≥ 2), and the Ramanujan conjecture (weight ≥ 2) by Brylinski-Labesse. By the work of Blasius-Rogawski **[B-Ro]**, one has a very different construction, which gives insight even into the noncongruent weights mod 2 and associates (in the non-weight 2 case) Grothendieck motives. In the weight 2 case, to bring things up to the level of the situation over $\mathbb{Q}$, one still needs to solve the

[3] By the recent result of Wiles alluded to earlier, Fermat's last theorem is now a theorem!

following problems:

(i) construct (up to F-isogeny) the abelian variety A_f over F;

(ii) exhibit, as predicted by the Tate conjecture, an algebraic correspondence (modulo homological equivalence) between the relevant Hilbert modular variety and the abelian variety obtained by restriction of scalars to $\mathbb{Q}$ of A_f.

When f is associated to a *Hecke* character of a CM-quadratic extension of F so that $L(f, s) = L(\chi, s)$ as Euler products over F, (i) has a solution by the works of Shimura, Taniyama and Casselman **[Sh-T, Sh2]**. For general f, Takayuki Oda **[Od]** has indicated a procedure to define an abelian variety A_f^0 over $\mathbb{C}$ which remains to be proven to be isogenous to one over an arithmetic field. It should be noted that, when the automorphic representation π associated to f has a special or a supercuspidal component at some finite place v_0, then π corresponds to an automorphic representation π' of the (adèles of) the multiplicative group of a quaternion algebra over F which is ramified at all but one infinite place. This way one can realize A_f using **[Ca]** as a factor of a Jacobian of the Shimura curve associated to B, solving the problem in this case. One also knows in this case **[M-R]** that A_f^0 is isogenous to $A_f \times \mathbb{C}$.

The above deals with irreducible motives M of rank 2 over (totally real) F that are *totally odd*, i.e., where complex conjugation c acts by determinant -1 at each infinite place η. Suppose M is *totally even.* Then c acts by a scalar at each η, and this forces the Hodge structure of $M_{\mathrm{dR}} \otimes_{F,\eta} \mathbb{C}$ to be purely of type (p, p) for some integer p. Then the Tate-twisted motif $M(p)$ is of weight zero. For any Artin motif of rank 2 over F that is even at some η, the corresponding π should be a nonholomorphic cusp form which, when F is $\mathbb{Q}$, should be of Maass type and have eigenvalue $\frac{1}{4}$ relative to the hyperbolic Laplacian on $\mathfrak{H}$. Such automorphic forms of "Galois" type (of all parity distributions), together with the holomorphic Hilbert modular new forms (of mixed weights, all congruent mod 2, and with character), are expected to comprise the entire set of motivic cusp forms π on $\mathrm{GL}(2)/F$.

For forms of "weight two" over imaginary quadratic fields, one now has a construction of the conjectural ℓ-adic representations **[T2, H-S-T]**.

We end this long rumination on the $n = 2$ case by noting the following result of Langlands **[La3]**, as strengthened by Tunnell **[Tu]**:

Given any Artin Motif M of rank 2 (of *any* parity) over a number field F such that the associated Galois representation has solvable image in $\mathrm{GL}(2, \mathbb{C})$, there is an automorphic representation π of $\mathrm{GL}(2)/F$ such that the L-functions agree.[4]

For M Artin of *any* rank n with *nilpotent* Galois image, the work of Arthur-Clozel associates a corresponding π.

[4]This theorem implies in particular that every odd irreducible representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ onto $\mathrm{GL}_2(\mathbb{F}_3)$ corresponds to a holomorphic cusp form of weight 1.

In conclusion, we encourage the reader to consult the papers of Langlands **[La1, La2]** and of Clozel **[Cl1]** for a detailed discussion of other relevant questions. We shall try to minimize repetition from those sources and instead stress some different aspects. The idea that automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$, all $n \geq 1$, should collectively behave like the representation ring of a group appeared to have also occurred (independently) to J. A. Shalika. Geometers who may be familiar with automorphic forms only as sections of suitable vector bundles will benefit by looking at **[Bo1]** and **[Bo-J]** for understanding the passage to the point of view of automorphic representations used below. We see no point in our repeating such things in this article.

This article would not have come to exist if not for the prodding of the organizers, especially Uwe Jannsen and R. MacPherson, whom we thank. It is a pleasure to acknowledge some fruitful discussions with P. Deligne and with J.-P. Serre during the motives conference.

We thank Deligne for mentioning the problem of constructing a local component functor on automorphic categories, and for referring us to Gallagher's article **[Ga]**. We are also grateful to the referee for valuable suggestions, such as the inclusion of Lemma 1.1.2 and the material in §1.4, leading to improvement over the first version. We thank D. Grayson for helpful conversations and criticism concerning certain categorical questions and constructions. Thanks are also due to F. Bogomolov, D. Blasius, H. Jacquet, J. Rogawski, R. Taylor, and T. Zink for helpful comments, and to R. P. Langlands for encouragement and suggestions. Finally we are pleased to acknowledge the hospitality of the University of Bielefeld where the first version of the article was written and then typeset by Frau Rahner.

1. Tensor categories; finite groups

The principal references for this section are **[De-M]** and **[De2]**.

1.1. A *tensor category* is (according to **[De-M]**) a category C equipped with the following:

(1.1.1)

(a) a functor $\otimes: C \times C \to C$, $(X_1, X_2) \mapsto X_1 \otimes X_2$;

(b) functorial isomorphisms:

$$a = a_{X_1, X_2, X_3}: X_1 \otimes (X_2 \otimes X_3) \xrightarrow{\sim} (X_1 \otimes X_2) \otimes X_3$$

for all objects X_1, X_2, X_3 (associativity); and

$$c = c_{X_1, X_2}: X_1 \otimes X_2 \xrightarrow{\sim} X_2 \otimes X_1$$

for all objects X_1, X_2 (commutativity);

(c) constraints: (for all objects X_1, X_2, X_3, X_4)

(i) $c_{X_2, X_1} \circ c_{X_1, X_2} = \mathrm{id}_{X_1 \otimes X_2}$.
The following diagrams commute:

(ii)

$$\begin{array}{ccccc} X_1 \otimes (X_2 \otimes (X_3 \otimes X_4)) & \xrightarrow{a} & (X_1 \otimes X_2) \otimes (X_3 \otimes X_4) & \xrightarrow{a} & ((X_1 \otimes X_2) \otimes X_3) \otimes X_4 \\ \Big\downarrow {\scriptstyle 1\otimes a} & & & & \Big\downarrow {\scriptstyle a \otimes 1} \\ X_1 \otimes ((X_2 \otimes X_3) \otimes X_4) & \xrightarrow{a} & & & (X_1 \otimes (X_2 \otimes X_3)) \otimes X_4 \end{array}$$

(iii)

$$\begin{array}{ccccc} X_1 \otimes (X_2 \otimes X_3) & \xrightarrow{a} & (X_1 \otimes X_2) \otimes X_3 & \xrightarrow{c\otimes 1} & X_3 \otimes (X_1 \otimes X_2) \\ \Big\downarrow {\scriptstyle 1\otimes c} & & & & \Big\downarrow {\scriptstyle a \otimes 1} \\ X_1 \otimes (X_3 \otimes X_2) & \xrightarrow{a} & (X_1 \otimes X_3) \otimes X_2 & \xrightarrow{c\otimes 1} & (X_3 \otimes X_1) \otimes X_2 \end{array}$$

(d) (identity object) $\exists U \in \mathrm{ob}(C)$ together with an isomorphism $\underline{u}: U \xrightarrow{\sim} U \otimes U$ such that $X \mapsto U \otimes X$ is an autoequivalence of categories of C.

Such a category is called a tensor category ACU in **[Sa]** and a symmetric monoidal category with identity in **[Mac]**. A basic example is provided by the category Vec_k of finite-dimensional vector spaces V over a field k together with the familiar tensor product and the canonical associativity and commutativity isomorphisms. An identity object is given by a one-dimensional vector space U together with a basis vector e so that the requisite isomorphism $u: U \to U \otimes U$ is the unique one sending e to $e \otimes e$. A more interesting example is the category $\mathrm{Rep}_k(G)$ of finite-dimensional k-representations $\sigma: G \to \mathrm{GL}(V)$ of a pro-algebraic k-group G with the usual $\otimes$, a, and c. The trivial representation defines an identity object.

In many instances, such as in the automorphic setup (see §2), one has a $\mathbb{C}$-linear abelian, semisimple category, and one would like to define a tensor structure $\otimes$ by suitably defining it on the simple objects. It may be useful to have a concrete description of the inherent difficulties one encounters.

LEMMA 1.1.2. *Let C be a $\mathbb{C}$-linear abelian, semisimple category relative to $\oplus$ and let $\sum = \{S_j \mid j \in J\}$ be a set of representations in* $\mathrm{ob}(C)$ *of the isomorphism classes of the simple objects. Then the following are equivalent*:

(1) $\exists$ *a $\mathbb{C}$-bilinear functor $\otimes: C \times C \to C$ together with functorial isomorphisms a, c of* (1.1.1)(b);

(2) $\exists$ *a family of finite-dimensional $\mathbb{C}$-vector spaces $\{V_{i,j}^k \mid i, j, k \in J\}$ with $V_{i,j}^k = 0$ for almost all k given (i, j) such that $\exists$ $\mathbb{C}$-linear isomorphisms* $(\forall i, j, k, m$ *in* $J)$:

$$a' = a'_{i,j,k,m}: \bigoplus_{r \in J} (V_{i,j}^r \otimes V_{r,k}^m) \xrightarrow{\sim} \bigoplus_{r \in J} (V_{i,r}^m \otimes V_{j,k}^r)$$

and

$$c' = c'_{i,j,k}\colon V^k_{i,j} \xrightarrow{\sim} V^k_{j,i}.$$

PROOF. Suppose (1) holds. Set $V^k_{i,j} = \operatorname{Hom}(S_k, S_i \otimes S_j)$ and $[k; i, j] = \dim V^k_{i,j}$. The associativity isomorphism a_{S_i,S_j,S_k} gives $(\forall m)$ an isomorphism: $\operatorname{Hom}(S_m, S_i \otimes (S_j \otimes S_k)) \xrightarrow{\sim} \operatorname{Hom}(S_m, (S_i \otimes S_j) \otimes S_k)$ which yields $a'_{i,j,k,m}$. Similarly c_{S_i,S_j} gives $c'_{i,j,k}$. Hence (2) follows. Conversely suppose (2) holds. We shall construct $(\otimes, a, c)$ up to some choices, which will not affect the natural equivalence class. For every simple object Y fix an isomorphism $\varphi_Y\colon Y \xrightarrow{\sim} S_i$ such that $\varphi_Y = \mathrm{id}$ if $Y = S_i$. If Y, Y' are simple objects isomorphic to S_i, S_j, respectively, put $Y \otimes Y' = S_i \otimes S_j = \bigoplus_k X^k_{i,j}$, where $X^k_{i,j}$ is the object isomorphic to $S_k^{[k;i,j]}$ such that $\operatorname{Hom}(S_m, X^k_{i,j}) = \delta_{m,k}$, $V^k_{i,j}$, $\forall m$. (Here S^r_k denotes the sum of r copies of S_k. A choice of an isomorphism $\beta\colon \mathbb{C}^{[k;i,j]} \xrightarrow{\sim} V^k_{i,j}$ gives $V^k_{i,j} \xrightarrow{\sim} \operatorname{Hom}(S_k, S_k^{[k;i,j]})$ by $v \mapsto \beta^{-1}(v)$.) Let T, T', Y, Y' be simple objects and let $f \in \operatorname{Hom}(T, Y)$ and $g \in \operatorname{Hom}(T', Y')$. Define $f \otimes g$ as follows. If f or $g = 0$, put $f \otimes g = 0$. If not, f and g must be isomorphisms and $\exists i, j$ such that $T \simeq Y \simeq S_i$ and $T' \simeq Y' \simeq S_j$. To define $f \otimes g \in \operatorname{End}(S_i \otimes S_j)$ it suffices to give elements u_k in $\mathrm{GL}(V^k_{i,j})$. Let $\psi = \varphi_Y \circ f \circ \varphi_T^{-1} \in \operatorname{Hom}(S_i, S_i)$ and $\xi = \varphi_{Y'} \circ g \circ \varphi_{T'}^{-1} \in \operatorname{Hom}(S_j, S_j)$. Then ψ and ξ are isomorphisms and by Schur's lemma are multiplication by scalars b_ψ, b_ξ, in $\mathbb{C}^*$. Put $u_k = b_\psi b_\xi$ id $(\forall k)$. Extend $\otimes$ to all objects of C by semisimplicity. The definition of $f \otimes g$ can accordingly be extended to arbitrary morphisms f, g. One thus gets a $\mathbb{C}$-bilinear functor $\otimes\colon C \times C \to C$. The (functorial) associativity and commutativity isomorphisms follow from a' and c', respectively. If the choices are replaced by new ones, it is clear how to get a natural equivalence between $\otimes$ and the new functor. □

The commutative diagrams of (1.1.1)(c) translate to (corresponding) ones involving $\{V^k_{i,j}\}$. For instance, the hexagonal constraint (ii) becomes the equality $\beta_1 = \beta_2$, where β_1 is the composite map

(1.1.3)
$$\bigoplus_r (V^m_{i,r} \otimes V^r_{j,k}) \overset{(1)}{\to} \bigoplus_r (V^r_{i,j} \otimes V^m_{r,k}) \overset{(2)}{\to} \bigoplus_r (V^m_{k,r} \otimes V^r_{i,j}) \overset{(3)}{\to} \bigoplus_r (V^r_{k,i} \otimes V^m_{r,j})$$

and β_2 is the composite map

$$\bigoplus_r (V^m_{i,r} \otimes V^r_{j,k}) \overset{(4)}{\to} \bigoplus_r (V^m_{i,r} \otimes V^r_{k,j}) \overset{(5)}{\to} \bigoplus_r (V^r_{i,k} \otimes V^m_{r,j}) \overset{(6)}{\to} \bigoplus_r (V^r_{k,i} \otimes V^m_{r,j}).$$

Here (1), (3), (5) are given by a'^{-1}, (2) by $1 \otimes c'$ followed by the flip $V^r_{i,j} \otimes V^m_{k,r} \to V^m_{k,r} \otimes V^r_{i,j}$, (4) by $1 \otimes c'$ and (6) by $c' \otimes 1$.

The tensor structure $\{V^k_{i,j}, a', c'\}$ has invariants modulo the action of $\Pi\,\mathrm{GL}(V^k_{i,j})$. Such invariants were studied by Gallagher for $C = \mathrm{Rep}_{\mathbb{C}}(G)$, G a finite group. See §1.4 below.

Note that the knowledge of just the dimensions $[k; i, j]$ is not sufficient to turn C into a tensor category. (One needs additional information.) In the automorphic setting (see §3.2.6), these integers have a conjectural meaning as the orders of poles at the edge of absolute convergence of certain L-functions.

1.2. Suppose there is an exact, faithful $\mathbb{C}$-linear functor $\eta: C \to \mathrm{Vec}_{\mathbb{C}}$ such that $Y_j := \eta(S_j)$ is one dimensional $\forall j \in J$ and $[k; i, j] = [k; j, i]$. (See §2.4 for automorphic examples.) If we set: $V^k_{i,j} = Y^\vee_k \otimes Y_i \otimes Y_j \otimes \mathbb{C}^{[k;i,j]}$, there is a natural definition of c' induced by $Y_i \otimes Y_j \xrightarrow{\sim} Y_j \otimes Y_i$, but the data is (evidently) insufficient to define a' and satisfy the constraints. There is a very simple example, however, where things work. This is when we have for all (i, j) a unique element, denoted ij, of J such that $(ij)k = i(jk)$ and $[k; i, j] = 1$ (resp. 0) if $k = ij$ (resp. $\neq ij$). One can define $a': V^{ij}_{i,j} \otimes V^{ijk}_{ij,k} \xrightarrow{\sim} V^{ijk}_{i,jk} \otimes V^{jk}_{j,k}$ by means of the natural associativity and commutativity isomorphisms in $\mathrm{Vec}_{\mathbb{C}}$, together with the evaluation isomorphism ev: $Y^\vee_\ell \otimes Y_\ell \xrightarrow{\sim} \mathbb{C}$ $(\forall \ell)$, such that the constraints are satisfied.

1.3. Suppose $(C, \otimes, a, c)$ is a tensor category with an identity object 1. One says (cf. **[De2,** §2.1]) that C is *rigid* iff $\exists$ an involution $X \mapsto X^\vee$ of $\mathrm{ob}(C)$, and maps $(\forall X)$: $\mathrm{ev} = \mathrm{ev}_X \in \mathrm{Hom}(X \otimes X^\vee, 1)$ and $\delta = \delta_X \in \mathrm{Hom}(1, X^\vee \otimes X)$, such that the composites $X \to X \otimes 1 \xrightarrow{\mathrm{id}\otimes\delta} X \otimes (X^\vee \otimes X) \xrightarrow{a} (X \otimes X^\vee) \otimes X \xrightarrow{\mathrm{ev}\otimes\mathrm{id}} 1 \otimes X \to X$ and $X^\vee \to 1 \otimes X^\vee \xrightarrow{\delta\otimes\mathrm{id}} (X^\vee \otimes X) \otimes X^\vee \xrightarrow{a} X^\vee \otimes (X \otimes X^\vee) \xrightarrow{\mathrm{id}\otimes\mathrm{ev}} X^\vee \otimes 1 \to X^\vee$ are both identity maps. In the definition of **[De2]**, a tensor category (over $\mathbb{C}$) is a $\mathbb{C}$-linear rigid abelian tensor category with $\mathrm{End}(1) \simeq \mathbb{C}$. A criterion for rigidity is supplied by the following

PROPOSITION 1.3.1 **[De-M]**. *Let C be a $\mathbb{C}$-linear abelian category with a tensor structure $(\otimes, a, c)$ satisfying* (1.1.1) (a), (b), *and an identity object* 1 *with* $\mathrm{End}(1) \simeq \mathbb{C}$. *Suppose $\exists$ a faithful, exact $\mathbb{C}$-linear functor $\eta: C \to \mathrm{Vec}_{\mathbb{C}}$ such that* (i) $\eta \circ \otimes = \otimes \circ (\eta \times \eta)$; (ii) $\eta(a)$ *and* $\eta(c)$ *are the usual associativity and commutativity isomorphisms in* $\mathrm{Vec}_{\mathbb{C}}$; *and* (iii) $\forall L \in \mathrm{ob}(C)$ *with* $\dim(\eta(L)) = 1$, $\exists M \in \mathrm{ob}(C)$ *such that* $M \otimes L \simeq 1$. *Then C is rigid.*

A $\mathbb{C}$-linear rigid abelian tensor category C with $\mathrm{End}(1) \simeq \mathbb{C}$ is a *neutral Tannakian category* if it admits a fiber functor $/\mathbb{C}$, i.e., a faithful, exact $\mathbb{C}$-linear functor $\eta: C \to \mathrm{Vec}_{\mathbb{C}}$ that is a tensor functor (see **[De-M,** p. 113], for a precise definition). Roughly speaking, being a tensor functor means that η is compatible with the respective structures $(\otimes, a, c)$ and the constraints (1.1.1) (c). One of the main theorems of **[De-M,** Theorem 2.11] asserts that every $\mathbb{C}$-linear neutral Tannakian category C is equivalent (in possibly many ways) to $\mathrm{Rep}_{\mathbb{C}} G$ for an affine group scheme G over $\mathbb{C}$. Moreover, if C is semisimple, the connected component G^0 of G is pro-reductive. Finally we note that a theorem of Deligne **[De2]** asserts that a $\mathbb{C}$-linear rigid abelian tensor category C with $\mathrm{End}(1) \simeq \mathbb{C}$ (which is simply called a tensor category $/\mathbb{C}$ in **[De2]**) is Tannakian iff for every $X \in \mathrm{ob}(C)$ $\exists$ an integer $n \geq 0$ such that $\Lambda^n(X) = 0$. Here $\Lambda^n(X)$ denotes the image of the anti-symmetrization map $\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\sigma\colon \otimes^n X \to \otimes^n X$.

We conclude this subsection by mentioning two relevant examples. Let F be a field $\subseteq \overline{F} \subseteq \mathbb{C}$ and let $M_{F,\mathbb{C}}$ (resp. $M_{F,\mathbb{C}}^{\mathrm{num}}$) denote the (true) category of (pure) motives over F as in **[De-M,** §6] (resp. **[Ja]**), with the homs given by absolutely Hodge cycles (resp. algebraic cycles modulo numerical equivalence), with the additional proviso that the coefficients be extended to $\mathbb{C}$. Then $M_{F,\mathbb{C}}$ is a $\mathbb{C}$-linear semisimple neutral Tannakian category relative to the fiber functor $X \mapsto H_B(X) \otimes \mathbb{C}$. $M_{F,\mathbb{C}}^{\mathrm{num}}$ is a $\mathbb{C}$-linear abelian semisimple category, but it is Tannakian only under the hypothesis that the Künneth components of the diagonal (relative to some Weil cohomology) are algebraic. Let $\mathfrak{G}_F$ denote the motivic Galois group associated to $M_{F,\mathbb{C}}$ by the Tannakian formalism. Then there is a short exact sequence

$$1 \to \mathfrak{G}_F^0 \to \mathfrak{G}_F \to \mathrm{Gal}(\overline{F}/F) \to 1 \tag{1.3.2}$$

where $\mathfrak{G}_F^0$ is a connected pro-reductive group over $\mathbb{C}$. (We refer to **[De-M]** and **[Ja]** for details on this paragraph.) It should be stressed that for a comparison with the (conjectural) Langlands group, one needs to extend coefficients to an algebraically closed field since automorphic forms are not sensitive to coefficients.

1.4. Recovery of finite groups. Let G be a finite group of order n and let $\chi_1, \dots, \chi_m$ (resp. $C_1, \dots, C_m$) denote its distinct irreducible (complex) characters (resp. conjugacy classes) with χ_1 (resp. C_1) being the principal character (resp. class) denoted sometimes by 1. Fix a set $\sum = \{\sigma_j \mid 1 \leq j \leq m\}$ of representatives for the irreducible representations of G over $\mathbb{C}$, with χ_i being the character of σ_i. For any pair of representations (τ_1, τ_2) of G, let $[\tau_1 : \tau_2]$ denote the dimension of $\mathrm{Hom}_G(\tau_2, \tau_1)$. Say that G is determined if it is so up to isomorphism. Denote by h_i the order of C_i, $\forall i$. For any pair of characters (ψ, ψ') of G, set $\langle \psi, \psi' \rangle = \frac{1}{n}\sum_{g \in G} \psi(g)\overline{\psi}'(g) = \frac{1}{n}\sum_{l=1}^m h_l \psi(C_l)\overline{\psi}'(C_l)$, where $\overline{\psi}'$ is the complex conjugate character of ψ'. Since $\langle \chi_i, \chi_j \rangle = \delta_{i,j}$, the integers $[\tau_i \otimes \tau_j ; \tau_k]$ are determined by the character table $(\chi(C_j))_{1 \leq i, j \leq m}$ as the sum $\frac{1}{n}\sum_{l=1}^m h_l \chi_i(C_l)\chi_j(C_l)\overline{\chi}_k(C_l)$. (The h_l and n are first determined by $1 = (h_l/n)(\sum_{i=1}^m \chi(C_l)\overline{\chi}_i(C_l))$ and $n = \sum_{l=1}^m h_l$.) It is well known that the character table does not determine G. The easiest counterexample is provided by the nonisomorphic pair (D_8, Q_8) having the same character table, where D_8 (resp. Q_8) denotes the dihedral (resp. quaternion) group of order 8. Note however that if σ (resp. σ') is the unique irreducible two-dimensional representation of D_8 (resp. Q_8), then $\det(\sigma)$ is a nontrivial quadratic character, while $\det(\sigma') = 1$. Thus these two groups are distinguished by the data $\{[\Lambda^2(\sigma_i)\colon \sigma_j]\}$. Such a phenomenon does not persist in general. One has

PROPOSITION 1.4.1 [**Da**]. *A finite group G is not determined by the knowledge of the integers $\{[\Lambda^r(\sigma_i) : \sigma_j]\}$ in addition to the character table.*

To be precise, Dade proved (in [**Da**]) the existence of a class of p-groups ($p \geq 5$) that provided a negative answer to the following question of Brauer: Is a finite group determined by the character table together with the knowledge of which conjugacy classes are rth powers of the C_i for all $r \geq 1$? The Proposition can be extracted as follows: Let $\chi_i^{(r)}$ denote the character of $\Lambda^r(\sigma_i)$. Then one has the following recursive formula (by combining (1.27) and Proposition 12.8 of [**C-R**, vol. I]:

$$(1.4.2) \qquad (-1)^{r-1} r\chi_i^{(r)}(C_k) = \chi_i(C_k^r) - \sum_{l=1}^{r-1}(-1)^{l-1}\chi_i^{(l)}(C_k)\chi_i(C_k^{r-l})$$

which gives $[\Lambda^{(r)}(\sigma_i) : \sigma_j] = \frac{1}{n}\sum_{k=1}^m h_k\chi_i^{(r)}(C_k)\overline{\chi}_j(C_k)$ by induction on r from Brauer's data.

As noted in §1.3, the Tannakian formalism allows one to recover a finite group G. It is shown in [**Ga**] that the new invariants of Tannaka can be formulated as G-invariant subspaces, determined up to equivalence, in the spaces of the tensor products $\sigma_i \otimes \sigma_j \otimes \sigma_k$. In fact, Gallagher goes further and shows that G is determined by a set of *numerical* invariants. More precisely one has

THEOREM 1.4.3 [**Ga**]. *A finite group G is determined by a list $\{\chi_1, \chi_2, \dots\}$ of its characters together with the values of the following expressions* $(\forall n, r)$:

$$M(W) := \int_{G^n} \prod_{i=1}^r \chi_i(W_i(g_1, \dots, g_n))dg_1 \cdots dg_n$$

where $W = (W_1, \dots, W_n)$ runs over all maps: $G^n \to G^r$.

It is useful to note that this result extends to any compact group G.

2. Automorphic forms on GL_n

2.1. The building blocks. Let F be a local field. When F is non-Archimedean, denote by $\mathfrak{O}$ the ring of integers with prime ideal $\mathfrak{P}$, discrete valuation v, residue field $k = \mathbb{F}_q$, $q = p^f$, and uniformizer $\varpi = \varpi_v$. Let $|\cdot|$ denote the absolute value on F normalized such that $|\varpi| = q^{-1}$ (resp. $|x| = \mathrm{sgn}(x)x$, resp. $|z| = z\overline{z}$) when F is $\mathfrak{P}$-adic (resp. $\mathbb{R}$, resp. $\mathbb{C}$). For any $n \geq 1$, denote by $G_n = G_{n,v}$ the group of F-rational points of GL_n equipped with the evident locally compact topology. Let $K_n = K_{n,v}$ denote its standard maximal compact subgroup, namely $\mathrm{GL}_n(\mathfrak{O})$ (resp. $O(n)$, resp. $U(n)$) when F is non-Archimedean (resp. $\mathbb{R}$, resp. $\mathbb{C}$).

In the $\mathfrak{P}$-adic case, an *algebraic* (or *smooth*) representation of G_n is a homomorphism $\pi\colon G_n \to \mathrm{GL}(V)$, written as (π, V) or even π for short, where V is a $\mathbb{C}$-vector space, such that the stabilizer of each v in V is open in G_n. It is admissible iff its restriction to K_n is isomorphic to an algebraic direct sum of irreducibles, necessarily finite dimensional, of K_n with finite multiplicities. It is known that every irreducible algebraic representation is automatically admissible. In the Archimedean case, we shall by abuse define an admissible representation of G_n to be an admissible (Lie G_n, K_n)-module [W].

The dual (or contragredient) of an admissible (π, V) is denoted by $(\check{\pi}, \check{V})$, and the matrix coefficients are defined by the function $m\colon V \times \check{V} \to \mathrm{Maps}(G_n, \mathbb{C})$, $(v, \check{v}) \to c_{v,\check{v}}\colon (g \to \langle \pi(g)v, \check{v}\rangle = \langle v, \check{\pi}(g)\check{v}\rangle)$, where $\langle\, ,\, \rangle$ denotes the canonical contraction map of $V \times \check{V}$ into $\mathbb{C}$. A cuspidal representation of G_n is an admissible representation (π, V) such that the matrix coefficients are square integrable modulo the center Z_n. Denote by $A_0(n, F)$ the set of all irreducible cuspidal representations of G_n. Note that for $n = 2$, this definition of cuspidals includes, besides the supercuspidals, the special representation.

DEFINITION 2.1.1. $\underline{A_0}(F)$ = the category whose objects are elements of $A_0(F) := \bigcup_{n\geq 1} A_0(n, F)$ with morphisms given by

$$\mathrm{Hom}(\pi, \pi') = \delta_{n,n'} \mathrm{Hom}_{G_n}(\pi, \pi')$$

for π, π' in $A_0(n, F)$, $A_0(n', F)$ respectively.

Here $\delta_{n,n'}$ signifies as usual the Kronecker delta, which is 1 if $n = n'$ and zero otherwise. $\mathrm{Hom}(\pi, \pi')$ is accordingly a $\mathbb{C}$-vector space of dimension one or zero.

For $n \geq 1$, let $\det\colon G_n \to F^*$ denote the determinant map. Then every $\pi \in A_0(n, F)$ has central character $\omega \circ \det$, which we shall write by abuse of notation as ω. The structure of F^* then allows one to write it as $\omega_0|\cdot|^{t(\pi)}$ for a real number $t(\pi)$ and a unitary character σ_0. Introduce a total order $\geq$ in $A_0(F)$ by setting $\pi \geq \pi'$ iff $t(\pi) \geq t(\pi')$.

Let $n = \sum_{1\leq i\leq m} n_i$ be a partition into a sum of nonnegative integers, and let π_i be in $A_0(n_i, F)$ for each i with $t_i = t(\pi_i)$. For $n \geq 1$, let P be the standard parabolic subgroup of G_n with Levi component $M = \prod G_{n_i}$, with the product taken only over those $i \leq m$ with nonzero n_i. Denote by $I(\pi_1, \dots, \pi_m)$ the admissible representation of G_n induced by the representation of P which is $\pi_1 \otimes \cdots \otimes \pi_m$ on M and trivial on the unipotent radical. The *isobaric Langlands subquotient*, denoted by $\pi_1 \boxplus \cdots \boxplus \pi_m$, is defined as follows: It is the unique quotient of $I(\pi_1, \dots, \pi_m)$ if $\pi_1 \geq \cdots \geq \pi_m$; otherwise it is the unique subquotient of $I(\pi_1, \dots, \pi_m)$ that is isomorphic to $\pi_1' \boxplus \cdots \boxplus \pi_m'$, where $(\pi_1', \dots, \pi_m')$ is an m-tuple in $A_0(F)$ obtained by permutation of $\{\pi_1, \dots, \pi_m\}$ such that $\pi_1' \geq \cdots \geq \pi_m'$. (For a quick background on the representation theory of G_n, the reader may consult the articles of S. Kudla and A. W. Knapp in this volume.) Let $\mathrm{Isob}(F)$ denote the set of representations of the form $\pi_1 \boxplus \cdots \boxplus \pi_m$ with π_i cuspidal. Note that the representation $|\cdot|^{(n-1)/2} \boxplus |\cdot|^{(n-3)/2} \boxplus \cdots \boxplus |\cdot|^{(3-n)/2} \boxplus |\cdot|^{(1-n)/2}$ is the trivial representation of $\mathrm{GL}_n(F)$. In particular, the irreducible subquotients of the representation of $\mathrm{GL}_2(F)$ induced by the character $\left(\begin{smallmatrix} a & x \\ 0 & b \end{smallmatrix}\right) \mapsto |a/b|^{1/2}$ (of the Borel subgroup) are $|\cdot|^{1/2} \boxplus |\cdot|^{-1/2}$ and the special representation with trivial central character (which is cuspidal).

Now let F be a global field with an adèle ring $\mathbb{A}_F = F_\infty \times \mathbb{A}_F^f$. Denote by $G_n = G_{n,\infty} \times G_n^f$ the group of $\mathbb{A}_F$-valued points of GL_n, equipped with the natural locally compact topology. By definition, an admissible G_n-module [F] is a representation (π, V), where V is a $\mathbb{C}$-vector space that is simultaneously an admissible (Lie G_∞, K_∞)-module and an admissible G_n^f-module such that the action of G_n^f commutes with that of Lie G_∞ and K_∞. Here K_∞ denotes a maximal compact subgroup of G_∞, and an admissible representation of the totally disconnected group G_n^f is defined exactly as it is done for the local groups $G_{n,v}$ at non-Archimedean places v. Denote by Z_n the center of GL_n.

Let ω be a character of $\mathbb{A}_F^*/F^*$. Choosing a quotient measure on $G(\mathbb{Q})Z_n(\mathbb{A}_F)\backslash G_n$, we denote by $L^2(n, \omega)$ the space of measurable functions $\varphi\colon G_n \to \mathbb{C}$ which are square integrable modulo $G_n(\mathbb{Q})Z_n(\mathbb{A}_F)$ (on the left) and transform by the rule: $\varphi(\gamma z g) = \omega(z)\varphi(g)$, for all $\gamma \in G_n(\mathbb{Q})$, $z \in Z_n(\mathbb{A}_F)$, and $g \in G_n$. G_n acts on this Hilbert space by right translation ρ. Denote by $L_0^2(n, \omega)$ the closed subspace of $L^2(n, \omega)$ consisting of functions φ that satisfy (for almost all g in G_n):

$$\int_{U(F)\backslash U(\mathbb{A}_F)} \varphi(ug)\, dg = 0 \tag{2.1.2}$$

for all the unipotent radicals U of standard parabolic subgroups $P = M \times U$ of GL_n. This space, called the *space of cusp forms* on G_n with central character ω, is G_n-invariant and is well known to be contained in the discrete part of $L^2(n, \omega)$. This way one gets a (Hilbert) direct sum decomposition as G_n-modules:

$$L_0^2(n, \omega) = \widehat{\bigoplus_\pi} m(\pi) H_\pi, \tag{2.1.3}$$

where (π, H_π) runs over a set of inequivalent unitary irreducible representations of G_n with $m(\pi)$ denoting the multiplicity of each π. A theorem of Shalika [**Sha1**] says that $m(\pi) = 1$ for every π. (This is a very special feature of GL_n and for general groups this is false. There is a conjectured formula of J. Arthur [**A**].) One knows that the space V_π of "admissible" vectors in H_π is a dense subspace invariant under G_n.

DEFINITION 2.1.4. An irreducible admissible representation (π, V_π) of G_n is *cuspidal* or cuspidal automorphic iff $\pi \otimes |\det(\cdot)|^s$ is, for some $s \in \mathbb{C}$, isomorphic to the admissible subspace of an irreducible summand (π, H_π) of $L_0^2(n, \omega)$, for some ω.

Now let $A_0(n, F)$ be the set of all *cuspidal* representations on G_n, and define the category $\underline{A}_0(F) = \bigcup_{n\geq 0} A_0(n, F)$ exactly as in Definition 2.1.1

(local case). *We have a global version of* $t(\pi) \in \mathbb{R}$, *which allows us to* introduce an ordering $\geq$ on $\underline{A}_0(F)$ as before.

Let

$$N_n = \left\{ u(\underline{x}) = \begin{pmatrix} 1 & x_1 & & * \\ 0 & & \ddots & x_{n-1} \\ & & & 1 \end{pmatrix} \right\}$$

denote the standard maximal unipotent subgroup of GL_n, and let ψ be a nontrivial unitary character of (the additive group of) $\mathbb{A}_F$, trivial on F. Define a character $\Theta: N_n(\mathbb{A}_F) \to U(1) \subseteq \mathbb{C}^*$ by sending $u(\underline{x})$ to $\psi(\sum_{i=1}^{n-1} x_i)$. We say that a representation (π, V_π) of G_n is generic if it admits a linear form λ in its admissible dual (see **[Sha1]**) such that $\pi^\vee(ng)\lambda = \Theta(n)\pi^\vee(g)\lambda$, for all $n \in N_n(\mathbb{A}_F)$ and $g \in G_n$. The notion of being generic is independent of the choice of Θ (for $\mathrm{GL}(n)$). A well-known result **[Sha1]** is that every *cuspidal* automorphic representation (π, V_π) of G_n is generic. This leads to the existence of a *new vector* u_0, uniquely defined up to a scalar multiple, in (any cuspidal) V_π such that u_0 is fixed by the *conductor* **[J-PS-S]** K_π, which is a suitable compact open subgroup of G_n^f.

DEFINITION 2.1.5. An irreducible admissible representation (π, V_π) of G_n is *automorphic* iff there is a standard parabolic subgroup $P = M \times U$ of $\mathrm{GL}(n)$ with Levi component $M = \prod_{i=1}^r \mathrm{GL}(n_i)$ and cuspidal representations $\sigma_1, \dots, \sigma_r$ of $G_{n_1}, \dots, G_{n_r}$ respectively, such that π is an irreducible subquotient of the induced representation $\mathrm{Ind}_{P(\mathbb{A}_F)}^{G_n}((\sigma_1 \times \cdots \times \sigma_r) \otimes 1)$.

That this agrees with the usual definition of an automorphic representation as being a subquotient of the space of automorphic forms on G_n is a consequence of a theorem of Langlands **[La5]**.

Now let (π, V_π) be an automorphic representation of G_n occurring in some $\mathrm{Ind}_{P(\mathbb{A}_F)}^{G_n}((\sigma_1 \times \cdots \times \sigma_r) \otimes 1)$. We shall now give a criterion for π to be *isobaric.* Since this property depends only on the isomorphism class of π, we shall now make use of the factorization $\pi \simeq \otimes_v \pi_v$ and note that each π_v is in the Jordan-Hölder series of $\mathrm{Ind}_{P(F_v)}^{G_{n,v}}((\sigma_{1,v} \times \cdots \times \sigma_{r,v}) \otimes 1)$. The $\sigma_{i,v}$ need no longer be cuspidal, but they are generic. This implies (by the local classification results of Bernstein-Zelevinsky and Vogan) that each $\sigma_{i,v}$ is a *full* induced representation $\mathrm{Ind}_{P_i(F_v)}^{G_{n_i,v}}((\tau_{1,v} \times \cdots \times \tau_{s(i),v}) \otimes 1)$, where $P_i = M_i \times U_i$ is a standard parabolic subgroup of $\mathrm{GL}(n_i)$ with $M_i = \mathrm{GL}_{m_1} \times \cdots \times \mathrm{GL}_{m_{s(i)}}$, and each $\tau_{j,v}$ is in $R_0(m_j, F_v)$. By the compatibility of induction by stages, this means that π_v is a subquotient of the representation of $G_{n,v}$ induced from the parabolic Q with Levi component $L = \prod_{i=1}^r M_i$, with the representation on L being given by $\prod_{i=1}^r \prod_{j=1}^{s(i)} \tau_{j,v}$.

DEFINITION 2.1.6. (π, V_π) as above is *isobaric* iff π_v is the isobaric Langlands subquotient $\boxplus_{i=1}^r \boxplus_{j=1}^{s(i)} \tau_{j,v}$ at *every* place v.

Let $\mathrm{Aut}(F) = \bigcup_n \mathrm{Aut}(n, F)$ and $\mathrm{Isob}(F) = \bigcup_n \mathrm{Isob}(n, F)$ respectively denote the sets of automorphic and isobaric representations of G_n for all n.

The following is a basic result on isobaric representations with the global part of part (a) (resp. part (b)) being due to Langlands (resp. Jacquet and Shalika).

THEOREM 2.1.7. *Let F be a local or a global field. Then*

(a) **[La5]** *Every $\pi \in \mathrm{Adm}(F)$, F: local, is isomorphic to an element of* $\mathrm{Isob}(F)$, *while in the global case, given any π in $\mathrm{Aut}(F)$ there exists an isobaric (automorphic) representation π' such that $\pi_v \simeq \pi'_v$ at almost all v.*

(b) **[J-S1]** *Suppose two isobaric representations $\pi_1 \boxplus \cdots \boxplus \pi_m$ and $\pi'_1 \boxplus \cdots \boxplus \pi'_n$ are isomorphic. Then $m = n$ and $\pi_i \simeq \pi'_{\sigma(i)}$ $\forall i \le m$, where σ is a permutation of $\{1, \ldots, m\}$.*

2.2. Certain categories. For F a global or a local field write: $A_0(F) = \{\pi_i \mid i \in I\}$ relative to an indexing set $I = I(F)$.

DEFINITION 2.2.1. Let $\underline{A}(F)$ be the category $\coprod_{i\in I} \mathrm{Vec}_{\mathbb{C}}$ with the hom sets being given by: $\mathrm{Hom}((W_i), (W'_i)) = (M_{i,j})_{i,j\in I}$, with $M_{i,j} = \mathrm{Hom}(W_i, W'_j) \otimes, \mathrm{Hom}(\pi_i, \pi_j)$.

Put $\deg(\pi_i) = n$ if $\pi_i \in A_0(n, F)$. Recall that $\mathrm{Hom}(\pi_i, \pi_j) = 0$ unless $\deg(\pi_i) = \deg(\pi_j)$ and $\pi_i \simeq \pi_j$, in which case it is of dimension 1 over $\mathbb{C}$. It is easy to see that $\underline{A}(F)$ can be endowed with the structure of a $\mathbb{C}$-linear abelian category relative to an obvious sum operation $\oplus$. It is also semisimple with the simple objects being of the form $T_j := (W_i)$ with $W_i = 0$ if $i \neq j$ and $\dim W_j = 1$. Moreover there is an involution $\vee$ on $\mathrm{ob}(\underline{A}(F))$ induced by the contragredient operation $\pi \mapsto \pi^\vee$ on $A_0(F)$.

Every finite set B of cuspidal representations defines an object $(W_i(B))$ of $\underline{A}(F)$ defined by $W_i(B) = \mathbb{C}^{[\pi_i : B]}$, where $[\pi_i : B]$ denotes the multiplicity of π_i in B (not up to isomorphism). Of course B also defines, once an ordering is chosen, an isobaric representation $\boxplus \pi_i^{[\pi_i, B]}$. As will be evident, the definition of $\underline{A}(F)$ is simply a formal device to circumvent the difficulties in turning $\mathrm{Isob}(F)$ into an abelian category. Compare with **[C1,** §1.4.2], where the inherent ambiguities (in the definition of morphisms and $\boxplus$) are bypassed by introducing a category $\mathrm{Isob}_1(F)$, whose simple objects are restricted to be certain "principal" models, one for each isomorphism class in $A_0(F)$.

When F is a *number field*, one says that a cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_F)$ is of *Galois* (or *Artin*) *type* iff for every *Archimedean* place w of F the representation $\sigma_w(\pi)$ of the local Weil group W_{F_w} associated to π_w in **[La4]** is trivial on $\overline{F}_w^*(\simeq \mathbb{C}^*)$, which is a subgroup of W_{F_w} of index 2 (resp. 1) when w is real (resp. complex). Set

$$A_0^{\mathrm{Gal}}(F) = \{\pi \in A_0(F) \mid \pi \text{ is of Galois type}\}. \tag{2.2.2}$$

Set $C_F = F^*$ (resp. $\mathbb{A}_F^*/F^*$) if F is a local (resp. global) field, and put

$$A_0^{\mathrm{ur}}(F) = \{\mu \in \mathrm{Hom}_{\mathrm{cont}}(C_F, \mathbb{C}^*) \mid \mu \text{ is unramified}\}. \tag{2.2.3}$$

In the global case we say that μ is unramified if it is so at every place. We will adopt the same terminology for any π in $A_0(F)$.

Finally, for every finite set S of places of a global field F, we set

$$A_0^S(F) = \bigcup_{n\geq 1} A_0^S(n, F), \tag{2.2.4}$$

where

$$A_0^S(n, F) = \left\{ \pi \in \mathrm{Adm}(\mathrm{GL}_n(\mathbb{A}_F^S)) \,\middle|\, \begin{array}{c} \exists\, \pi_S \in \mathrm{Adm}(\mathrm{GL}_n(F_S)) \\ \text{such that} \\ \pi_S \otimes \pi \in A_0(n, F) \end{array} \right\}.$$

Here Adm(−) denotes the set of irreducible admissible representations of −. Denote by $\underline{A}^{\mathrm{Gal}}(F)$ (resp. $\underline{A}^{\mathrm{ur}}(F)$, resp. $\underline{A}^S(F)$) the $\mathbb{C}$-linear abelian category generated by $A_0^{\mathrm{Gal}}(F)$ (resp. $A_0^{\mathrm{ur}}(F)$, resp. $A_0^S(F)$.)

2.3. Basic questions. As stated in the introduction, one wishes for the existence of a pro-reductive group of $\mathscr{L}_F$ over $\mathbb{C}$ whose irreducible representations of dimension n parametrize the cuspidal representations of G_n. This suggests the existence of a tensor product of objects of $\underline{A}_0(F)$ among other things. We shall now formulate it more precisely in the global case. (See **[Ku]** for a discussion of the local problems.) First we note that, given $\pi_i \in \mathrm{Isob}(n_i, F)$, $i = 1, 2$, there exists an Euler product $L^S(\pi_1, \pi_2; \otimes, s) = \prod_{v\notin S} L_v(\pi_1, \pi_2; \otimes, s)$ for every finite set S of places containing the ramified places of π_1, π_2, convergent in some right half-plane, and admitting a meromorphic continuation to the whole s-plane. (One has information at the ramified places as well, which we are ignoring here.) These functions will be introduced and discussed in §3. The unramified Euler factors have degree n_1, n_2. Denote by $L^S(\pi, s)$ the standard L-function (of degree n) of $\pi \in \mathrm{Isob}(n, F)$ with the Euler factors at $v \notin S$ removed.

CONJECTURE 2.3.1 (Langlands). Let F be a global field, and let $\pi_i \in \mathrm{Isob}(n_i, F)$, $i = 1, 2$. Then does there exist $\pi_1 \boxtimes \pi_2 \in \mathrm{Isob}(n_1 n_2, F)$ such that

$$L^S(\pi_1 \boxtimes \pi_2, s) = L^S(\pi_1, \pi_2; \otimes, s)\,? \tag{$*$}$$

This is one of the major open problems of the field. For a discussion of the approach via the converse theorems of Piatetski-Shapiro, et al, see §3.3.

By the strong multiplicity-one theorem **[J-S]**, the identity $(*)$ characterizes $\pi_1 \boxtimes \pi_2$ uniquely up to isomorphism. We note that for every $v \notin S$ we have a natural candidate for the v-component $(\pi_1 \boxtimes \pi_2)_v$. Indeed, there are unramified characters $\chi_{1,v}, \dots, \chi_{n_1,v}, \chi'_{1,v} \dots, \chi'_{n_2,v}$ of F_v^* such that $\pi_{1,v} \simeq \chi_{1,v} \boxplus \cdots \boxplus \chi_{n_1,v}$ and $\pi_{2,v} \simeq \chi'_{1,v} \boxplus \cdots \boxplus \chi'_{n_2,v}$. One may take $(\pi_1 \boxtimes \pi_2)_v$ to be $\boxplus_{1\leq j\leq n_1, 1\leq k\leq n_2} \chi_{j,v}\chi'_{k,v}$. Put $\pi^S := \bigotimes_{v\notin S}(\pi_1 \boxtimes \pi_2)_v$. Then $\pi_S \otimes \pi$ is, for any $\pi_S \in \mathrm{Adm}(n, F_S)$, an admissible representation of $\mathrm{GL}_n(\mathbb{A}_F)$. The problem is to show that, for a suitable π_S, it is automorphic.

The following is perhaps too optimistic.

QUESTION 2.3.2. Let F be a local or a global field. Can we endow $\underline{A}(F)$ with the structure of a neutral Tannakian category?

Obviously one has no idea at present how to attack such a question, which can be equally well asked about $\underline{A}^{\mathrm{Gal}}(F)$ and $\underline{A}^{S}(F)$. There is a canonical candidate for a set of representatives $\sum$ for the isomorphism classes of cuspidal representations, namely the set of $\pi_j \in A_0(F)$ such that $\pi_j \otimes |\det(\cdot)|^s$ is, for some s, in the space of cusp forms. There is a conjectural formula (see Proposition 3.2.7) for the "multiplicity" $[k; i, j]$ of π_k in the conjectural $\pi_i \boxtimes \pi_j$. The only other information, which may be of some (limited) use in understanding " $\mathrm{Hom}(\pi_k, \pi_i \otimes \pi_j)$ ", is the existence of the canonical one-dimensional subspace of the space of each π_j consisting of new vectors.

It should be mentioned that the general philosophy suggests that the conjectural group $\mathscr{L}_F^{\mathrm{Gal}}$ associated to $\underline{A}^{\mathrm{Gal}}(F)$ should be isomorphic to $\mathrm{Gal}(\overline{F}/F)$ for any number field F, and the functor $\underline{A}^{\mathrm{Gal}}(F) \to \underline{A}(F)$ should be fully faithful and correspond to a surjection $\gamma: \mathscr{L}_F \twoheadrightarrow \mathscr{L}_F^{\mathrm{Gal}}$ with connected kernel. Also, for every finite set S of places one could hope for a surjection $\gamma^S: \mathscr{L}_F^S \twoheadrightarrow \mathfrak{G}_F^S$, where $\mathfrak{G}_F^S$ is the S-motivic Galois group defined by the category (relative to absolutely Hodge cycles) of (pure) motives over F with coefficients extended to $\mathbb{C}$ and unramified outside S (see the end of §1.3). To sum up, we expect the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathscr{L}_F^0 & \longrightarrow & \mathscr{L}_F & \xrightarrow{\gamma} & \mathscr{L}_F^{\mathrm{Gal}} & \longrightarrow & 1 \\ & & \downarrow & & \downarrow r & & \downarrow r^{\mathrm{Gal}} & & \\ 1 & \longrightarrow & \mathfrak{G}_F^0 & \longrightarrow & \mathfrak{G}_F & \longrightarrow & \mathrm{Gal}(\overline{F}/F) & \longrightarrow & 1 \end{array} \tag{2.3.3}$$

To say that r^{Gal} is an isomorphism is essentially the strong Artin conjecture.

REMARK 2.3.4. One says that a cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_F)$, F a number field, is *algebraic* if for every Archimedean place w of F, the representation $\sigma_w(\pi)$ (given by **[La4]**) of W_{F_w} is of type A_0 (in the sense of Weil) when restricted to $\overline{F}_w^* \simeq \mathbb{C}^*$. This aspect is analyzed extensively in **[C1]**. One can define $A^{\mathrm{alg}}(F)$ and ask analogous questions. This setup may be better in the sense that one could hope for $\mathscr{L}_F^{\mathrm{alg}}$ to be isomorphic to $\mathfrak{G}_F$ (even over $\overline{\mathbb{Q}}$).

2.4. A local component functor on $\underline{A}^S(F)$. Let F be a global field and let S be a finite set of places of F containing the archimedean ones.

PROPOSITION 2.4.1. *For every $v \notin S$, there exists a functor*

$$\lambda_v: \underline{A}^S(F) \to \underline{A}^{\mathrm{ur}}(F_v)$$

that is $\mathbb{C}$-linear, exact, and faithful.

PROOF. Let π be an irreducible admissible representation of $\mathrm{GL}_n(\mathbb{A}_F^S)$ such that, for some π_S in $\mathrm{Adm}(n, F_S)$, the representation $\pi_S \otimes \pi$ of $\mathrm{GL}_n(\mathbb{A}_F)$ is cuspidal. Then one knows (cf. **[J-PS-S]**) that there is a canonical one-dimensional subspace Y_π, called the new subspace, of the space $\mathscr{H}_\pi$ of π. If K is the conductor of π, then Y_π is the space of K-invariants in $\mathscr{H}_\pi$. Set $(\forall v \notin S)$

$$l_v(\pi) = \text{the } \mathbb{C}\text{-linear } \mathrm{GL}_n(F_v)\text{-span of } Y_\pi . \tag{2.4.2}$$

This is an admissible representation of $\mathrm{GL}_n(F_v)$.

Now take π to be in $A_0^S(n, F)$ and fix $v \notin S$. Then $l_v(\pi)$ is unramified and generic. It must be isomorphic to a full induced representation $\mathrm{Ind}_{B_v}^{\mathrm{GL}(n, F_v)}((\mu_1 \times \cdots \times \mu_n) \otimes 1)$ where $\mu_i \in A_0^{\mathrm{ur}}(F_v)$ $\forall i$, and B_v denotes the F_v-points of the Borel subgroup of $\mathrm{GL}_n(F_v)$. For every $\chi \in A_0^{\mathrm{ur}}(F_v)$ put $[\chi : l_v(\pi)] =$ multiplicity of χ in $\{\mu_1, \ldots, \mu_n\}$. This is well defined because if $l_v(\pi)$ is also isomorphic to a full representation by the character $((\nu_1 \times \cdots \times \nu_n) \otimes 1)$ of B_v, then $\mu_i = \nu_{\sigma(i)}$ $\forall i$, for some permutation σ. If (W_i) is an object of $\underline{A}^S(F)$ with $W_i = 0$ except for $i = j$, then we set

$$\lambda_v((W_i)) = (V_k)_{k \in I_v} \tag{2.4.3}$$

where $V_k = Y_\pi \otimes \mathbb{C}^{[\chi_k \,:\, l_v(\pi)]} \otimes W_i$. Here I_v denotes the indexing set for $A_0^{\mathrm{ur}}(F) = \{\chi_k\}$. This can be extended to all objects by additivity. We need to define λ_v on morphisms.

Now let $\pi_j, \pi_l \in A_0^S(n, F)$ and let $U_j = (W_i)$, $U_l = (W_i')$ be in $\mathrm{ob}(\underline{A}^S(F))$ such that W_i (resp. W_i') is 0 unless $i = j$ (resp. $i = l$). Let $\psi \in \mathrm{Hom}(W_j, W_l)$ and $\varphi \in \mathrm{Hom}(\pi_j, \pi_l)$. If $\varphi = 0$ we set $\lambda_v(\psi \otimes \varphi) = 0$. If not, $\deg(\pi_j) = \deg(\pi_l)$ and $\pi_j \simeq \pi_l$. Set $\varphi_0 = \varphi|_{Y_{\pi_j}}$. Then φ_0 must map Y_{π_j} isomorphically onto Y_{π_l}. In addition, since $l_v(\pi_j) \simeq l_v(\pi_l)$, the multiplicity m_k of each χ_k in $l_v(\pi_j)$ is the same as in $l_v(\pi_l)$. Thus we have $\lambda_v(U_j) = (Y_{\pi_j} \otimes \mathbb{C}^{m_k} \otimes W_j)_k$ and $\lambda_v(U_l) = (Y_{\pi_l} \otimes \mathbb{C}^{m_k} \otimes W_l')_k$. We define $\lambda_v(\psi \otimes \varphi)$ to be the block diagonal map $(\lambda_v(\psi \otimes \varphi)_k)$ where

$$\lambda_v(\psi \otimes \varphi)_k = \varphi_0 \otimes \mathrm{id} \otimes \psi . \tag{2.4.4}$$

It is easy to see how to extend this to all morphisms by linearity to get a $\mathbb{C}$-linear functor. Since $\varphi_0 = 0$ iff $\varphi = 0$ for every $\varphi \in \mathrm{Hom}(\pi_j, \pi_l)$ with $\deg(\pi_j) = \deg(\pi_l)$, we see that λ_v is faithful. It is also evidently exact. □

There is a natural functor $\eta_v : \underline{A}^{\mathrm{ur}}(F_v) \to \mathrm{Vec}_{\mathbb{C}}$ taking (V_k) to $\bigoplus_{k \in I_v}(V_k \otimes \mathbb{C}_{\chi_k})$, where $\mathbb{C}_{\chi_k}$ denotes a one-dimensional space on which F_v^* acts by χ_k. The definition of η_v on the arrows is clear. This provides an example of the sort discussed at the end of §1.2 leading to the structure of a tensor category with an identity object 1 defined by the trivial representation with $\mathrm{End}(1) \simeq \mathbb{C}$. It is also rigid via the involution $\chi_k \mapsto \chi_k^\vee = \chi_k^{-1}$ and the evident maps ev and δ. Finally, η_v is evidently a fiber functor, thus making $\underline{A}^{\mathrm{ur}}(F_v)$ a neutral Tannakian category. We get a $\mathbb{C}$-linear exact, faithful functor

$$\eta_v \circ \lambda_v : \underline{A}^S(F) \to \mathrm{Vec}_{\mathbb{C}} . \tag{2.4.5}$$

A natural question to ask, assuming that there exists a tensor structure on $\underline{A}^S(F)$, is whether a suitably defined local component functor can be used to define a fiber functor by composing it with η_v.

REMARK 2.4.6. Sending a cuspidal π_j to its new subspace Y_{π_j} and $\varphi \in \mathrm{Hom}(\pi_j, \pi_l)$ to φ_0 as in (the proof of) Proposition 2.4.1, we get a functor $\Theta: \underline{A}^S(F) \to \mathrm{Vec}_{\mathbb{C}}$ that is $\mathbb{C}$-linear, exact, and faithful. This is not what we want, however, since the minimum we would like is the property that the dimension of the image of π_j is its degree $\forall j$, which is satisfied by $\eta_v \circ \lambda_v$.

2.5. Ramanujan/Purity conjecture for $\mathrm{GL}(n)$. Call an isobaric representation π of $\mathrm{GL}_n(\mathbb{A}_F)$ *essentially tempered* iff some twist $\pi(s) = \pi \otimes |\det(\cdot)|^s$ is unitary and has tempered components π_v at every place v of F. In particular, at every unramified v, $\pi(s)_v \simeq \chi_{1,v} \boxplus \cdots \boxplus \chi_{n,v}$ with $|\chi_{i,v}| = 1$ $\forall i$, whence the name purity.

CONJECTURE 2.5.1 (Ramanujan/Purity). Every *cuspidal* automorphic representation of $\mathrm{GL}_n(\mathbb{A}_F)$ is essentially tempered.

Besides the $\mathrm{GL}(2)$ examples discussed in the second half of the introduction, this is also known to be true by Rogawski, Kottwitz, et al **[Pic]** for forms on $\mathrm{GL}(3)$ descending to cohomological forms on some $U(3)$, and by Clozel and Kottwitz for regular self-dual algebraic forms on $\mathrm{GL}(n)/\mathbb{Q}$ **[Cl1, Cl2]**. For results for function fields we refer to **[Dr, F-K, Lau]**.

It should be noted that this conjecture is not known to be true for all forms π on GL_n/F for any $n \geq 2$ and $F \geq \mathbb{Q}$. For a cuspidal unitary automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_F)$, at any unramified placed set $v = \alpha_{i,v}(\pi) = \chi_{i,v}(y_v)$, $i \leq n$, if $\pi_v \simeq \chi_{1,v} \boxplus \cdots \boxplus \chi_{n,v}$ and y_v a uniformizer at v. Then we have

$$\begin{aligned} &\text{(i)} \quad |\alpha_{i,v}(\pi)| < (Nv)^{1/2} \quad (\forall n, F) \quad \textbf{[J-S1], [Shah4]} \\ &\text{(ii)} \quad |\alpha_{i,v}(\pi)| < (Nv)^{1/5} \quad (\forall F, n=2) \quad \textbf{[Shah4]} \\ &\text{(iii)} \quad |\alpha_{i,p}(\pi)| \leq p^{5/28} \quad (F = \mathbb{Q}, n=2) \quad \textbf{[Bu-D-H-I]}. \end{aligned} \tag{2.5.2}$$

Note that a noncuspidal automorphic representation η occurring discretely in $L^2(Z(\mathbb{A})\,\mathrm{GL}_n(F)\backslash \mathrm{GL}_n(\mathbb{A}_F), \omega)$ is not essentially tempered. Indeed, by **[M-W]** one knows that every such η is an isobaric representation of the form $\pi(\frac{d-1}{2}) \boxplus \cdots \boxplus \pi(\frac{3-d}{2}) \boxplus \pi(\frac{1-d}{2})$ for some divisor $d > 1$ of n and cuspidal π of $\mathrm{GL}_{n/d}(\mathbb{A}_F)$.

An interesting generalization of the Ramanujan conjecture for general reductive groups G is Arthur's conjecture on the shape of the discrete spectrum, based on relevant homomorphisms into ${}^LG$ of the Arthur group $\mathscr{A}_F =$

$\mathscr{L}_F \times \mathrm{SU}(2)$ **[A, B-Ro2]**. This is compatible with the counterexamples found by Howe, Piatetski-Shapiro, Kurokawa, et al, to the direct generalization. It is also compatible with the work of Burger, Li, and Sarnak on the Ramanujan duals.

3. L-functions of $\mathrm{GL}(n)$: Properties and conjectures

3.1. L-functions and functoriality. Let F be a local or a global field with fixed algebraic closure F and absolute Weil group W_F **[T]**. Let $\Theta: W_F \to \mathfrak{G}_F$ be the canonical homomorphism into the absolute Galois group over F. If K/F is a finite Galois extension, we shall denote by $W_{K/F}$ (resp. $\mathfrak{G}_{K/F}$) the relative Weil group (resp. Galois group) of K/F.

For every reductive group G over F, one has the dual group ${}^LG$, which is a semi-direct product $\widehat{G} \rtimes W_F$, where $\widehat{G}$ is the complex group associated to the root datum dual to that of G (see **[Bo2, La4]**). For instance, when $G = G_I := \prod_{j\in I} \mathrm{GL}(n_j)$, I a finite indexing set, the product is direct and $\widehat{G}$ identifies with $G_I(\mathbb{C})$. When K/F is a finite extension and $G = R_{K/F}\,\mathrm{GL}(n)/K$, the group defined by restriction of scalars, $\widehat{G}$ is a product of copies of $\mathrm{GL}(n, \mathbb{C})$ indexed by $J := \mathrm{Hom}_F(K, \overline{F})$, the action of W_F being given, via Θ, by the Galois action on J.

Let V be a finite-dimensional $\mathbb{C}$-vector space, and $r: {}^LG \to \mathrm{GL}(V)$ a continuous semisimple representation. One expects to associate to every irreducible admissible (resp. automorphic) representation π of $G(F)$ (resp. $G(\mathbb{A}_F)$) L- and ε-functions $L(\pi, r; s)$ and $\varepsilon(\pi, r; s)$ **[La2]**. In the local unramified case, this is done for $G = G_I$ (our main case of interest here) as follows. There are unramified representations $\pi_j \in \mathrm{Adm}(G_{n_j}(F))$ for all $j \in I$ such that $\pi = \prod_{j\in I} \pi_j$ (acting on the tensor product of the spaces of π_j.) Each π_j, being unramified, defines a semisimple conjugacy class C_j in $G_{n_j}(\mathbb{C})$ which can be represented by a diagonal matrix $[\alpha_1, \dots, \alpha_j]$, well defined up to permutation of the α_i's. The C_j's combine to define a semisimple conjugacy class C in $G_I(\mathbb{C})$ determining the isomorphism class of π. Denoting by Fr the Frobenius conjugacy class in W_F, we may associate $\widetilde{C} = (C, \mathrm{Fr}) \subset {}^LG$ to π. One sets (with q being the order of the residue field)

(3.1.1) $$L(\pi, r; s) = \det(1 - r(\widetilde{C})q^{-s}|V)^{-1}.$$

By abuse of notation, we shall sometimes write $L(\pi_1, \dots, \pi_k, r; s)$ for $L(\pi, r; s)$ when $I = \{j_1, \dots, j_k\}$.

Now let F be global and π a unitary automorphic representation. If $\pi \simeq \otimes_\nu \pi_\nu$ with π_ν unramified outside a (necessarily finite) set S of places, the incomplete L- and ε-functions relative to (S, r) can be defined as follows:

(3.1.2)
$$L^S(\pi, r; s) = \prod_{\nu\notin S} L(\pi_\nu, r; s) \quad \text{and} \quad \varepsilon^S(\pi, r; s) = \prod_{\nu\notin S} \varepsilon(\pi_\nu, r, s).$$

By Langlands, the Euler product defining $L^S(\pi, r; s)$ converges in some right half-plane. If one admits the local Langlands conjecture (see **[Ku]**), one can also treat the ramified places and define the complete functions $L(\pi, r; s)$ and $\varepsilon(\pi, r; s)$. An independent definition is available at least in the following cases:

(3.1.3)

(a) $I = \{n\}$, $r =$ standard representation, (We write $L(\pi, s)$ for $L(\pi, r, s)$ in this case.)

(b) $I = \{n_1, n_2\}$, $r = \otimes$, the tensor product representation into $\mathrm{GL}(n_1 n_2, \mathbb{C})$,

(c) $I = \{n_1, \dots, n_k\}$, $r = \oplus$, the direct sum representation into $\mathrm{GL}(n_1 + \cdots + n_k, \mathbb{C})$,

(d) $I = \{n\}$, $r =$ symmetric square representation,

(e) $I = \{n\}$, $r =$ exterior square representation,

(f) $I = \{n_1, n_2, n_3\}$, $r = \otimes$ (triple product),

(g) $I = \{2\}$, $r =$ symmetric cube representation.

Clearly (a) is a special case of (c). It is also a special case of (b) by taking $n_1 = n_1$, $n_2 = 1$, and r the tensor product of the standard representation of $G_{n_1}(\mathbb{C})$ with the trivial representation of $G_{n_2}(\mathbb{C})$.

More importantly, one knows in these cases that the L-function admits a meromorphic continuation to the whole s-plane and admits a functional equation

$$L(\pi, r; s) = \varepsilon(\pi, r; s) L(\pi^\vee, r; 1 - s). \tag{3.1.4}$$

We refer to **[Go-J, J-PS-S1, J-S 1-3, Shah 1–5, M-W** (appendix), **Ge-J, Bu-F, Bu-G, PS-G-R]** for details. Admitting such a result for general r, we are now in a position to state a weak form of the principle of functoriality:

CONJECTURE 3.1.5. Let G, G' be reductive groups over F either of the form G_I, for some $I = \{n_1, \dots, n_k\}$, or of the form $R_{K/F}(\mathrm{GL}(n)/K)$, for K/F a finite extension, and let $f\colon {}^LG \to {}^LG$ be a continuous homomorphism that restricts to an algebraic map $\widehat{G} \to \widehat{G}'$. Then, given any irreducible automorphic (resp. admissible) representation π of $G(\mathbb{A}_F)$ (resp. $G(F)$) for a global (resp. local) field F, there exists a corresponding irreducible automorphic (resp. admissible) representation $f(\pi)$ of $G(\mathbb{A}_F)$ (resp. $G(F)$) such that we have for every continuous semisimple representation $r\colon {}^LG \to \mathrm{GL}(V)$ the following identity:

$$L(\pi, r \circ f; s) = L(f(\pi), r; s),$$

and

$$\varepsilon(\pi, r \circ f; s) = \varepsilon(f(\pi), r; s).$$

Note that this lifting problem concerns itself only with the *isomorphism classes* of representations. For general groups G, G', one may have to restrict to those representations that transfer to isobaric representations of some $\mathrm{GL}(n)$.

This has been completely solved to date only in the local Archimedean case **[La4]** and (recently) for linear groups over local fields of characteristic p. In the global (and local non-Archimedean) situation it has been solved in the following important cases:

(3.1.6)

(a) $f\colon {}^L\mathrm{GL}(2)/F \to {}^L\mathrm{GL}(3)/F =$ the *symmetric square* representation **[Ge-J]**

(b) (Base change **[A-C]**) K/F: finite cyclic extension,
$f\colon {}^L\mathrm{GL}(n)/F \to {}^L(R_{K/F}\,\mathrm{GL}(n)/K)$,

(c) $I = \{n_1, \dots, n_k\}$, $n = \sum_{i=1}^{k} n_i$,
$f\colon {}^L\mathrm{GL}_I/F \to {}^L\mathrm{GL}(n)/F$: the *direct sum* representation **[J-S1]**.

The base change result of Arthur and Clozel **[A-C]** also gives in addition a characterization of the isobaric automorphic representations of $\mathrm{GL}_n(\mathbb{A}_K)$ that are in the image of f. One also gets base change for solvable Galois extensions by successively achieving it for cyclic extensions.

Perhaps the biggest open problem is to prove the conjecture for the map (over a fixed base field F):

$$f\colon {}^L\mathrm{GL}(m) \times {}^L\mathrm{GL}(n) \to {}^L\mathrm{GL}(mn)$$

whose restriction to $\mathrm{GL}(m, \mathbb{C}) \times \mathrm{GL}(n, \mathbb{C})$ is the tensor product.

For every pair (π_1, π_2) in $\mathrm{Isob}(m, F) \times \mathrm{Isob}(n, F)$, one hopes for the existence of an isobaric automorphic representation $\pi_1 \boxtimes \pi_2 = f(\pi_1, \pi_2)$ of $\mathrm{GL}_{mn}(\mathbb{A}_F)$ such that $L(\pi_1 \boxtimes \pi_2, s) = L(\pi_1, \pi_2, \otimes; s)$. Let S be the set of places where π_1 or π_2 is ramified. Then, for every $v \notin S$, (the isomorphism class of) $\pi_{1,v} \boxtimes \pi_{2,v}$ can be defined as the one corresponding to the tensor product of the corresponding semisimple conjugacy classes $\widetilde{C}_{1,v}$ and $\widetilde{C}_{2,v}$. Hence, the problem is to show that, for some admissible irreducible π_S of $\mathrm{GL}(mn, F_S)$, the global admissible representation $\pi_S \otimes (\bigotimes_{v \notin S} \pi_{1,v} \boxtimes \pi_{2,v})$ is automorphic. If this happens for two choices π_S and π'_S, then they will be isomorphic by the strong multiplicity-one theorem.

We shall conclude this section by indicating how everything fits together in terms of the conjectural group $\mathcal{L}_F$, whose finite-dimensional representations should parametrize the isobaric (resp. admissible) irreducible representations of $\mathrm{GL}(\mathbb{A}_F)$ (resp. $\mathrm{GL}(F)$) for F global (resp. local).

First let F be local. One expects $\mathcal{L}_F$ to be isomorphic to $\mathrm{SL}_2(\mathbb{C}) \times W_F$ in the non-Archimedean case. It has become customary now to replace $\mathrm{SL}_2(\mathbb{C})$

with the compact group $\mathrm{SU}(2)$, which does not change the finite-dimensional representation theory. For any I, an L-parameter is a homomorphism

$$\varphi: \mathscr{L}_F \to {}^L G_I \tag{3.1.7}$$

such that $\varphi(W_F)$ consists of semisimple elements and the composite of φ with the natural projection pr: ${}^L G_I \to W_F$ is the identity map. Two L-parameters are equivalent iff they differ by conjugation by an element of $G_I(\mathbb{C})$. Given any continuous semisimple representation $r: {}^L G_I \to \mathrm{GL}(V)$, we may then attach L- and ε-functions $L(\varphi, r; s)$ and $\varepsilon(\varphi, r; s)$. We refer to **[De3]** and **[Ta]** for details, except that they use the Weil-Deligne group W_F' in the non-Archimedean case instead of $\mathscr{L}_F$. See §1.10 of **[B-Ro2]** for a translation to $\mathscr{L}_F$ of the definition of $L(\varphi, r; s)$. It is a fact that the L- and ε-functions depend only on the equivalence class of φ.

Now let F be global. Then there is not even a good guess as to how $\mathscr{L}_F$ should look, but nevertheless there should exist natural morphisms $i_v: \mathscr{L}_{F_v} \to \mathscr{L}_F$ at all places v. One can define a global version of an L-parameter $\varphi: \mathscr{L}_F \to {}^L G_I$ such that the local homomorphisms $\varphi_v := \varphi \circ i_v$ are all L-parameters. One sets for every $r: {}^L G_I \to \mathrm{GL}(V)$:

$$L(\varphi, r; s) = \prod_v L(\varphi_v, r; s) \quad \text{and} \quad \varepsilon(\varphi, r; s) = \prod_v \varepsilon(\varphi_v, r; s). \tag{3.1.8}$$

Let $\Phi_F(G_I)$ (resp. $\Phi_{F_v}(G_I)$) denote the set of equivalence classes of L-parameters of G_I over F (resp. F_v). Denote by $\mathrm{Isob}(G_I)$ (resp. $\mathrm{Adm}(G_I, v)$) the set of irreducible isobaric (resp. admissible) representations of $G_I(\mathbb{A}_F)$ (resp. $G_I(F_v)$). Then the most optimistic conjecture will be to have the following commutative diagram at every v:

$$\begin{array}{ccc} \Phi_F(G_I) & \xrightarrow{\beta} & \mathrm{Isob}(G_I)/\sim \\ {\scriptstyle i_v^*}\downarrow & & \downarrow{\scriptstyle \rho_v} \\ \Phi_{F_v}(G_I) & \xrightarrow{\beta_v} & \mathrm{Adm}(G_{I,v})/\sim \end{array} \tag{3.1.9}$$

where β and β_v are bijections, $/\sim$ denotes "modulo isomorphism", i_v^* sends φ to φ_v, and ρ_v associates to each π its local component π_v (well defined up to isomorphism). Moreover, β and β_v should preserve L- and ε-functions relative to any $r: {}^L G_I \to \mathrm{GL}(V)$.

REMARK 3.1.10. (a) For any $n \geq 1$, the subset of $\Phi_F(\mathrm{GL}_n)$ corresponding to *irreducible* representations of $\mathscr{L}_F$ should be in bijection with the set of (isomorphism classes of) cuspidal automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$. It is natural to want to parametrize the automorphic representations that occur discretely in $L^2(\omega) := L^2(\mathrm{GL}_n(F)Z_n(\mathbb{A}_F)\backslash \mathrm{GL}_n(\mathbb{A}_F), \omega)$ for any character of the center $Z_n(\mathbb{A}_F)$, trivial on $Z_n(F)$. For any reductive group G over F, the Arthur parameters conjecturally describe packets of such representations **[A]**. For $\mathrm{GL}(n)$, this is actually a theorem **[M-W]**. See §2.5.

(b) There has been a lot of progress in the function field case. We refer to the papers of **[Dr, F-K, Lau]**.

Now we are in a position to "explain" the principle of functoriality in terms of $\mathscr{L}_F$. Indeed, given an L-homomorphism $f: {}^LG_I \to {}^LG_J$, the conjectural map on the isobaric automorphic representations π of $G_I(\mathbb{A}_F)$ can be described as follows: By (3.1.10), there exist a unique (up to equivalence) $\varphi: \mathscr{L}_F \to {}^LG_i$ such that $\pi = \beta(\varphi)$. Then the composite map $\varphi \circ f$ defines an L-parameter of G_J. Thus $\beta(\varphi \circ f)$ is the desired class $f(\pi)$ of isobaric representations of $G_J(\mathbb{A}_F)$. The identity $L(\pi, r \circ f; s) = L(f(\pi), r; s)$ follows immediately for any $r: {}^LG_J \to \mathrm{GL}(V)$.

REMARK 3.1.11. (a) The symmetric space of GL_n/F has a complex structure iff $n = 2$ and F is totally real. Hence there are no algebro-geometric automorphic forms on $\mathrm{GL}(n)$ in general. But if π is an isobaric automorphic representation of $\mathrm{GL}(n, \mathbb{A}_F)$ corresponding to an L-parameter $\varphi: \mathscr{L}_F \to {}^L\mathrm{GL}(n)/F$, it might happen that there is a reductive group G over F of hermitian type and an L-homomorphism $j: {}^LG/F \to {}^L\mathrm{GL}(n)/F$ such that φ factors as $j \circ \rho$, for some $\rho: \mathscr{L}_F \to {}^LG/F$. Then the analog of (3.1.10) will involve a bijection β sending ρ to a packet $\prod$ of automorphic representations. If $\pi \in \prod$ admits relative Lie algebra cohomology of $(\mathfrak{G}_\infty, K_\infty)$ with coefficients in a finite-dimensional module, then $\prod$ should give rise to a "motive" M in the cohomology of a suitable Shimura variety. But this takes us far afield from our task here.

(b) Suppose $\mathscr{L}_F$ is defined as the "Galois group" of a Tannakian category $\underline{A}(F)$ relative to a tensor product $\otimes$ and a fiber functor η. Then, given any object π in this category, we could look at the tensor subcategory $\underline{A}(\pi)$ generated by π and define a group $\mathscr{L}(\pi)$ to be $\mathrm{Aut}^{\otimes}_{\underline{A}(\pi)}(\eta)$. This will be the automorphic analog of the Galois group $\mathfrak{G}(M)$ of a pure motive M. If M and π correspond, there should be an isomorphism of $\mathscr{L}(\pi)$ with $\mathfrak{G}(M)$ over $\mathbb{C}$.

3.2. The Tate conjecture and an analog. Let F be a global field and let S_∞ be the set of its Archimedean places. Then a striking result of Jacquet and Shalika **[J-S1,** Theorem 4.7] asserts the following: Let σ be an n-dimensional continuous irreducible representation over $\mathbb{C}$ of $\mathfrak{G}_F = \mathrm{Gal}(\overline{F}/F)$ which can be written as a virtual representation $\sigma_2 - \sigma_1$. Suppose each σ_i corresponds to an isobaric automorphic representation π_i such that $L^S(\sigma_i, s) = L^S(\pi_i, s)$, $i = 1, 2$, for S a finite set of places containing S_∞. Then there exists a cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_F)$ such that $L^S(\sigma, s) = L^S(\pi, s)$; in particular, $L^S(\sigma, s)$ is holomorphic outside $s = 1$. The purpose of this section is to show that this result remains valid in the context of motives of arbitrary weight if one admits the analytic Tate conjecture (see below).

Let $\underline{M}(F)$ denote a category of semisimple pure motives over F. Let M

be a pure motive of weight w and (a compatible system of) ℓ-adic realization(s) M_ℓ. Then $\mathfrak{G}_F = \mathrm{Gal}(\overline{F}/F)$ acts on the finite-dimensional $\mathbb{Q}_\ell$-vector space(s) M_ℓ. Denote by $L(M_\ell, s)$ the L-function of M_ℓ, defined in the usual way as a convergent Euler product in $\mathrm{Re}(s) > \frac{w}{2} + 1$. Granting the meromorphic continuation to the "edge" $\frac{w}{2} + 1$, one has the following

CONJECTURE 3.2.1 (Tate).
(a) w: odd $\Rightarrow L(M_\ell, \frac{w}{2} + 1) \in \mathbb{C}^*$;
(b) w: even $\Rightarrow \dim_{\mathbb{Q}_\ell} M_\ell(\frac{w}{2})^{\mathfrak{G}_F} = -\mathrm{ord}_{s=\frac{w}{2}+1} L(M_\ell, s)$.

For any $j \in \mathbb{Z}$, $M_\ell(j)$ denotes $M_\ell \otimes_{\mathbb{Q}_\ell} \mathbb{Q}_\ell(j)$. This conjecture is well known to be true for Artin-Hecke motives. Now let M be of the form $N_1 \otimes N_2^\vee$ for some pure motives N_1, N_2 over F. Then we get the following consequence of Tate's conjecture:

$$\dim_{\mathbb{Q}_\ell} \mathrm{Hom}_{\mathfrak{G}_F}(N_{1,\ell}, N_{2,\ell}) = -\mathrm{ord}_{s=\frac{w}{2}+1} L(N_{1,\ell} \otimes N_{2,\ell}^\vee, s). \tag{3.2.2}$$

Now let π_1, π_2 be unitary irreducible automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$, for some $n \geq 1$. Then a basic result of **[J-S1]** says that π_1 is isomorphic to π_2 iff $L(\pi_1, \pi_2^\vee, \otimes; s)$ has a pole at the "edge" $s = 1$. (In the unitary normalization used for automorphic forms, the functional equation always relates s to $1 - s$.) If M is a motive of weight w corresponding to π, then one should have $L(\pi_f, s) = L(M_\ell, s + \frac{w}{2})$.) We may define L- and ε-functions for objects in $\underline{A}(F)$ by additivity. Then a consequence of the above result is the following:

$$\dim_{\mathbb{C}} \mathrm{Hom}(\pi_1, \pi_2) = -\mathrm{ord}_{s=1} L(\pi_1, \pi_2^\vee, \otimes; s) \tag{3.2.3}$$

for all π_1, π_2 in $\mathrm{ob}(\underline{A}(F))$. This can be equivalently stated in terms of isobaric representations.

The following assertion is the motivic formulation of the theorem of Jacquet and Shalika alluded to above.

PROPOSITION 3.2.4. *Let M, M_1, M_2 be semisimple motives over F (with coefficients in $E \subset \mathbb{C}$ of pure weight w and ranks n, n_1, n_2 respectively such that $M \oplus M_1 = M_2$ with M simple. Fix a finite set S of places containing S_∞ such that the ℓ-adic realizations of the motives in question are unramified Galois representations outside S. Assume the Tate conjecture in the form* (3.2.2) *for the Homs among M, M_1, and M_2. Finally, suppose there exist isobaric unitary automorphic representations π_i of $\mathrm{GL}_{n_i}(\mathbb{A}_F)$, $i = 1, 2$, such that $L^S(M_i, s) = L^S(\pi_i, s - w/2)$. Then there exists a cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_F)$ such that $L^S(M, s) = L^S(\pi, s - w/2)$.*

PROOF. The proof follows directly from the argument of **[J-S1]**. For any pair of isobaric automorphic representations π, π' unramified outside S, let $\langle \pi', \pi \rangle$ denote the order of the pole at $s = 1$ of $L^S(\pi', \pi^\vee, \otimes; s)$. The

$\mathbb{Z}$-valued map $\langle\, ,\rangle$ can evidently be extended to the group of virtual isobaric automorphic representations (obtained by formally adding negatives). Similarly define $\langle N', N\rangle$ for virtual motives N, N' (of pure weight w). The hypothesis (of the Proposition) that $L^S(\pi_i, s - w/2) = L^S(M_i, s)$ implies that the respective semisimple conjugacy classes at each $v \notin S$ agree. From this one sees easily that $\langle \pi_i, \pi_j\rangle = \langle M_i, M_j\rangle$ for $i, j = 1, 2$, and also that $\langle \pi, \pi\rangle = \langle M, M\rangle$ where π is the virtual representation $\pi_2 \boxminus \pi_1$. We may write π as a finite sum: $\sum_{i\in I} m_i\eta_i$ with $m_i \in \mathbb{Z}$ and η_i cuspidal. Then by (3.2.3), $\langle \pi, \pi\rangle = \sum_{i\in I} m_i^2$. On the other hand, by (3.2.2) and by the irreducibility of the Galois representation defined by M we get $\langle M, M\rangle = 1$. This forces I to be a singleton i_0 with $m_{i_0} = \pm 1$. Write η for η_{i_0}. Suppose $\pi = -\eta$. then $\pi_1 = \eta \boxplus \pi_2$ and this implies that $\langle \pi_1, \pi_1\rangle \geq \langle \eta, \eta\rangle + \langle \pi_2, \pi_2\rangle > \langle \pi_2, \pi_2\rangle$. However, since $M_2 = M \oplus M_1$, we get by a similar reasoning that $\langle M_2, M_2\rangle > \langle M_1, M_1\rangle$, leading to a contradiction. Thus π must be η. □

As a corollary one obtains the following result of Jacquet and Shalika (see Remark 4.8 of **[J-S1]**).

COROLLARY 3.2.5 (Jacquet–Shalika). *Assume that for any finite (not necessarily normal) extension K/F with $n = [K : F]$, and for every continuous character χ: $\mathrm{Gal}(\overline{\mathbb{Q}}/K) \to \mathbb{C}^*$, we can find an isobaric automorphic representation $I(\chi)$ of $\mathrm{GL}_n(\mathbb{A}_F)$ such that $L(\chi, s) = L(I(\chi), s)$. Then Artin's conjecture holds, i.e., for every number field k and for every continuous representation σ: $\mathrm{Gal}(\overline{\mathbb{Q}}/k) \to \mathrm{GL}_n(\mathbb{C})$, $L(\sigma, s)$ has no pole outside $s = 1$.*

Indeed, being continuous, σ factors through the Galois group G of a finite extension E/F. Furthermore, by Brauer's theorem, σ can be written in the Grothendieck group of representations of G as a $\mathbb{Z}$-linear sum: $\sum_{H\in X} n_H \mathrm{Ind}_H^G(\chi_H)$, where X is a set of elementary subgroups of G and χ_H a linear character: $H \to \mathbb{C}^*$ for each H.

Let K_H be the subfield of E corresponding to H. Then by our hypothesis, there is an isobaric automorphic representation $I(\chi_H)$ of $\mathrm{GL}_n(\mathbb{A}_F)$, $n = [K_H : F]$, such that $L(\mathrm{Ind}_H^G(\chi_H), s) = L(\chi_H, s) = L(I(\chi_H), s)$. Let X^+ (resp. X^-) denote the subset of X consisting of subgroups H of G such that n_H is positive (resp. negative). Let $\pi^\pm$ be the isobaric representation: $\boxplus_{H\in X^\pm} |n_H| I(\chi_H)$. Let $\sigma_\pm$ denote the Galois representation $\bigoplus_{H\in X^\pm} |n_H| \mathrm{Ind}_H^G(\chi_H)$. The corollary now follows from Proposition 3.2.4 by setting $M = \sigma$, $M_1 = \sigma^-$, $M_2 = \sigma^+$, $\pi_1 = \pi^-$, and $\pi_2 = \pi^+$.

REMARK 3.2.6. (i) The hypothesis of Proposition 3.2.4 which is predicted by the principle of functoriality, is known to be true by **[A-C]** for cyclic extensions K/F. Hence Artin's conjecture will follow if one can define $I(\chi)$ for nonnormal extensions. For nonnormal *cubic* extensions, this is a theorem (see **[J-PS-S3]**).

(ii) In fact one can ask for the hypothesis to hold for all continuous characters χ of the global Weil group W_k (for all K/F), which is again known to be true in the cyclic and nonnormal cubic cases. Under this assumption the conclusion of the Proposition extends to Artin-Hecke motives in $\underline{M}(F)$. Interesting examples of such motives, besides the Artin ones, are given by the cohomology of abelian varieties over F of potential CM type, especially when the complex multiplications are not defined over F.

Let $\{\pi_j \mid j \in J\}$ be a set of representations for the isomorphism classes of elements of $A_0^S(F)$ for some fixed finite set S of places of F. We set

$$[k\,;\, i\,,\, j] = -\operatorname{ord}_{s=1} L^S(\pi_k^\vee\,,\, \pi_i\,,\, \pi_j\,;\, \otimes\,,\, s) \tag{3.2.6}$$

which is well defined only when the triple product L-function on the right is meromorphic at $s = 1$.

PROPOSITION 3.2.7. *Suppose* $\boxtimes$ *is defined on* $\mathrm{Isob}^S(F)$ (*Conjecture* 2.3.1). *Then the integers* $[k\,;\, i\,,\, j]$ *are well defined, nonnegative and satisfy* $(\forall i\,,\, j\,,\, k\,,\, m)$

(i) $[k\,;\, i\,,\, j] = [k\,;\, j\,,\, i]$,

(ii) $\sum_r [r\,;\, i\,,\, j][m\,;\, r\,,\, k] = \sum_r [m\,;\, i\,,\, r][r\,;\, j\,,\, k]$.

PROOF. By **[J-S1]**, assuming $\pi_i \boxtimes \pi_j$ is isobaric, we have $[k\,;\, i\,,\, j] = -\operatorname{ord}_{S=1} L^S(\pi_k^\vee\,,\, \pi_i \boxtimes \pi_j\,;\, \otimes\,,\, s)$, the multiplicity of π_k in the decomposition of $\pi_i \boxtimes \pi_j$ into a sum of (isomorphism classes of) cuspidals. Part (i) is immediate from the definition of the L-function. For part (ii) consider the multiplicity of π_m in $\pi_i \boxtimes \pi_j \boxtimes \pi_k$. (See Lemma 1.1.2 for the relevance of this Proposition.)

3.3. Converse theorems and a question. Let F be a global field and let π be an irreducible admissible representation of $\mathrm{GL}_n(\mathbb{A}_F)$ for some $n \geq 1$. A basic question is to ask under what conditions one can find an automorphic representation π' such that π and π' are *nearly equivalent*, i.e., $\pi_v \simeq \pi'_v$ for all v outside a finite set S. We note that every irreducible automorphic representation is (cf. **[La]**) essentially equivalent to an isobaric automorphic representation. A companion question is to know when π is nearly equivalent to a *cuspidal* automorphic representation. Of particular interest to us are those π such that $L^S(\pi\,,\, s)$ equals, for some finite set S, either $L^S(\chi\,,\, s)$ for a Grössencharacter of a nonnormal extension K of F, or $L^S(\pi_1\,,\, \pi_2\,,\, \otimes\,;\, s)$ for some cuspidal automorphic representations π_i of $\mathrm{GL}(n_i\,,\, \mathbb{A}_F)$, $i = 1\,,\, 2$, with $n = n_1 n_2$.

This converse problem has a long history with the pivotal $n = 2$ case solved by A. Weil and Jacquet-Langlands with further refinements due to Piatetski-Shapiro and W. Li. It was then extended to $n = 3$ by Piatetski-Shapiro. It has been known for some time, due to the work of Jacquet, Piatetski-Shapiro and Shalika, that for any n, the automorphy of π will follow once one has good analytic control over the L-functions $L(\pi\,,\, \pi'\,,\, \otimes\,;\, s)$

for all cuspidal automorphic representations π' of $\mathrm{GL}_m(\mathbb{A}_F)$ for all $m \leq n-1$. Recently J. Cogdell and I. Piatetski-Shapiro have proved two general converse theorems (for all n), which we shall now describe.

Following their lead, we shall say that π is nice relative to an irreducible cuspidal automorphic representation τ of $\mathrm{GL}_m(\mathbb{A}_F)$, $m < n$, iff $L(\pi, \tau, \otimes; s)$ and $L(\pi^\vee, \tau^\vee, \otimes; s)$ are absolutely convergent in some right half-plane, admit analytic continuations (to the whole s-plane) which are bounded in vertical strips, and admit the functional equation

$$L(\pi, \tau, \otimes; s) = \varepsilon(\pi, \tau, \otimes; s)L(\pi^\vee, \tau^\vee, \otimes; 1-s). \tag{3.3.1}$$

THEOREM 3.3.2 **[Co-PS]**. *Let π be an irreducible admissible representation of* $\mathrm{GL}(n, \mathbb{A}_F)$ *whose central character is invariant under F^*.*

(a) *Suppose π is nice relative to every cuspidal automorphic representation τ of* $\mathrm{GL}_m(\mathbb{A}_F)$, *for all $m \leq n-1$. Then π is a cuspidal automorphic representation of* $\mathrm{GL}_n(\mathbb{A}_F)$.

(b) *Fix any finite set S of places containing all the Archimedean ones, and let $\Omega_S(m)$ denote, for any $m < n$, the set of cuspidal automorphic representations of* $\mathrm{GL}_m(\mathbb{A}_F)$ *that are unramified outside S. Suppose $n \geq 3$ and π is nice relative to every τ in $\bigcup_{m \leq n-1} \Omega_S(m)$. Then π is nearly equivalent to an irreducible automorphic representation π'. In fact, $\pi_v \simeq \pi'_v$ at every non-Archimedean place v where π is unramified.*

Jacquet has suggested that part (a) above **[Co-PS**, Theorem 1] should hold assuming only that π is nice relative to all τ in $\mathrm{GL}_m(\mathbb{A}_F)$, for all $m \leq [\frac{n}{2}]$. This can be explained via $\mathscr{L}_F$ as follows. Indeed, cuspidal automorphic representations should correspond to irreducible representations of $\mathscr{L}_F$, and the reducibility of a given n-dimensional semisimple representation σ of $\mathscr{L}_F$ can be detected by tensoring σ with irreducibles τ of dimension $\leq [\frac{n}{2}]$. In fact, going much further than anyone had hoped, Cogdell and Piatetski-Shapiro have made the following conjecture.

CONJECTURE 3.3.3 **[Co-PS]**. Let π be an irreducible admissible representation of $\mathrm{GL}_n(\mathbb{A}_F)$ whose central character ω_π is invariant under F^*. Assume that π is nice relative to every idèle class character χ of F. Then π is nearly equivalent to an automorphic representation π' of $\mathrm{GL}_n(\mathbb{A}_F)$.

For $n \leq 3$, this conjecture is known to hold with $\pi \simeq \pi'$. For $n \geq 4$, π' cannot in general be isomorphic to π (see **[PS]** and also **[He]**).

REMARK 3.3.4. Let π_1, π_2 be cuspidal automorphic representations of $\mathrm{GL}(2, \mathbb{A}_F)$. Let π be an irreducible admissible representation of $\mathrm{GL}_4(\mathbb{A}_F)$ such that for all v outside the set S where π_1 or π_2 is ramified, $\pi_v \simeq \pi_{1,v} \boxtimes \pi_{2,v}$. Assume the truth of a stronger form of Theorem 3.3.2 asserting that it suffices to test the niceness of π relative to (all) cuspidals τ of $\mathrm{GL}_m(\mathbb{A}_F)$ for $m \leq 2$. For every idèle class character χ of F, we have $L^S(\pi \otimes \chi, s) = L^S(\pi_1, \pi_2 \otimes \chi, \otimes; s)$, which is well understood. If π_3 is a cuspidal representation of $\mathrm{GL}_2(\mathbb{A}_F)$, then the triple product L-function $L(\pi_1, \pi_2, \pi_3, \otimes; s)$ now has an integral representation **[G, PS-R]** giving a meromorphic continuation. If one can prove the boundedness in vertical strips and control the poles, one will then be able to prove that π is nearly equivalent to an automorphic representation.

We end this section with a question: it is motivated by the following simple, but hopefully useful, observation. Let σ be an n-dimensional semisimple complex representation of $\mathcal{L}_F$, or of W_F if one wants to be concrete. Then to detect the possible reducibility of σ, there is another way besides tensoring σ with irreducibles of dimension $\leq [\frac{n}{2}]$, which gives a sufficient, but not necessary, condition. What one can do is to take the jth exterior power $\bigwedge^j(\sigma)$, for every $j \leq [\frac{n}{2}]$, and look for a line stable under $\mathcal{L}_F$.

QUESTION 3.3.5. Let π be a unitary irreducible admissible representation of $\mathrm{GL}_n(\mathbb{A}_F)$ whose central character is trivial on F^*. Assume that, for every $j \leq [\frac{n}{2}]$ and for every unitary character χ of W_F, the L-functions $L(\pi, \bigwedge^j \otimes \chi; s)$ and $L(\pi^\vee, \bigwedge^j \otimes \overline{\chi}; s)$ are absolutely convergent in some right half-plane, admit meromorphic continuations to the whole s-plane with no pole anywhere, bounded in vertical strips, and satisfy the functional equation: $L(\pi, \bigwedge^j \otimes \chi, s) = \varepsilon(\pi, \bigwedge^j \otimes \chi, s) L(\pi^\vee, \bigwedge^j \otimes \overline{\chi}, 1-s)$. Then is π nearly equivalent to a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_F)$? When $n = 4$, does the conclusion still hold if the analytic behavior of $L(\pi, \bigwedge^2, \chi, s)$ is relaxed to allow it to have a simple pole at $s = 1$ for a unique character χ?

It must be noted, however, that the exterior jth power L-functions are not at all understood for $j > 2$ so that the only cases where this approach may be of some use are $n = 4$ and $n = 5$.

Given Conjecture 3.3.3 and the remark above, we only have to justify the second part of the question. Suppose σ is a reducible 4-dimensional $\mathbb{C}$-representation of $\mathcal{L}_F$. It must admit an irreducible summand η of dimension ≤ 2. If $\dim(\eta) = 1$, then it can be detected by the pole of the standard (degree 4) L-function of σ twisted by η^{-1}. So assume that $\sigma = \eta \oplus \eta'$ with η, η' irreducible of dimension 2. Then $\bigwedge^2(\sigma) = \eta \otimes \eta' \oplus \det(\eta) \oplus \det(\eta')$. Write $\chi^{-1} = \det(\eta)$ and $\chi'^{-1} = \det(\eta')$. If $\chi = \chi'$, then $L(\bigwedge^2(\sigma) \otimes \chi, s)$ has a pole of order 2 at the edge. If $\chi \neq \chi'$, then $L(\bigwedge^2(\sigma) \otimes \chi, s)$ and $L(\bigwedge^2(\sigma) \otimes \chi', s)$ both have poles. The disadvantage of this formulation is that it will miss those irreducibles σ (of $\mathcal{L}_F$) that have two distinct polarizations. However, here is an interesting case where it does apply. Let η be an irreducible 2-dimensional representation of $\mathcal{L}_F$ such that the symmetric powers $S^j(\eta)$ are irreducible for all $j \leq 4$. Put $\sigma = S^3(\eta)$. Then $\bigwedge^2(\sigma) = (\det(\eta) \otimes S^4(\eta)) \oplus \det(\eta)^3$. For $n = 5$, a crucial case is $\sigma = S^4(\eta)$.

3.4. Self-dual representations. Let π be a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_F)$ with unitary central character such that $\pi \simeq \pi^\vee$. Then the corresponding representation σ of $\mathscr{L}_F$ will be self-dual and admit an invariant symmetric or alternating bilinear form. It is of interest to know which one it is.[5] One possible way to decide is to use the factorization of the L-function:

$$L(\pi, \pi; \otimes, s) = L\left(\pi, \bigwedge^2, s\right) L(\pi, S^2, s). \tag{3.4.1}$$

We know that $L(\pi, \pi; \otimes, s)$ has a simple pole at $s = 1$ (since $\pi \simeq \pi^\vee$). One also knows that $L(\pi, \bigwedge^2, s)$ and $L(\pi, S^2, s)$ do not have poles of order ≥ 2 or zeros at $s = 1$. Thus exactly one of the two L-functions on the right of (3.4.1) has a pole (of order 1) at $s = 1$. By the Tate conjecture this determines the type of the bilinear form σ admits. One has by **[J-S3]** the following criterion for $n = 2m$:

(3.4.2)

$$\operatorname{ord}_{S=1} L\left(\pi, \bigwedge^2, s\right) \neq 0 \Leftrightarrow \iint \varphi\left[\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} g & 0 \\ 0 & g \end{pmatrix}\right] \psi(\operatorname{tr}(x))\, dx\, dg \neq 0$$

for some φ in the space of π, where ψ denotes a nontrivial character of F^+ and the integral is over the space: $(Z(\mathbb{A}_F)\,\mathrm{GL}_m(F)\backslash \mathrm{GL}_m(\mathbb{A}_F)) \times (M_m(F)\backslash M_m(\mathbb{A}_F))$. An essentially equivalent criterion is given in **[Bu-F]** by a different approach. Moreover, Bump and Ginzburg have given (in **[Bu-G]**) a consequence of the symmetric square L-function having a pole at $s = 1$ as the nonvanishing of the scalar product of φ with a suitable Θ-function.

A very different approach to self-dual representations π of $\mathrm{GL}_n(\mathbb{A}_\mathbb{Q})$ is taken by Clozel in **[Cl2]**. Suppose (for technical reasons) that π has supercuspidal components at two primes. The method of **[Cl2]** is to base change π to an imaginary quadratic extension E and then to descend to a suitable unitary group (defined by a division algebra), which has a Shimura variety. Using the work of Kottwitz he shows in particular that the (self-dual) π which are algebraic and *regular* at infinity satisfy the Ramanujan (purity) conjecture. For $n = 3$, Rogowski's work (see **[Ro, Pic]**) leads to very complete results on representations descending to $U(3)$.

We also refer to **[H]** and **[B-H-R]** for some other directions, the latter dealing with Galois conjugation for symplectic self-dual π on GL(4) which are algebraic and *semi-regular*, i.e., when the restriction of $\sigma(\pi_\infty)$ to $\mathbb{C}^*$ has (algebraic) constituents of multiplicity at most 2.

3.5. A refinement of the strong multiplicity-one theorem. Let F be a number field and S a set of finite places of F. One says that S has

[5]One could ask the same question about the invariant bilinear form $B = \otimes_v B_v$ on the space of $\pi = \otimes_v \pi_v$. It turns out that for every finite v, B_v is always symmetric. For $n = 2$ and π_v in the discrete series, if π'_v denotes the associated (finite-dimensional) representation of D_v^*, D_v: the unique quaternion division algebra over F_v, then the invariant bilinear form on the space of π'_v is symmetric whenever σ_v is symplectic. For details and consequences we refer to the forthcoming article, "On the symplectic root numbers of two-dimensional Galois representations", by D. Prasad and the author.

(Dirichlet) density δ iff the limit $\lim_{s\to 1+}\{\sum_{v\in S}(Nv)^{-s}/\log(1/(s-1))\}$ exists and equals δ. We would like to propose the following

CONJECTURE 3.5.1. Let F be a number field, n an integer ≥ 1, and S a set of finite places of F of Dirichlet density $\delta < 1/(2n^2)$.

(a) If π, π' are cuspidal automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$ such that $\pi_v \simeq \pi'_v$ $\forall v \notin S$, then $\pi \simeq \pi'$.

(b) If σ_ℓ, σ'_ℓ are (compatible systems of) ℓ-adic absolutely irreducible representations of $\mathrm{Gal}(\overline{F}/F)$ of dimension n such that $\mathrm{tr}\,\sigma_\ell(\mathrm{Fr}_v) = \mathrm{tr}(\sigma'_\ell(\mathrm{Fr}_v)$ $\forall v \notin S$, then $\sigma_\ell \simeq \sigma'_\ell$.

For S a finite set part (a) is the (usual) strong multiplicity-one theorem, while part (b) holds for any S of density zero by the Čebotarev density theorem. Part (a) of the conjecture is proved for $n = 2$ in **[R1]**; for $n = 1$ it is a well-known consequence of Hecke's equidistribution results. We refer to **[R2]** for weaker results for $n > 2$ and for a finer formulation. The obstruction to proving part (a) (resp. part (b)) is the lack of knowledge of the truth of the general Ramanujan conjecture for $\mathrm{GL}(n)$ (resp. the conjecture on meromorphic continuation to the edge line and the analytic Tate conjecture at the edge).

In the case of a finite group G, if σ, σ' are two $\mathbb{C}$-representations of dimension n with identical traces outside a set S with $|S|/|G| < 1/(2n^2)$, then $\langle \mathrm{tr}(\sigma), \mathrm{tr}(\sigma')\rangle$ can be checked to be nonzero; hence $\sigma \simeq \sigma'$. In the other direction one has the following example of K. Buzzard, B. Edixhoven, and R. Taylor. Let $m \geq 1$, C the center of Q_8 and H the quotient of Q_8^m/K where K is the degree 0 part of C^m. Denote by τ the unique irreducible of Q_8 of dimension 2 and let η be the representation of H defined by $\tau^{\otimes m}$. Put $G = H \times \{\pm 1\}$ and define σ (resp. σ') to be $\eta \otimes 1$ (resp. $\eta \otimes \mathrm{sgn}$). Then σ, σ' are irreducible of dimension $n = 2^m$ with identical traces outside $S = \{(1, -1), (-1, -1)\}$. Note that $|S|/|G| = 1/2^{2m+1} = 1/(2n^2)$.

This example shows that part (a) of the conjecture is also sharp at least for $n = 2^m$.[6] Indeed by the work of Arthur and Clozel **[A-C]** there exist, since G is nilpotent, cuspidal automorphic representations π, π' of $\mathrm{GL}_n(\mathbb{A}_F)$ and Galois extensions K/F with group G such that $L(\pi_v, s) = L(\pi'_v, s)$ for all v outside S', the union of S with the finite set of ramified places. On the other hand, if $\pi \simeq \pi'$ we should have $1 = -\mathrm{ord}_{S=1} L^{S'}(\overline{\pi} \times \pi', s) = -\mathrm{ord}_{S=1} L^{S'}(\overline{\sigma} \otimes \sigma', s)$, which would imply that $\sigma \simeq \sigma'$, a contradiction.

REMARK 3.5.2. Given a specific pair (π, π') of cuspidal automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$, it suffices by a theorem of C. Moreno **[Mo]** to check the isomorphism $\pi_v \simeq \pi'_v$ for all finite v with N_v bounded by a constant

[6] We have recently learned from J.-P. Serre such a Galois construction for *all* n. His argument appeals to the existence, proved in [**Se**, vol. II, pp. 611–612], of a finite group H and a representation $\tau\colon H \to \mathrm{GL}_n(\mathbb{C})$ such that the subset X of elements in H of nonzero trace has cardinality $\frac{1}{n^2}|H|$. The representations $\tau \otimes 1$ and $\tau \otimes \mathrm{sgn}$ of $G := H \times \{\pm 1\}$ furnish the desired example.

C depending on the conductors of π, π' and their infinity types.

Bibliography

[A] J. Arthur, *Unipotent automorphic representations: Conjectures*, Astérisque **171–172** (1989), 13–71.

[A-Cl] J. Arthur and L. Clozel, *Simple algebras, base change, and the advanced theory of the trace formula*, Ann. of Math. Stud., vol. 120, Princeton Univ. Press, Princeton, NJ, 1989.

[B-H-R] D. Blasius, M. Harris, and D. Ramakrishnan, *Coherent cohomology, limits of discrete series and Galois conjugation*, to appear in the Duke Math Journal (1993/94).

[B-Ro1] D. Blasius and J. Rogawski, *Motives for Hilbert modular forms*, preprint (1992).

[B-Ro2] ——, *Zeta functions of Shimura varieties*, these Proceedings, vol. 2, pp. 525–571.

[Bo1] A. Borel, *Introduction to automorphic forms*, Proc. Sympos. Pure Math., vol. 9, Amer. Math. Soc., Providence, RI, 1966, pp. 199–210.

[Bo2] ——, *Automorphic L-functions*, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 27–61.

[Bo-J] A. Borel and H. Jacquet, *Automorphic forms and automorphic representations*, Proc. Sympos. Pure Math., part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 189–202.

[Bu-D-H-I] D. Bump, W. Duke, J. Hoffstein, and H. Iwaniec, *An estimate for the Hecke eigenvalues of Maass forms*, Duke Math J., no. 4 (1992), 75–81.

[Bu-F] D. Bump and S. Friedberg, *The exterior square automorphic L-functions on* $GL(n)$, Israel Math. Conference Proceedings III, 1990, pp. 47–66.

[Bu-G] D. Bump and D. Ginzburg, *Symmetric square L-functions on* $GL(r)$, Ann. Math. **136** (1992), 137–205.

[Ca] H. Carayol, *Sur les représentations l-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. École Norm. Sup. (4) (1986), 409–468.

[Cl1] L. Clozel, *Motifs et formes automorphes: Applications du principe de fonctorialité*, Perspect. Math., vol. 11, Academic Press, San Diego, CA, 1990.

[Cl2] ——, *Représentations galoisiennes associées aux représentations automorphes autoduales de* $GL(n)$, Inst. Hautes Études Sci. Publ. Math. **73** (1991).

[Co-PS] J. Cogdell and I. Piatetski-Shapiro, *Converse theorems for* GL_n, preprint (1990).

[Da] E. C. Dade, *Answer to a question of R. Brauer*, J. Algebra **1** (1964), 1–4.

[De1] P. Deligne, *Formes modulaires et représentations l-adiques*, Sém. Bourbaki 1968/69, no. 355, Lecture Notes in Math., vol. 179, Springer, Berlin and New York.

[De2] ——, *Catégories Tannakiennes*, Grothendieck Festschrift, Prog. Math., part II, vol. 86, 1990, pp. 111–195.

[De3] ——, *Les constantes des équations fonctionnelles des fonctions* L, Lectures Notes in Math., vol. 349, Springer, Berlin and New York, 1973, pp. 501–595.

[De4] ——, *Problems of present day mathematics* V, Mathematical Developments Arising from Hilbert's Problems (F. Browder, ed.), Proc. Sympos. Pure Math **28**, part 1, 1976, pp. 41–44.

[De-M] P. Deligne and J. S. Milne, *Tannakian categories*, Lecture Notes in Math., vol. 900, Springer, Berlin and New York, 1982, pp. 101–228.

[De-Se] P. Deligne and J.-P. Serre, *Formes modulaires de poids* 1, Ann. Sci. École Norm. Sup. (4) 7 (1974), 507–530.

[Dr] V. G. Drinfeld, *Two-dimensional l-adic representations of the fundamental group of a curve over a finite field and automorphic forms on* $GL(2)$, American J. Math. **105** (1983), 85–114.

[F] G. Faltings, *Endlichleitssätz für Abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), 349–366.

[Fl] D. Flath, *Decomposition of representations into tensor products*, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI (1979), 179–183.

[F-K] Y. Flicker and D. Kazhdan, *Geometric Ramanujan conjecture and Drinfeld reciprocity law*, Number Theory, Trace Formulas and Discrete Groups, Academic Press, San Diego, CA, 1989, pp. 201–218.

[Ga] P. X. Gallagher, *Invariants for finite groups*, Adv. Math. **34** (1979), 46–57.

[G] P. Garrett, *Decomposition of Eisenstein series: Rankin triple products*, Ann. of Math. (2) **125** (1987), 209–235.

[Ge-J] S. Gelbart and H. Jacquet, *A relation between automorphic representations of* GL(2) *and* GL(3), Ann. Sci. École Norm. Sup. (4) **11** (1979), 471–542.

[Ge-Shah] S. Gelbart and F. Shahidi, *Analytic properties of automorphic L-functions*, Perspect. Math., vol. 6, Academic Press, San Diego, CA, 1988.

[Go-J] R. Godement and J. Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Math., vol. 260, Springer, Berlin and New York, 1972.

[H] M. Harris, *Hodge-de Rham structures and periods of automorphic forms*, these Proceedings, vol. 2, pp. 573–624.

[H-S-T] M. Harris, D. Soudry, and R. Taylor, *ℓ-adic representations associated to modular forms over imaginary quadratic fields*. I, Invent. Math. **112** (1993), 377–411.

[He] G. Henniart, *Quelques remarques sur les théorèmes réciproques*, Israel J. Math. Conference Proceedings, vol. 2, 1990.

[Hi] H. Hida, *Galois representations into* $GL_2(\mathbb{Z}_p[[X]])$ *attached to ordinary cusp forms*, Invent. Math. **85** (1986), 546–613.

[J-L] H. Jacquet and R. P. Langlands, *Automorphic forms on* GL(2), Lecture Notes in Math., vol. 114, Springer, Berlin and New York, 1970.

[J-PS-S1] H. Jacquet, I. Piatetski-Shapiro, and J. Shalika, *Rankin-Selberg convolutions*, Amer. J. Math. **105** (1983), 367–464.

[J-PS-S2] ——, *Conducteur des représentations génériques du groupe linéaire*, C. R. Acad. Sci. Paris Sér. I Math. **292** (1981), 611–616.

[J-PS-S3] ——, *Automorphic forms on* GL(3), Ann. of Math. (2) **109** (1979), 213–258.

[J-S1] H. Jacquet and J. Shalika, *Euler products and the classification of automorphic forms* I & II, Amer. J Math. **103** (1981), 499–558, 777–815.

[J-S2] ——, *Rankin-Selberg convolutions: Archimedean theory*, Israel J. Math. Conference Proceedings **1** (1990), 125–207.

[J-S3] ——, *Exterior square L-functions*, Perspect. Math., vol. 11, Academic Press, San Diego, CA, 1990, pp. 143–226.

[Ja] U. Jannsen, *Motives, numerical equivalence and semisimplicity*, Invent. Math. **107** (1992), 447–452.

[Kn] A. W. Knapp, *Representations of* $GL_2(\mathbb{R})$ *and* $GL_2(\mathbb{C})$, in *Automorphic Forms, Representations, and L-functions*, Proc. Sympos. Pure Math., vol. 33, part I, Amer. Math. Soc., Providence, RI, 1977, pp. 87–91.

[Ku] S. Kudla, *The local Langlands conjecture: The non-Archimedean case*, these Proceedings, vol. 2, pp. 365–391.

[La1] R. P. Langlands, *Automorphic representations, Shimura varieties and motives. Ein Märchen*, Proc. Sympos. Pure Math. **33**, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 205–246.

[La2] ——, *Problems in the theory of automorphic forms*, Lecture Notes in Math., vol. 170, Springer, Berlin and New York, 1970.

[La3] ——, *Base change for* GL(2), Ann. of Math. Stud., vol. 96, Princeton Univ. Press, Princeton, NJ, 1980.

[La4] ——, *On the classification of real algebraic groups*, Math. Surveys Monographs, vol. *31, Amer. Math. Soc., Providence, RI, 1989, pp. 101–170.*

[La5] ——, *On the notion of an automorphic representation*, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 203–207.

[Lau] G. Laumon, *Cohomology with compact supports of Drinfel'd modular varieties*, Prepublications (91-01), Univ. Paris-sud, Orsay (1991).

[L-R-S] G. Laumon, M. Rapoport, and Stuhler, *D-elliptic sheaves and the Langlands correspondence*, preprint (1992).

[Mac] S. Mac Lane, *Categories for the working mathematician*, Springer, Heidelberg, 1972.

[M-W] C. Moeglin and J.-L. Waldspurger, *Le spectre residuel de* GL(n), Ann. Sci. École Norm. Sup. (4) **22** (1989), 605–674.

[Mo] C. Moreno, *Analytic proof of the strong multiplicity one theorem*, Amer. J. Math. **147** (1985), 163–206.

[M-R] V. K. Murty and D. Ramakrishnan, *Cycles on quaternionic Shimura varieties*, article in preparation.

[Od] T. Oda, *Hodge structures of Shimura varieties attached to the unit groups of Shimura varieties*, Adv. Stud. Pure Math. **2** (1983), 15–36.

[Oh] M. Ohta, *Hilbert modular forms of weight one and Galois representations*, Progr. Math. **46** (1984), 333–352.

[Pic] *Zeta functions of Picard modular surfaces* (R. P. Langlands and D. Ramakrishnan, eds.), CRM Publications, 1991.

[PS] I. Piatetski-Shapiro, *Zeta functions of* $\mathrm{GL}(n)$, Univ. of Maryland, preprint, 1976.

[PS-G-R] I. Piatetski-Shapiro, S. Gelbart, and S. Rallis, *Explicit constructions of automorphic L functions*, Lecture Notes in Math., no. 1254, Springer-Verlag, Berlin and New York, 1987, vi+152 pp.

[R1] D. Ramakrishnan, *A refinement of the strong multiplicity one for* $\mathrm{GL}(2)$, appendix to [T2], Invent. Math. (1993) (to appear).

[R2] ——, *A fine multiplicity one theorem for* $\mathrm{GL}(n)$, in preparation.

[R-Z] H. Reimann and T. Zink, *The good reduction of Shimura varieties associated to quaternion algebras over a totally real number field*, preprint (1991).

[Ri1] K. Ribet, *On modular representations of* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), 431–476.

[Ri2] ——, *Galois representations attached to eigenforms with Nebentypus*, Modular Functions of One Variable V, Lecture Notes in Math., no. 601, Springer-Verlag, Berlin and New York, 1977, pp. 17–52.

[Ro] J. Rogawski, *Automorphic representations of unitary groups in three variables*, Ann. of Math. Stud., vol. 123, Princeton Univ. Press, Princeton, NJ, 1990.

[Ro-T] J. Rogawski and J. Tunnell, *On Artin L-functions associated to Hilbert modular forms of weight one*, Invent. Math. **14** (1983), 1–42.

[Sa] N. Saavedra Rivano, *Catégories Tannakiennes*, Lecture Notes in Math., no. 265, Springer-Verlag, Berlin and New York, 1972.

[Sch] A. J. Scholl, *Motives for modular forms*, Invent. Math. **100** (1990), 419–430.

[Se] J.-P. Serre, *Collected papers*, Springer, Berlin and New York, 1985.

[Shah1] F. Shahidi, *On certain L-functions*, Amer. J. Math. **103** (1981), 297–356.

[Shah2] ——, *Local coefficients as Artin factors for real groups*, Duke Math. J. **52** (1985), 973–1007.

[Shah3] ——, *Local coefficients and normalization of intertwining operators for* $\mathrm{GL}(n)$, Compositio Math. **48** (1983), 271–295.

[Shah4] ——, *On the Ramanujan conjecture and finiteness of poles for certain L-functions*, Ann. of Math. (2) **127** (1988), 547–584.

[Shah5] ——, *Automorphic L-functions: A survey*, Automorphic Forms, Shimura Varieties, and *L*-functions (L. Clozel and J. S. Milne, eds.); Perspect. Math., vol. 10, Academic Press, San Diego, CA, 1990, pp. 415–437.

[Shal] J. Shalika, *The multiplicity one theorem for* GL_n, Ann. of Math. (2) **100** (1974), 171–193.

[Sh1] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Iwanami Shoten, Tokyo, and Princeton Univ. Press, Princeton, NJ, 1971.

[Sh2] ——, *On the zeta function of an abelian variety with complex multiplication*, Ann. of Math. (2) **94** (1971), 504–533.

[Sh-T] G. Shimura and Y. Taniyama, *Complex multiplication of abelian varieties and its applications to number theory*, Publ. Math. Soc. Japan **6** (1961).

[Ta] J. Tate, *Number theoretic background*, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 3–26.

[T1] R. Taylor, *On Galois representations associated to Hilbert modular forms*, Invent. Math. **98** (1989), 265–280.

[T2] ——, *l-adic representations associated to modular forms over imaginary quadratic fields*. I, Invent. Math. (1993) (to appear).

[Tu] J. Tunnell, *Artin's conjecture for representations of octahedral type*, Bull. Amer. Math. Soc. (N. S.) **5** (1981), 173–175.

[W] N. Wallach, *Representations of reductive Lie groups*, Proc. Sympos. Pure Math., vol. 33, part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 71–91.

[We] A. Weil, *Dirichlet series and automorphic forms*, Lecture Notes in Math., vol. 189, Springer, Berlin and New York, 1970.

[Wi] A. Wiles, *On ordinary λ-adic representations associated to modular forms*, Invent. Math. **94** (1988), 529–573.

DEPARTMENT OF MATH, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Shimura Varieties and Motives

J. S. MILNE

ABSTRACT. Deligne has expressed the hope that a Shimura variety whose weight is defined over $\mathbb{Q}$ is the moduli variety for a family of motives. Here we prove that this is the case for "most" Shimura varieties. As a consequence, for these Shimura varieties, we obtain an explicit interpretation of the canonical model and a modular description of its points in any field containing the reflex field. Moreover, when we assume the existence of a sufficiently good theory of motives in mixed characteristic, we are able to obtain a description of the points on the Shimura variety modulo a prime of good reduction.

Contents

Introduction
Notations and conventions
1. Abelian motives and their Mumford-Tate groups
2. Moduli of motives
3. Shimura varieties as moduli varieties
4. The points on a Shimura variety modulo a prime of good reduction

Introduction

A Shimura variety $\mathrm{Sh}(G, X)$ is a projective system of algebraic varieties over $\mathbb{C}$. The data needed to define it are a reductive group G over $\mathbb{Q}$ together with a $G(\mathbb{R})$-conjugacy class X of homomorphisms $h: \mathbb{C}^{\times} \to G(\mathbb{R})$ satisfying conditions sufficient to ensure that X is, in a natural way, a finite union of bounded symmetric domains.

In a small number of cases, the Shimura variety can be interpreted as a moduli variety for abelian varieties with the additional structure of an endomorphism ring, a polarization, and a level structure. Such a Shimura

1991 *Mathematics Subject Classification*. Primary 14K20, 14D20, 11F55.
Partially supported by the National Science Foundation.
This paper is in final form and no version of it will be submitted for publication elsewhere.

variety is said to be of PEL-type. For example, the Shimura variety defined by a group of symplectic similitudes (a Siegel modular variety) or by a group GL_2 over a totally real field (a Hilbert-Blumenthal variety) is of PEL-type. When such a modular interpretation exists, it is a great help in studying the variety, for example, in constructing a model of the variety over a number field or, better, the ring of integers in the number field, and in studying the compactification of the model. See Faltings and Chai **[27]** for Siegel modular varieties and Rapoport **[58]** for Hilbert-Blumenthal varieties. In fact, the realization of elliptic modular curves as the moduli varieties of elliptic curves with level structure has been extraordinarily fruitful both for the study of the modular curves and for the study of elliptic curves.

If the group G can be embedded into a group of symplectic similitudes $G(\psi)$ in such a way that the elements h of X define on V a Hodge structure of type $\{(-1,0),(0,-1)\}$ for which $\pm 2\pi i\psi$ is a polarization, then $\mathrm{Sh}(G, X)$ has an interpretation as a moduli variety for abelian varieties with (absolute) Hodge cycle and level structure. Such a Shimura variety is said to be of Hodge type. Unfortunately, this interpretation is valid only in characteristic zero because Hodge cycles only make sense there (so long as the Hodge conjecture for abelian varieties remains open), but the interpretation can again be used to study models and compactifications of the Shimura variety over number fields—see Brylinski **[12]**.

A Shimura variety whose weight is defined over $\mathbb{Q}$ can always be interpreted (over $\mathbb{C}$) as a parameter space for Hodge structures, and Deligne notes: "Pour interpréter des structures de Hodge de type plus compliqué, on aimerait remplacer les variétés abéliennes par des "motifs" convenables, mais il ne s'agit encore que d'un rêve." (Deligne **[18,** p. 248]). The main purpose of this article is to provide such an interpretation when G has no factors of type E_6, E_7, or certain types D, and hence to realize the Shimura variety as a moduli variety for motives. As for Shimura varieties of Hodge type, the interpretation is valid only in characteristic zero and depends in a crucial way on Deligne's theorem that all Hodge cycles on abelian varieties are absolutely Hodge **[19]**).

I now describe the contents of the article in more detail.

Betti cohomology provides a functor from the category of motives over $\mathbb{C}$ (defined using algebraic cycles) to the category $\mathbf{Hdg}_{\mathbb{Q}}$ of polarizable rational Hodge structures. The Hodge conjecture predicts that the functor is fully faithful, but there is no description, not even conjectural, for its essential image.

Define the category of abelian motives $\mathbf{Mot}^{\mathrm{ab}}(\mathbb{C})$ over $\mathbb{C}$ to be the tensor subcategory of the category of motives over $\mathbb{C}$ (defined using absolute Hodge cycles) generated by the motives of abelian varieties. The main theorem of **[19]** implies that the Betti fibre functor

$$\mathbf{Mot}^{\mathrm{ab}}(\mathbb{C}) \longrightarrow \mathbf{Hdg}_{\mathbb{Q}}$$

is fully faithful. In §1 we describe the essential image of this functor, i.e., we describe the Hodge structures that are the Betti realization of an abelian motive, and we classify the reductive groups that arise as the Mumford-Tate group of an abelian motive. The key ingredients in the proof of the classification are Satake's results on symplectic embeddings of semisimple groups and the well-known fact that every polarizable Hodge structure with commutative Mumford-Tate group is the Betti realization of an abelian motive.

In §2 we investigate, along the lines of Griffiths **[35]**, Deligne **[18]**, the problem of realizing a motive over $\mathbb{C}$, endowed with the structure provided by a family of tensors, as a member of a universal family. In general it is not known how to do this, but we show that it is possible when the motive is abelian.

The study of the moduli of motives leads very naturally to the notion of a Shimura variety, and in §3 we classify the Shimura varieties that are moduli varieties for abelian motives; the class includes all but those whose defining group has factors of type E_6, E_7, or certain types D. For these Shimura varieties we are able to obtain a more direct proof of the existence of canonical models than that in **[18]**, and, for those whose weight is rational, we deduce a modular interpretation of the canonical model.

As noted above, the theory is restricted to characteristic zero. However, when we assume the existence of a sufficiently good theory of motives in mixed characteristic, we can extend the description of the Shimura variety as a moduli variety for abelian motives to characteristic p and use this to obtain an explicit description of the points of a reduction of the Shimura variety with coordinates in the algebraic closure of a finite field. The statement we arrive at is (essentially) the main conjecture of Langlands and Rapoport **[46]**.

Roughly speaking, when one combines the results of §4 of this paper with the results of §§5–7 of Milne **[50]** and Theorem 7.1 of Kottwitz **[41]**, then one arrives at the following statement: for Shimura varieties of abelian type and rational weight, Langlands's conjecture on the contribution of the variety itself (i.e., ignoring its boundary) to the local component of the zeta function at a good prime is a consequence of standard conjectures in algebraic geometry and representation theory, the most significant of which are the existence of a good theory of motives in mixed characteristic and the fundamental lemma.

It is a pleasure to thank J.-M. Fontaine, W. Messing, and A. Ogus for their help with p-adic cohomology, G. Prasad for his help with buildings, S. Zucker for his help with the proof of (2.41), and C.-L. Chai for his comments on an earlier draft.

Notations and conventions

An affine group scheme over a field k is said to be *algebraic* when it is of finite type over k. Every affine group scheme is the projective limit of its algebraic quotients. We allow a simple algebraic group to have a finite centre. For an algebraic group G over a field k, G^0 denotes the connected component of G containing 1 for the Zariski topology; when $k = \mathbb{R}$, $G(\mathbb{R})^+$ denotes the connected component containing 1 for the *real topology.*

For a reductive group G, G^{der} denotes the derived group of G, G^{sc} the simply connected covering group of G^{der}, $Z(G)$ the centre of G, G^{ad} the adjoint group $G/Z(G)$ of G, and G^{ab} the maximal abelian quotient G/G^{der} of G. If $G = G^{\text{ad}}$, then G is called an *adjoint* group. These notations extend in an obvious way to pro-reductive affine group schemes: write $G = \varprojlim G'$ (limit over the algebraic quotients of G), and set $G^* = \varprojlim G'^*$.

The map $\text{ad}\colon G \to \text{Aut}(G)$, sending an element of G to the inner automorphism it defines, factors through G^{ad},

$$G \xrightarrow{\text{ad}} G^{\text{ad}} \longrightarrow \text{Aut}(G).$$

The map sending an element g of G to the differential of $\text{ad}(g)$,

$$G \longrightarrow \text{GL}(\mathfrak{g}), \quad \mathfrak{g} = \text{Lie}\, G,$$

is denoted by Ad.

Let $F\colon \mathbf{A} \to \mathbf{B}$ be a functor, and let A be an object of $\mathbf{A}$. A map $\beta\colon B \to F(A)$ is said to *generate* A if the following holds: for any subobject $A' \hookrightarrow A$ of A such that β factors through $F(A') \to F(A)$, the map $A' \to A$ is an isomorphism.

If S is a set of objects in a Tannakian category $\mathbf{T}$ over a field k, we define the *tensor category generated by* S to be the smallest full subcategory $\mathbf{T}'$ of $\mathbf{T}$ containing S and closed under the formation of subobjects, quotient objects, direct sums, duals, and finite tensor products (hence it contains with any object X of $\mathbf{T}$, all objects isomorphic to X). It is again a Tannakian category over k.

We often write $V(R')$ for $V \otimes_R R'$. Other notations agree with those of Milne **[51]** except that here we denote the field of fractions of the Witt vectors by B.

1. Abelian motives and their Mumford-Tate groups

Throughout this section, k is an algebraically closed field of characteristic zero.

Hodge structures: definitions. We write $\mathbb{S}$ for the torus $\text{Res}_{\mathbb{C}/\mathbb{R}}\, \mathbb{G}_m$ over $\mathbb{R}$; thus

$$\mathbb{S}(\mathbb{R}) = \mathbb{C}^\times, \qquad \mathbb{S}_\mathbb{C} = \mathbb{G}_m \times \mathbb{G}_m.$$

The last identification is made in such a way that the map

$$\mathbb{S}(\mathbb{R}) = \mathbb{C}^\times \hookrightarrow \mathbb{C}^\times \times \mathbb{C}^\times = \mathbb{S}(\mathbb{C})$$

induced by $\mathbb{R} \hookrightarrow \mathbb{C}$ is $z \mapsto (z, \bar{z})$. Let $U^1 = \text{Ker}(\mathbb{S} \xrightarrow{\text{Nm}} \mathbb{G}_m)$, so that $U^1(\mathbb{R}) = \{z \in \mathbb{C}^\times \mid z\bar{z} = 1\}$.

With any homomorphism $h: \mathbb{S} \to G$ of real algebraic groups there are associated homomorphisms

$$\mu_h: \mathbb{G}_m \longrightarrow G_{\mathbb{C}}, \qquad \mu_h(z) = h_{\mathbb{C}}(z, 1), \quad z \in \mathbb{G}_m(\mathbb{C}) = \mathbb{C}^\times,$$

and

$$w_h: \mathbb{G}_m \longrightarrow G, \qquad w_h(r) = h(r)^{-1}, \quad r \in \mathbb{G}_m(\mathbb{R}) = \mathbb{R}^\times \subset \mathbb{C}^\times = \mathbb{S}(\mathbb{R})$$

(the *weight homomorphism*). The following formulas are useful:

(1.1.1) $$h_{\mathbb{C}}(z_1, z_2) = \mu_h(z_1) \cdot \bar{\mu}_h(z_2); \qquad h(z) = \mu(z) \cdot \overline{\mu(z)};$$

(1.1.2) $$h(i) \equiv \mu_h(-1) \mod w_h(\mathbb{G}_m).$$

A *real Hodge structure* on an $\mathbb{R}$-vector space V can be variously defined as:

(1.2.1) a representation h of $\mathbb{S}$ on V;

(1.2.2) a (Hodge) decomposition $V \otimes \mathbb{C} = \bigoplus V^{p,q}$ such that $\overline{V^{p,q}} = V^{q,p}$ for all p, q;

(1.2.3) a (weight) gradation $V = \bigoplus V_m$ and a descending (Hodge) filtration

$$\cdots \supset F^p \supset F^{p+1} \supset \cdots$$

such that $V_m = (V_m \cap F^p) \oplus (V_m \cap \overline{F^q})$ for all m, p, q with $p + q = m + 1$.

To pass from one definition to another, use the following rules:

$$v \in V^{p,q} \Leftrightarrow h(z)v = z^{-p} \cdot \bar{z}^{-q} v, \quad \text{all } z \in \mathbb{C}^\times;$$

$$V_m \otimes \mathbb{C} = \bigoplus_{p+q=m} V^{p,q}, \qquad F^p = \bigoplus_{p' \geq p} V^{p',q'};$$

$$V^{p,q} = V_{p+q} \cap F^p \cap \overline{F^q}.$$

Note that the weight gradation is defined by w_h. From the definition (1.2.1) it is clear that the real Hodge structures form a Tannakian category $\mathbf{Hdg}_{\mathbf{R}}$ over $\mathbb{R}$ with the forgetful functor ω as a fibre functor, and $Aut^{\otimes}(\omega) = \mathbb{S}$.

A *rational Hodge structure* is a vector space V over $\mathbb{Q}$ together with a representation h of $\mathbb{S}$ on $V \otimes \mathbb{R}$ such that w_h is defined over $\mathbb{Q}$. Thus to give a rational Hodge structure on V is the same as to give a gradation $V = \bigoplus V_m$ of V together with a real Hodge structure of weight m on $V_m \otimes \mathbb{R}$ for each m. The rational Hodge structure $\mathbb{Q}(m)$ has $(2\pi i)^m \mathbb{Q}$ as its underlying vector space with $h(z)$ acting as multiplication by $(z\bar{z})^m$. There are similar definitions with $\mathbb{Q}$ replaced by a subring $R \subset \mathbb{R}$.

A *polarization* of a real Hodge structure (V, h) is a family of morphisms of Hodge structures

$$\psi_m: V_m \times V_m \longrightarrow \mathbb{R}(-m), \qquad m \in \mathbb{Z},$$

such that

$$(x, y) \mapsto (2\pi i)^m \psi_m(x, h(i)y): V_m \times V_m \longrightarrow \mathbb{R}$$

is symmetric and positive-definite for each m; equivalently, such that $(2\pi i)^m \psi_m$ *is symmetric or skew-symmetric according as* m *is even or odd*, and $(2\pi i)^m \psi_m(x, h(i)x) > 0$ for all $x \neq 0$. A *polarization* of a rational Hodge structure is a family of morphisms of rational Hodge structures $\psi_m : V_m \times V_m \to \mathbb{Q}(-m)$ such that the family $(\psi_m \otimes \mathbb{R})_m$ is a polarization of real Hodge structures. The polarizable rational Hodge structures form a Tannakian category $\mathbf{Hdg}_{\mathbb{Q}}$ over $\mathbb{Q}$ with the forgetful functor ω as fibre functor, and we let $G_{\mathrm{Hdg}} = Aut^{\otimes}(\omega)$. The tensor functor

$$\mathbf{Hdg}_{\mathbb{Q}} \to \mathbf{Hdg}_{\mathbb{R}}, \qquad V \mapsto V \otimes \mathbb{R},$$

defines a homomorphism $h_{\mathrm{Hdg}} : \mathbb{S} \to G_{\mathrm{Hdg}}$.

The conditions (SV). We list some conditions on a homomorphism $h : \mathbb{S} \to G$ of real algebraic groups:

(SV1) the Hodge structure on the Lie algebra $\mathfrak{g}$ of G defined by $\mathrm{Ad} \circ h : \mathbb{S} \to \mathrm{GL}(\mathfrak{g})$ is of type $\{(1, -1), (0, 0), (-1, 1)\}$;

(SV2) $\mathrm{ad}\, h(i)$ is a Cartan involution of G^{ad}.

When G is connected, (SV1) implies that $w_h(\mathbb{G}_m) \subset Z(G)$. In the presence of this condition, we sometimes need to consider a stronger form of (SV2):

(SV2*) $\mathrm{ad}\, h(i)$ is a Cartan involution of $G/w_h(\mathbb{G}_m)$.

Note that (SV2*) implies that G is reductive.

Let G be an algebraic group over $\mathbb{Q}$, and let h be a homomorphism $\mathbb{S} \to G_{\mathbb{R}}$. We say that (G, h) satisfies the condition (SVx) when $(G_{\mathbb{R}}, h)$ satisfies (SVx). For such a pair, we shall also need to consider the condition:

(SV3) The weight homomorphism $w_h : \mathbb{G}_m \to G_{\mathbb{R}}$ is defined over $\mathbb{Q}$ and maps into the centre of G.

Finally, when G is an affine group scheme over $\mathbb{Q}$, we say that (G, h) satisfies (SVx) if $(H, q_{\mathbb{R}} \circ h)$ satisfies (SVx) for every algebraic quotient $q : G \to H$ of G.

Abelian motives: definition. Let $\mathbf{Mot}(k)$ be the category of motives over k, defined using absolute Hodge cycles (see Deligne and Milne [**22,** §6]). We shall be concerned with the tensor subcategory $\mathbf{Mot}^{\mathrm{ab}}(k)$ of $\mathbf{Mot}(k)$ generated by the motives $h_1(A)$ for A an abelian variety over k. An object of $\mathbf{Mot}^{\mathrm{ab}}(k)$ will be called an *abelian motive* over k.

EXAMPLE 1.3. (a) The Tate motive, being isomorphic to $\Lambda^2 h_1(E)$ for any elliptic curve E, is an abelian motive.

(b) Let X be a smooth projective variety over k. Then $h(X)$ is an abelian motive if X is a curve, a unirational variety of dimension ≤ 3, a Fermat hypersurface, or a K3-surface [**22,** 6.26].

(c) Recall that the *level* of a pure Hodge structure (V, h) is the maximum value of $|q - p|$ for which $V^{p,q} \neq 0$. Let (V, h) be a polarizable

Hodge structure of level ≤ 1. If (V, h) has even weight $2m$, then it is isomorphic to a sum of copies of $\mathbb{Q}(-m)$; if it has odd weight $2m-1$, then $V \otimes \mathbb{Q}(-m)$ is of type $\{(-1, 0), (0, -1)\}$, which Riemann's theorem shows to equal $h_1(A)$ for some abelian variety. In either case, (V, h) is the Betti realization of an abelian motive.

(d) Write $V_n(a_1, \ldots, a_d)$ for the complete intersection of d smooth hypersurfaces of degrees $a_1, \ldots, a_d$ in general position in $\mathbb{P}^{n+d}$ over $\mathbb{C}$. The varieties $V_n(2)$, $V_n(2, 2)$, $V_2(3)$, $V_n(2, 2, 2)$ (n odd), $V_3(3)$, $V_3(2, 3)$, $V_5(3)$, $V_3(4)$ have rational cohomology groups with Hodge structures of level ≤ 1 (see Rapoport **[57]**), and so, if all Hodge cycles are absolutely Hodge, their motives are abelian.

By definition, the abelian motives over k form a Tannakian category over $\mathbb{Q}$, and Betti cohomology provides a fibre functor ω_B over $\mathbb{Q}$. Let $G_{\mathrm{Mab}} = Aut^{\otimes}(\omega_B)$. The functor

$$M \mapsto \omega_B(M) \otimes \mathbb{R}\colon \mathbf{Mot}^{\mathrm{ab}}(\mathbb{C}) \longrightarrow \mathbf{Hdg}_{\mathbb{R}}$$

defines a homomorphism $h_{\mathrm{Mab}}\colon \mathbb{S} \to G_{\mathrm{Mab}}$. In Corollary 1.34, we shall exhibit a universal property for the pair $(G_{\mathrm{Mab}}, h_{\mathrm{Mab}})$.

Polarizable rational Hodge structures. Let (V, h) be a polarizable rational Hodge structure, and let $G = Aut^{\otimes}(\omega)$ where ω is the forgetful functor on the tensor category generated by (V, h). Then G can be identified with a subgroup of $\mathrm{GL}(V)$, and h can be regarded as a homomorphism $\mathbb{S} \to G_{\mathbb{R}}$. We call the pair (G, h) the *Mumford-Tate group*[1] of (V, h).

LEMMA 1.4. *The following conditions on a* $t \in V^{\otimes r} \otimes V^{\vee \otimes s}$ *are equivalent*:

(a) t *is of type* $(0, 0)$;
(b) t *is fixed by* $h(\mathbb{C}^{\times})$;
(c) t *is fixed under the action of* G *on* $V^{\otimes r} \otimes V^{\vee \otimes s}$.

PROOF. The implications (a) $\Leftrightarrow$ (b) and (c) $\Rightarrow$ (b) are obvious. For (a) $\Rightarrow$ (c), note that if t is of type (0, 0), then the map $\mathbb{Q}(0) \to V^{\otimes r} \otimes V^{\vee \otimes s}$ sending 1 to t is a morphism of Hodge structures, and that, by definition, the action of G commutes with morphisms of Hodge structures. □

A tensor t satisfying the conditions in the lemma is called a *Hodge tensor* of V.

Let G be a real algebraic group, and let C be an element of $G(\mathbb{R})$ whose square is central. A C*-polarization* of a real representation (V, ξ) of G is a G-invariant bilinear form $\psi\colon V \times V \to \mathbb{R}$ such that $(x, y) \mapsto \psi(x, Cy)$ is symmetric and positive-definite.

---

[1] Sometimes the Mumford-Tate group of (V, h) is defined to be the group associated with the tensor category generated by (V, h) and $\mathbb{Q}(1)$. For the relation between the two notions, see the penultimate subsection of this section.

LEMMA 1.5. *Let G and C be as above. The following conditions are equivalent*:

(a) $\operatorname{ad} C$ *is a Cartan involution of* G;
(b) *every real representation of* G *is* C*-polarizable*;
(c) G *admits a faithful representation that is* C*-polarizable.*

PROOF. See [17, 2.8]. □

PROPOSITION 1.6. *Let G be a connected algebraic group over $\mathbb{Q}$, and let h be a homomorphism $\mathbb{S} \to G_{\mathbb{R}}$. The pair (G, h) is the Mumford-Tate group of a polarizable rational Hodge structure if and only if it satisfies the conditions* (SV2*, 3) *and G is generated by h (i.e., there is no proper $\mathbb{Q}$-rational subgroup H of G such that $\operatorname{Im}(h) \subset H_{\mathbb{R}}$).*

PROOF. Let (G, h) be the Mumford-Tate group of a polarizable rational Hodge structure (V, h). That w_h is defined over $\mathbb{Q}$ is part of the definition of a rational Hodge structure. For any $a \in \mathbb{Q}^{\times}$, $w_h(a): V \to V$ is a morphism of Hodge structures and hence commutes with the action of G. Therefore (G, h) satisfies (SV3).

Let H be a subgroup of G such that $H_{\mathbb{R}}$ contains $h(\mathbb{S})$. Then

$$t \text{ fixed by } H \Rightarrow t \text{ is a Hodge tensor} \Rightarrow t \text{ fixed by } G,$$

and it follows that $H = G$ by [19, 3.1, 3.5]. Thus h generates G.

Let $C = h(i)$. Then $C^2 = h(-1) = w_h(-1)$, which lies in the centre of $G(\mathbb{R})$. Let (W, ξ) be a representation of $G/w_h(\mathbb{G}_m)$. The rational Hodge structure $(W, \xi \circ h)$ is in the tensor category generated by (V, h), and so it is polarizable. Let $\psi: W \otimes W \to \mathbb{Q}(0)$ be a polarization of $(W, \xi \circ h)$ as a rational Hodge structure. Then ψ is fixed under the action of $h(\mathbb{S})$, and because $G/w_h(\mathbb{G}_m)$ is generated by h, it is also fixed under the action of $G/w_h(\mathbb{G}_m)$, and so it is a C-polarization. Therefore, Lemma 1.5 shows that $\operatorname{ad} C$ is a Cartan involution for $(G/w_h(\mathbb{G}_m))_{\mathbb{R}}$, i.e., that (G, h) satisfies (SV2*).

Conversely, let (G, h) be a pair satisfying (SV2*, 3) and such that G is generated by h. We first show that, for any representation $\xi: G \to \mathrm{GL}(V)$ of G on a $\mathbb{Q}$-vector space, the rational Hodge structure $(V, \xi \circ h)$ is polarizable. Again let $C = h(i)$. If $(V, \xi \circ h)$ has weight 0, then any C-polarization of (V, ξ) is also a polarization of $(V, \xi \circ h)$. If the weight is nonzero, there will be a smallest $m > 0$ such that $\mathbb{Q}(m)$ lies in the tensor category generated by (V, h), and we let G_1 be the subgroup of G that acts trivially on $\mathbb{Q}(m)$. The element C acts as 1 on $\mathbb{Q}(m)$ and so lies in $G_1(\mathbb{R})$. The map $G_1 \to G/w(\mathbb{G}_m)$ is an isogeny, and so the condition (SV2*) for (G, h) implies that $\operatorname{ad} C$ is a Cartan involution of $G_1(\mathbb{R})$. Therefore, there is a G_1-invariant C-polarization ψ of V. After replacing V with a homogeneous component, we may suppose that $(V, \xi \circ h)$ has weight n, and then the map

$$(2\pi i)^{-n}\psi: V \otimes V \longrightarrow \mathbb{Q}(-n)$$

is a polarization of $(V, \xi \circ h)$.

Now choose ξ to be a faithful representation of G, and let G' be the Mumford-Tate group of the polarizable rational Hodge structure $(V, \xi \circ h)$. Both G and G' are algebraic subgroups of $\mathrm{GL}(V)$ generated by h, and so they must be equal. □

COROLLARY 1.7. *The pair* $(G_{\mathrm{Hdg}}, h_{\mathrm{Hdg}})$ *satisfies the conditions* (SV2*, 3); *moreover, for any algebraic group* G *and map* h *satisfying these conditions, there is a unique homomorphism* $\rho(h)\colon G_{\mathrm{Hdg}} \to G$ *such that* $h = \rho(h)_{\mathbb{R}} \circ h_{\mathrm{Hdg}}$.

PROOF. The algebraic quotients of $(G_{\mathrm{Hdg}}, h_{\mathrm{Hdg}})$ are precisely the Mumford-Tate groups of polarizable rational Hodge structures, and so the proposition shows that $(G_{\mathrm{Hdg}}, h_{\mathrm{Hdg}})$ satisfies (SV2*, 3) and is generated by h_{Hdg}. Let (G, h) be a pair satisfying (SV2*, 3), and let G' be the subgroup of G generated by h. Then (G', h) is the Mumford-Tate group of a polarizable Hodge structure, and so there is a homomorphism $\rho(h)\colon G_{\mathrm{Hdg}} \to G' \subset G$ such that $h = \rho(h)_{\mathbb{R}} \circ h_{\mathrm{Hdg}}$. It is unique because h generates G_{Hdg}. □

Note that the conditions in the corollary determine the pair $(G_{\mathrm{Hdg}}, h_{\mathrm{Hdg}})$ uniquely (up to a unique isomorphism).

The Mumford-Tate group of a polarizable Hodge structure is reductive (because it satisfies (SV2*)) and connected (because it is generated by h, and $\mathbb{S}$ is connected). Consequently G_{Hdg} is pro-reductive and connected.

The functor

$$\omega_B\colon \mathbf{Mot}^{\mathrm{ab}}(\mathbb{C}) \longrightarrow \mathbf{Hdg}_{\mathbb{Q}}$$

induces a homomorphism

$$\rho\colon G_{\mathrm{Hdg}} \longrightarrow G_{\mathrm{Mab}}$$

such that $\rho_{\mathbb{R}} \circ h_{\mathrm{Hdg}} = h_{\mathrm{Mab}}$, and ρ is the unique homomorphism satisfying this condition. Because ω_B is fully faithful, ρ is surjective (i.e., faithfully flat).

Hodge structures of CM-type. A polarizable rational Hodge structure is said to be of *CM-type* if its Mumford-Tate group is commutative and, hence, a torus. The Hodge structures of CM-type form a Tannakian subcategory $\mathbf{Hdg}_{\mathbb{Q}}^{\mathrm{cm}}$ of $\mathbf{Hdg}_{\mathbb{Q}}$ over $\mathbb{Q}$.

PROPOSITION 1.8. *Every Hodge structure of CM-type is the Betti realization of an abelian motive.*

PROOF. This is well known. For a proof, see **[51**, 4.6]. □

COROLLARY 1.9. *The kernel of* $\rho\colon G_{\mathrm{Hdg}} \to G_{\mathrm{Mab}}$ *is contained in* $(G_{\mathrm{Hdg}})^{\mathrm{der}}$.

PROOF. Let $S = \mathit{Aut}^{\otimes}(\omega)$ where ω is the forgetful functor on $\mathbf{Hdg}_{\mathbb{Q}}^{\mathrm{cm}}$. The proposition shows that the inclusion $\mathbf{Hdg}_{\mathbb{Q}}^{\mathrm{cm}} \hookrightarrow \mathbf{Hdg}_{\mathbb{Q}}$ factors through $\mathbf{Mot}^{\mathrm{ab}}(\mathbb{C}) \hookrightarrow \mathbf{Hdg}_{\mathbb{Q}}$, and so the homomorphism $G_{\mathrm{Hdg}} \to S$ factors through

$G_{\mathrm{Hdg}} \to G_{\mathrm{Mab}}$. Thus,

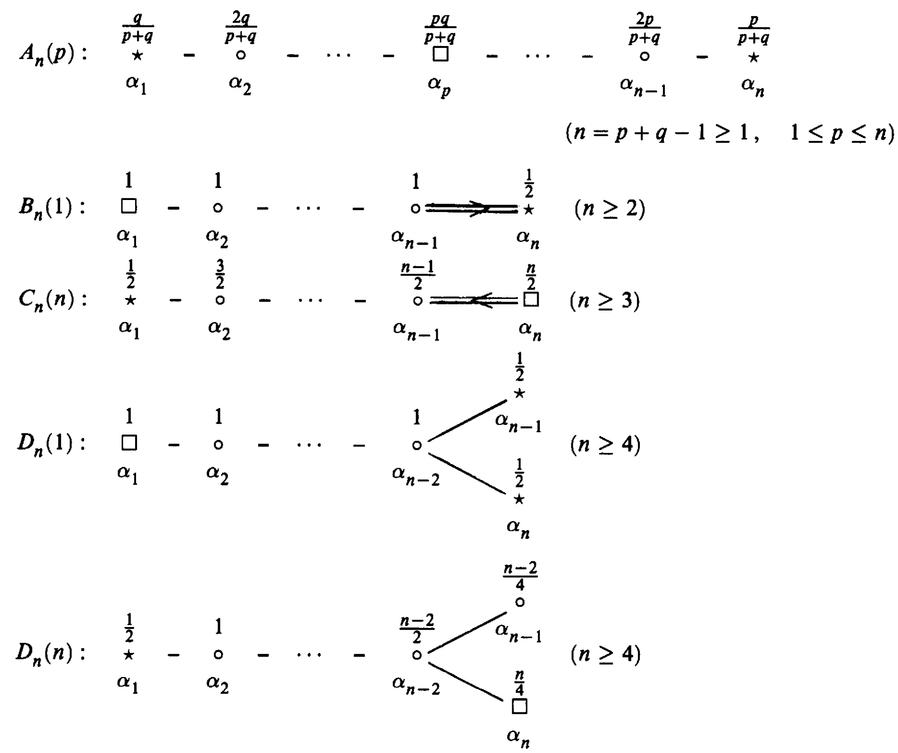
$$\mathbf{Hdg}_{\mathbb{Q}}^{\mathrm{cm}} \hookrightarrow \mathbf{Mot}^{\mathrm{ab}}(\mathbb{C}) \hookrightarrow \mathbf{Hdg}_{\mathbb{Q}},$$
$$S \twoheadleftarrow G_{\mathrm{Mab}} \twoheadleftarrow G_{\mathrm{Hdg}}.$$

Hence $\mathrm{Ker}(\rho) \subset \mathrm{Ker}(G_{\mathrm{Hdg}} \twoheadrightarrow S) = (G_{\mathrm{Hdg}})^{\mathrm{der}}$. □

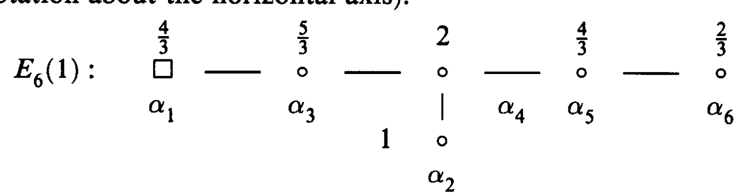
The pro-torus S in the proof is called the *Serre group.* For a description of it in terms of its character group, see for example [**51,** §4].

Dynkin diagrams. For future reference, we provide a table of Dynkin diagrams.

TABLE 1.10

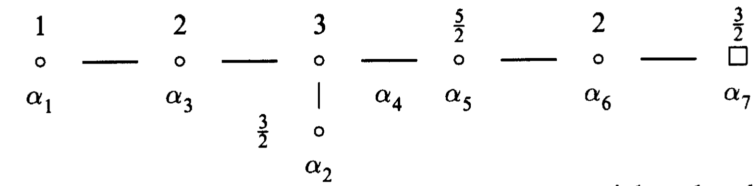


$D_n(n-1)$: Same as $D_n(n-1)$ but with α_{n-1} and α_n interchanged (rotation about the horizontal axis).



$E_6(6)$: Same as $E_6(1)$ but with (α_1, α_3) interchanged with (α_6, α_5) (rotation about the vertical axis).

$E_7(7)$:



As will be explained below, nodes marked by squares are special, and nodes marked by stars correspond to symplectic representations. The number in parenthesis indicates the position of the special node. Let α be the simple root corresponding to the special node; the number attached to the i^{th} node is the coefficient of α in the expansion of the i^{th} fundamental weight ϖ_i.

Special Hodge structures. A *special* Hodge structure is a polarizable rational Hodge structure whose Mumford-Tate group (G, h) satisfies (SV1). The next two results were noted in [**17**, 7.3].

PROPOSITION 1.11. *The special Hodge structures form a Tannakian subcategory of* $\mathbf{Hdg}_{\mathbb{Q}}$.

PROOF. A direct sum of special Hodge structures is obviously special. Let (V, h) be a special Hodge structure, and let (G, h) be its Mumford-Tate group. The Mumford-Tate group of any Hodge structure in the tensor category generated by (V, h) is a quotient of (G, h) and, hence, satisfies (SV1).

PROPOSITION 1.12. *The Betti realization of an abelian motive is special.*

PROOF. After Proposition 1.11, it suffices to prove this for an abelian variety. The Betti realization (V, h) of an abelian variety is of type $\{(-1, 0), (0, -1)\}$; let (G, h) be its Mumford-Tate group. Then $\mathfrak{g} \subset \operatorname{End}(V) = V \otimes V^\vee$, which is of type $\{(-1, 1), (0, 0), (1, -1)\}$. □

It is reasonable to hope that the following statement may be true.

HYPOTHESIS 1.13. *Every special Hodge structure is the Betti realization of a motive.*

More explicitly, this means the following: for every special Hodge structure (V, h), there is a projective algebraic variety X and an integer m such that (V, h) is a direct factor of $H_B(X)(m)$ and the projection $H_B(X)(m) \to V \subset H_B(X)(m)$ is an absolute Hodge cycle on X (e.g., the class of an algebraic cycle).

A motive whose Betti realization is special will be called a *special motive.* The next example indicates that they are indeed exceptional among motives. Apparently, no special motive is known that is not already abelian, although the hypothesis predicts their existence. In §3 we shall see that Deligne's hope that all Shimura varieties with rational weight are moduli varieties for motives implies the hypothesis.

EXAMPLE 1.14. Let $X \to \mathbb{P}^1$ be a Lefschetz pencil of hypersurfaces of degree d and odd dimension $2r-1$ over $\mathbb{C}$, and let X_s be the fibre over $s \in \mathbb{P}^1(\mathbb{C})$. It is known (see Deligne [**17**, 7.6] that for s outside a countable subset of $\mathbb{P}^1(\mathbb{C})$, the Mumford-Tate group of the rational Hodge structure

$H^{2r-1}(X_s, \mathbb{Q})$ is the full group of symplectic similitudes. It follows that the Hodge structure is not special unless it has level ≤ 1.

PROPOSITION 1.15. *A pair* (G, h) *is the Mumford-Tate group of a special Hodge structure if and only if it satisfies* (SV1, 2*, 3) *and* h *generates* G.

PROOF. Immediate consequence of Proposition 1.6 and the definition of a special Hodge structure. □

Classification. Following [18], we classify the pairs (G, h) as in Proposition 1.15 with G a simple adjoint group. Note that for an adjoint group, (SV3) simply says that h is a homomorphism $\mathbb{S}/\mathbb{G}_m \to G_\mathbb{R}$ and that (SV2) implies (SV2*).

LEMMA 1.16. *A simple adjoint group* G *over* $\mathbb{Q}$ *for which there exists a homomorphism* $h\colon \mathbb{S}/\mathbb{G}_m \to G_\mathbb{R}$ *satisfying* (SV2) *is of the form* $\operatorname{Res}_{F/\mathbb{Q}} G_0$ *for some absolutely simple group* G_0 *over a totally real number field* F.

PROOF. Every simple adjoint group over $\mathbb{Q}$ is of the form $\operatorname{Res}_{F/\mathbb{Q}} G_0$ for some absolutely simple group G_0 over a number field F, and so the only problem is to show that F is totally real. A compact simple group over $\mathbb{R}$ is absolutely simple, and an inner form of an absolutely simple group is also absolutely simple. The condition (SV2) implies that $G_\mathbb{R}$ is an inner form of its compact form, and hence its simple factors are absolutely simple. Since

$$G_\mathbb{R} = \prod_{v|\infty} \operatorname{Res}_{F_v/\mathbb{R}} G_{0,F_v},$$

this shows that F must be totally real. □

Thus to give a pair (G, h) as in Proposition 1.15 with G simple and adjoint is the same as to give an absolutely simple adjoint group G_0 over a totally real field F together with homomorphisms

$$h_v\colon \mathbb{S}/\mathbb{G}_m \to G_v, \quad G_v \underset{\text{df}}{=} G_0 \otimes_F F_v, \qquad v \text{ a real prime of } F,$$

satisfying (SV1,2) and such that at least one h_v is nontrivial.

We fix a simple adjoint group G_∞ over $\mathbb{C}$ and consider the triples (G, γ, h) consisting of a real inner form[2] (G, γ) of the compact form G_c of G_∞ and a nontrivial homomorphism $h\colon \mathbb{S}/\mathbb{G}_m \to G$ satisfying (SV1, 2).

[2]Let G_0 be an algebraic group over a field k_0 of characteristic zero and let k be an algebraic closure of k_0. An *inner form* of G_0 is an algebraic group G over k_0 together with a $G_0(k)$-conjugacy class γ of isomorphisms $c\colon G_{0,k} \to G_k$ such that $c^{-1} \circ \tau c$ is an inner automorphism of $G_{0,k}$ for all $\tau \in \operatorname{Gal}(k/k_0)$. Two inner forms (G, γ) and (G', γ') are *isomorphic* if there is an isomorphism of algebraic groups $\varphi\colon G \to G'$ (over k_0) such that

$$c \in \gamma \implies \varphi \circ c \in \gamma'.$$

Such a φ is uniquely determined up to an inner automorphism of G over k_0. If (G, γ) is an inner form of G_0 and $c \in \gamma$, then $c_\tau = c^{-1} \circ \tau c$ is a 1-cocycle for G_0^{ad} whose cohomology class does not depend on the choice of c. In this way, the set of isomorphism classes of inner forms of G_0 becomes identified with $H^1(k_0, G_0^{\text{ad}})$.

Choose a maximal torus T in G_∞. Let $R \subset X^*(T)$ be the corresponding system of roots, and fix a system of simple roots B. The nodes of the Dynkin diagram D are parametrized by the elements of B. Recall Bourbaki [**11**, VI.1.8] that there is a unique root $\tilde{\alpha} = \sum_{\alpha\in B} n(\alpha)\alpha$ such that, for any root $\sum_{\alpha\in B} m(\alpha)\alpha$, $n(\alpha) \geq m(\alpha)$ for all $\alpha \in B$. We call a node s_α of D *special* if $n(\alpha) = 1$.

From (G, γ, h) we obtain a $G_\infty(\mathbb{C})$-conjugacy class of cocharacters $\gamma \circ \mu_h$ of G_∞. This class contains a unique element $\mu \in X_*(T)$ such that

$$\langle \alpha, \mu \rangle \geq 0 \quad \text{for all } \alpha \in B.$$

The condition (SV1) implies that

$$\langle \alpha, \mu \rangle \in \{1, 0, -1\} \text{ for } \alpha \in R.$$

Since μ is nontrivial, not all the values $\langle \alpha, \mu \rangle$ can be zero; so these conditions imply that $\langle \alpha, \mu \rangle = 1$ for exactly one $\alpha \in B$, which must in fact be special (otherwise $\langle \tilde{\alpha}, \mu \rangle > 1$); moreover, this condition is also sufficient for (SV1) to hold.

PROPOSITION 1.17. *The map* $(G, \gamma, h) \mapsto s_\alpha$ *defines a one-to-one correspondence between the set of isomorphism classes of triples* (*as above*) *and the set of special nodes of* D. *The isomorphism class of the pair* (G, γ) *itself determines* s_α *unless the opposition involution* τ *moves* s_α, *in which case*

$$(G, \gamma, h) \leftrightarrow s_\alpha, \qquad (G, \gamma, h^{-1}) \leftrightarrow \tau s_\alpha.$$

PROOF. We explain only how to construct the triple (G, γ, h) corresponding to a special node s_α—see [**18**, 1.2] for more details. There is a unique $\mu \in X_*(T)$ such that

$$\text{(1.17.1)} \qquad \langle \alpha, \mu \rangle = 1, \quad \langle \alpha', \mu \rangle = 0 \qquad \text{for all } \alpha' \in B, \quad \alpha' \neq \alpha.$$

Let (G, γ) be the real form of the compact form G_c corresponding to the Cartan involution $\mathrm{ad}\,\mu(-1)$, i.e., such that

$$c(\mu(-1) \cdot \iota g \cdot \mu(-1)) = \iota c(g), \qquad g \in G_c(\mathbb{C}), \text{ some } c \in \gamma,$$

and set $h(z_1, z_2) = \mu(z_1) \cdot (\iota\mu)(z_2)$ (cf. (1.1.1), (1.1.2)). Then h, when regarded as a map into $G_{\mathbb{C}}$, is defined over $\mathbb{R}$, and because s_α is special, (G, h) satisfies (SV1).

The final statement follows from [**18**, 1.2.7]. □

An examination of the tables in [**11**, pp. 250–275] reveals that every node of the Dynkin diagram of type A_n is special, that the Dynkin diagrams of type B_n, C_n, and E_7 each have one special node, that the Dynkin diagrams of type D_n each have three special nodes, and that the Dynkin diagram of type E_6 has two special nodes. This is illustrated in Table 1.10, where the special nodes are marked by squares. The Dynkin diagrams of type E_8, F_4, and G_2 have no special nodes, and so groups of these types can not occur as factors of the adjoint group of the Mumford-Tate group of a special Hodge structure.

Following Deligne, we write $D_n^{\mathbb{R}}$ for the diagrams $D_n(1)$, $D_4(3)$, and $D_4(4)$, and we write $D_n^{\mathbb{H}}$ for the remaining diagrams of type D_n. A simple adjoint group G over $\mathbb{R}$ will be said to be of type A_n, B_n, C_n, $D_n^{\mathbb{R}}$, $D_n^{\mathbb{H}}$, E_6, or E_7 if it corresponds to a diagram of that type in Table 1.10.

Let G be a simple group over $\mathbb{Q}$ such that $G_{\mathbb{R}}$ is noncompact. We say G is of type A_n, B_n, C_n, $D_n^{\mathbb{R}}$, $D_n^{\mathbb{H}}$, E_6, or E_7 if all the noncompact factors of $G_{\mathbb{R}}^{\mathrm{ad}}$ are of this type. When noncompact factors of type $D_n^{\mathbb{R}}$ and $D_n^{\mathbb{H}}$ both occur, we say G is of mixed type D.

Symplectic representations. In this subsection, we review the symplectic representations of groups. These were studied by Satake in a series of papers (see especially Satake **[61]–[63]**). Our exposition follows that of **[18]**.

A *symplectic space* (V, ψ) over a field k is a finite-dimensional vector space V over k together with a nondegenerate alternating form ψ on V. The corresponding *symplectic group* is the subgroup of $\mathrm{GL}(V)$ such that

$$\mathrm{Sp}(\psi)(k) = \{g \in \mathrm{GL}(V) \mid \psi(gx, gy) = \psi(x, y),\ \text{all } x, y \in V\},$$

and the group of *symplectic similitudes*, $G(\psi)$, is $\mathrm{Sp}(\psi)\cdot\mathbb{G}_m$ (here $\mathbb{G}_m$ is identified with the group of nonzero diagonal matrices). Write $\mathrm{PSp}(\psi)$ for the adjoint group of $\mathrm{Sp}(\psi)$. The *Siegel upper half-space* $X(\psi)^+$ corresponding to a real symplectic space (V, ψ) is the set of Hodge structures h on V of type $\{(-1, 0), (0, -1)\}$ for which $2\pi i\psi$ is a polarization. Each $h \in X(\psi)^+$ factors through $G(\psi)$, and the map $h \mapsto \bar{h} = \mathrm{ad}\circ h$ identifies $X(\psi)^+$ with an $\mathrm{Sp}(\psi)(\mathbb{R})$-conjugacy class of maps $\mathbb{S}/\mathbb{G}_m \to \mathrm{PSp}(\psi)$.

The real case. Let H be a semisimple group over $\mathbb{R}$, and let $\bar{h}$ be a homomorphism $\mathbb{S}/\mathbb{G}_m \to H^{\mathrm{ad}}$, none of whose components are trivial, satisfying (SV1,2).

We shall say that a representation $\xi\colon H \to \mathrm{GL}(V)$ with finite kernel is *symplectic* if there exists a nondegenerate alternating form ψ on V, a reductive group G, and a homomorphism $h\colon \mathbb{S} \to G$ such that

(1.18.1) $G^{\mathrm{der}} = \xi(H)$ (hence $G^{\mathrm{ad}} = H^{\mathrm{ad}}$), and $\mathrm{ad}\circ h = \bar{h}$;

(1.18.2) ξ extends to G in such a way that $\xi(G) \subset G(\psi)$ and $\xi\circ h \in X(\psi)^+$.

Assume H is simply connected, and set $\mu = \mu_{\bar{h}}$. Choose a maximal torus T in $H_{\mathbb{C}}$, and let $R \subset X^*(T)$ be the corresponding system of roots. Let $B = \{\alpha_1, \dots, \alpha_n\}$ be a system of simple roots such that $\langle\alpha, \mu\rangle \geq 0$ for all $\alpha \in B$. Recall that the lattice of weights is

$$P(R) = \{\varpi \in X^*(T)\otimes\mathbb{Q} \mid \langle\varpi, \alpha^\vee\rangle \in \mathbb{Z} \text{ all } \alpha^\vee \in R^\vee\},$$

that the fundamental weights are the elements of the dual basis $\{\varpi_1, \dots, \varpi_n\}$ to $\{\alpha_1^\vee, \dots, \alpha_n^\vee\}$, and that the dominant weights are the elements $\sum n_i\varpi_i$, $n_i \in \mathbb{N}$. The quotient $P(R)/Q(R)$ of $P(R)$ by the lattice $Q(R)$ generated by R is the character group of $Z(H)$. Write τ for the opposition involution acting on the Dynkin diagram (or on R or on the set of fundamental weights): it preserves each connected component of the diagram and acts as the unique nontrivial involution on a component of type A_n, D_n (n odd), or E_6, and trivially on the other components.

PROPOSITION 1.19. *Let ξ be an irreducible representation of H on a real vector space, and let ϖ be the highest weight of an irreducible component of $\xi_{\mathbb{C}}$. The representation ξ is symplectic if and only if*

$$(1.19.1) \qquad \langle \varpi + \tau\varpi , \bar{\mu} \rangle = 1.$$

PROOF. See [**18**, 1.3.6]. □

COROLLARY 1.20. *If ξ is symplectic, then ϖ is a fundamental weight. Therefore, the representation factors through a simple factor of H.*

PROOF. For any dominant weight ϖ, $\langle \varpi + \tau\varpi , \mu \rangle \in \mathbb{N}$ because $\varpi + \tau\varpi \in Q(R)$. If $\varpi \neq 0$, $\langle \varpi + \tau\varpi , \mu \rangle > 0$ unless μ kills all the weights of the representation corresponding to ϖ. Hence a dominant weight satisfying (1.19.1) can not be a sum of two dominant weights. □

PROPOSITION 1.21. *Let H be a simply connected simple group over $\mathbb{R}$, and let $\bar{h}\colon \mathbb{S}/\mathbb{G}_m \to H^{\mathrm{ad}}$ be a nontrivial homomorphism satisfying* (SV1, 2). *There exists a nontrivial symplectic representation of $(H, \bar{h})$ if and only if H is of type A, B, C, or D. If H is of type A, B, C, or $D^{\mathbb{R}}$, then the symplectic representations form a faithful family of representations of H; if H is of type $D^{\mathbb{H}}$ they form a faithful family of representations of the double covering of the adjoint group corresponding to the subgroup of $P(R)/Q(R)$ generated by ϖ_1.*

PROOF. The proof proceeds by an examination of the tables. We treat only the cases B, D, and E_6, since the remaining cases are similar. In each case, $(H, \bar{h})$ corresponds to a special simple root α of $H_{\mathbb{C}}$, and (see 1.17.1) $\langle \varpi_i , \mu \rangle$ is the coefficient of α in the expression of ϖ_i as a $\mathbb{Q}$-linear combination of the simple roots. These coefficients are listed in Table 1.10.

(B_n). In this case $\mu \leftrightarrow \alpha_1$, and the opposition involution acts trivially on the Dynkin diagram. Thus we seek a fundamental weight ϖ_i such that $\varpi_i = \frac{1}{2}\alpha_1 + \cdots$. From Table 1.10 we see that only ϖ_n has this property. Because ϖ_n generates P/Q, the representation with highest weight ϖ_n is a faithful representation of H.

$(D_n^{\mathbb{R}})$. Suppose first that $n = 4$ and $\mu \leftrightarrow \alpha_4$. The opposition involution acts trivially, and so ϖ_1 and ϖ_3 give rise to symplectic representations; they generate P/Q. The case $\mu \leftrightarrow \alpha_3$ is similar. Otherwise $\mu \leftrightarrow \alpha_1$. The opposition involution acts trivially if n is even and switches α_{n-1} and α_n if n is odd. In the first case, ϖ_n and ϖ_{n-1} give rise to symplectic representations and together they generate P/Q. In the second case, ϖ_n and ϖ_{n-1} give rise to symplectic representations and each generates P/Q.

$(D_n^{\mathbb{H}})$. In this case, $\mu \leftrightarrow \alpha_n$ (or α_{n-1}). Only ϖ_1 gives rise to a symplectic representation, and it generates a subgroup of order 2 (and index 2) in $P(R)/Q(R)$. The corresponding representation factors through H/C where C is the kernel of ϖ_1 regarded as a character of $Z(H)$.

(E_6). In this case, $\mu \leftrightarrow \alpha_1$ or α_6. An examination of Table 1.10 shows that no fundamental weight qualifies. □

REMARK 1.22. Let H be the identity component of the group of automorphisms of a nondegenerate anti-Hermitian form on a vector space of dimension n over a quaternion algebra $\mathbb{H}$ over $\mathbb{R}$. Then H is an inner form of $\mathrm{SO}(2n)$, and it is of type $D_n^{\mathbb{H}}$. It is the double covering of G^{ad} corresponding to the subgroup of $P(R)/Q(R)$ generated by ϖ_1.

The rational case. Now let H be a semisimple group over $\mathbb{Q}$, and let $\bar{h}\colon \mathbb{S}/\mathbb{G}_m \to H_{\mathbb{R}}^{\mathrm{ad}}$ be a homomorphism satisfying (SV1, 2) and generating H^{ad}. We choose the maximal torus T in $H_{\mathbb{C}}$ to be rational over $\mathbb{Q}^{\mathrm{al}}$.

We shall say that a representation $\xi\colon H \to \mathrm{GL}(V)$ (over $\mathbb{Q}$) with finite kernel is *symplectic* if there exists a nondegenerate alternating form ψ on V, a reductive group G, and an $h\colon \mathbb{S} \to G_{\mathbb{R}}$ such that

(1.23.1) $G^{\mathrm{der}} = \xi(H)$ (hence $G^{\mathrm{ad}} = H^{\mathrm{ad}}$), and $\mathrm{ad} \circ h = \bar{h}$;

(1.23.2) ξ extends to a faithful representation of G on V in such a way that $\xi(G) \subset G(\psi)$ and $\xi_{\mathbb{R}} \circ h \in X(\psi)^+$.

LEMMA 1.24. *Let H be a simply connected and simple group over $\mathbb{Q}$, and let $\bar{h}\colon \mathbb{S}/\mathbb{G}_m \to H_{\mathbb{R}}^{\mathrm{ad}}$ be a nontrivial homomorphism satisfying* (SV1, 2). *If $(H, \bar{h})$ has a symplectic representation over $\mathbb{Q}$, then H cannot be of exceptional type or of mixed type D.*

PROOF. Because $(H, \bar{h})$ satisfies (SV2), $H = \mathrm{Res}_{F/\mathbb{Q}} H_0$ for some absolutely simple group H_0 over a totally real field F (see Lemma 1.16). Thus

$$H_{\mathbb{R}} = \prod_{v \in I} H_v, \quad H_v = H_0 \otimes_{F,v} \mathbb{R}, \quad I = \mathrm{Hom}(F, \mathbb{R}).$$

Let I_{nc} be the subset of I of v for which H_v is not compact, and let $H_{\mathrm{nc}} = \prod_{v \in I_{\mathrm{nc}}} H_v$. Because $\bar{h}$ generates H^{ad}, I_{nc} is nonempty. The Galois group $\mathrm{Gal}(\mathbb{Q}^{\mathrm{al}}/\mathbb{Q})$ acts on the Dynkin diagram of $H_{\mathbb{C}}$ in a manner consistent with its projection to I.

Let (V, ξ) be a symplectic representation of $(H, \bar{h})$. The restriction of $\xi_{\mathbb{R}}$ to H_{nc} is a real symplectic representation of H_{nc}, and so (Corollary 1.20 and Proposition 1.21) any nontrivial irreducible component of $\xi_{\mathbb{C}}|H_{\mathrm{nc}}$ factors through H_v for some $v \in I_{\mathrm{nc}}$ and corresponds to a node of the Dynkin diagram D_v of H_v marked in Table 1.10 with a star.

An irreducible component W of $\xi_{\mathbb{C}}$ is of the form $\bigotimes_{v \in T} W_v$ where T is a subset of I and W_v is the representation of $G_{v\mathbb{C}}$ corresponding to a node $s_v \in D_v$. Let $\mathcal{S}(\xi)$ be the set of all nodes that arise in this fashion from an irreducible component of $\xi_{\mathbb{C}}$. Then $\mathcal{S}(\xi)$ is nonempty, stable under the action of $\mathrm{Gal}(\mathbb{Q}^{\mathrm{al}}/\mathbb{Q})$, and, if $s \in \mathscr{S}(\xi) \cap D_v$, $v \in I_{\mathrm{nc}}$, then s is marked by a star in Table 1.10. Since the diagrams for E_6 and E_7 have no starred nodes, no such set exists in this case. If G is of mixed type D, then there is no such set $\mathscr{S}(\xi)$ because there is no automorphism of a Dynkin diagram of type D_n, $n \geq 5$, carrying the node s_1 into either s_{n-1} or s_n. □

We shall need to consider the following condition on a semisimple group H over $\mathbb{Q}$:

(1.25) there exists an isogeny $H' \to H$ with H' a product of simple groups H_i' such that either

(a) H_i' is simply connected of type A, B, C, or $D^{\mathbb{R}}$, or,

(b) H_i' is of type $D_n^{\mathbb{H}}$ $(n \geq 5)$ and equals $\mathrm{Res}_{F/\mathbb{Q}} H_0$ for H_0 the double covering of an adjoint group that is a form of $\mathrm{SO}(2n)$ (cf. Remark 1.22).

THEOREM 1.26. *Let H be a semisimple group over $\mathbb{Q}$, and let $\bar{h}$ be a homomorphism $\mathbb{S}/\mathbb{G}_m \to H_{\mathbb{R}}^{\mathrm{ad}}$ satisfying* (SV1, 2) *and generating H^{ad}. There exists an isogeny $H' \to H$ such that $(H', \bar{h})$ admits a faithful family of symplectic representations if and only if H satisfies* (1.25).

PROOF. Suppose that H satisfies (1.25), and let $H' \to H$ be an isogeny as in the statement of (1.25). It suffices to show that H' admits a faithful family of symplectic representations, and for this it suffices to show that each simple factor of H' admits such a family. This is proved in [**18**, 2.3.10].

Conversely, suppose H has a covering H' such that $(H', \bar{h})$ admits a faithful family of symplectic representations. According to Lemma 1.24, H' (hence H) cannot be of exceptional type or mixed type D. Let H^{sc} be the universal covering group of H' (hence of H), and let H'' be the quotient of H^{sc} by the intersection of the kernels of the rational symplectic representations of H^{sc}. Then H'' is still a covering of H', and it follows from (Corollary 1.20 and Proposition 1.21) that it satisfies (1.25). □

Abelian motives: Mumford-Tate groups. As we noted above, Deligne's theorem [**19**] shows that $\omega_B\colon \mathbf{Mot}^{\mathrm{ab}}(\mathbb{C}) \to \mathbf{Hdg}_{\mathbb{Q}}$ is fully faithful, and so the homomorphism $G_{\mathrm{Hdg}} \to G_{\mathrm{Mab}}$ it defines is surjective. When we identify rational Hodge structures with representations of G_{Hdg} on $\mathbb{Q}$-vector spaces, abelian motives become identified with those representations that factor through G_{Mab} (cf. [**21**, 8.17]).

THEOREM 1.27. *Let G be an algebraic group over $\mathbb{Q}$, and let $h\colon \mathbb{S} \to G_{\mathbb{R}}$ be a homomorphism satisfying* (SV1, 2*, 3) *and generating G. The pair (G, h) is the Mumford-Tate group of an abelian motive if and only if G^{der} satisfies* (1.25).

The proof will occupy the rest of this subsection.

PROPOSITION 1.28. *For any semisimple group H over $\mathbb{Q}$ and homomorphism $\bar{h}\colon \mathbb{S}/\mathbb{G}_m \to H_{\mathbb{R}}^{\mathrm{ad}}$ satisfying* (SV1, 2), *there exists a reductive group G*

with $G^{\mathrm{der}} = H$ *and a homomorphism* $h\colon \mathbb{S} \to G_{\mathbb{R}}$ *lifting* $\bar{h}$ *and satisfying* (SV1, 2*, 3).

PROOF. (Cf. Milne [**48,** Appendix]). For any finite extension L of $\mathbb{Q}$ splitting H, there exists a central extension defined over $\mathbb{Q}$

$$1 \longrightarrow N \longrightarrow G \longrightarrow H^{\mathrm{ad}} \longrightarrow 1$$

such that $G^{\mathrm{der}} = H$ and N is a product of copies of $(\mathbb{G}_m)_{L/\mathbb{Q}}$ (the $\mathbb{Q}$-torus obtained from $\mathbb{G}_{m,L}$ by restriction of scalars). For a proof, see for example Milne and Shih [**53,** 3.1].

Assume first that $\bar{h}$ is "special", i.e., that it factors through $T_{\mathbb{R}}$ for some maximal torus T in H^{ad}. Then (SV2) implies that $T_{\mathbb{R}}$ is anisotropic, and so T splits over a CM-field L, which we may choose to be Galois over $\mathbb{Q}$. Construct G as above using this L. According to Borel [**6,** 12.4, 13.17], there is a maximal torus $T' \subset G$ mapping onto T. Since T' is its own centralizer, it contains N, which is therefore the kernel of $T' \to T$. Hence $X_*(T') \to X_*(T)$ is surjective, and we can choose $\mu \in X_*(T')$ mapping to $\mu_{\bar{h}} \in X_*(T)$. The weight $w \underset{\mathrm{df}}{=} -\mu - \iota\mu$ of μ lies in $X_*(N)$. Because $X_*(N)$ is an induced Galois module, its cohomology groups are zero; in particular, the zeroth Tate (modified) group

$$H^0_{\mathrm{Tate}}(\mathrm{Gal}(\mathbb{C}/\mathbb{R}), X_*(N)) \underset{\mathrm{df}}{=} \frac{X_*(N)^{\mathrm{Gal}(\mathbb{C}/\mathbb{R})}}{(\iota+1)X_*(N)} = 0.$$

Clearly $\iota w = w$, and so there exists a $\mu_0 \in X_*(N)$ such that $(\iota+1)\mu_0 = w$. When we replace μ with $\mu + \mu_0$, we find that the weight becomes 0; in particular, it is defined over $\mathbb{Q}$. Choose h so that $h(z) = \mu(z) \cdot \overline{\mu(z)}$.

For a general $\bar{h}$, there will exist a $\bar{g} \in H^{\mathrm{ad}}(\mathbb{R})$ such that $\operatorname{ad}\bar{g} \circ \bar{h}$ is special (Deligne [**19,** p. 75]). Construct G and h as in the last paragraph corresponding to $\operatorname{ad}\bar{g} \circ \bar{h}$. Because $H^1(\mathbb{R}, N) = H^1(L \otimes_{\mathbb{Q}} \mathbb{R}, \mathbb{G}_m) = 0$, $\bar{g}$ will lift to an element $g \in G(\mathbb{R})$, and we take the pair $(G, \operatorname{ad} g^{-1} \circ h)$.

The pair (G, h) we have constructed satisfies (SV1, 2, 3), and its centre is split by a CM-field. Let T be the subtorus of G^{ab} generated by h. Then $T_{\mathbb{R}}$ is anisotropic, and when we replace G with the inverse image of T, we obtain a pair (G, h) satisfying (SV1, 2*, 3). □

COROLLARY 1.29. *Let* H *be a semisimple group over* $\mathbb{Q}$, *and let* $\bar{h}\colon \mathbb{S}/\mathbb{G}_m \to H^{\mathrm{ad}}_{\mathbb{R}}$ *be a homomorphism satisfying* (SV1, 2). *There exists a unique homomorphism* $\rho(H, \bar{h})\colon (G_{\mathrm{Hdg}})^{\mathrm{der}} \to H$ *such that the following diagram commutes*:

$$\begin{array}{ccc} (G_{\mathrm{Hdg}})^{\mathrm{der}} & \xrightarrow{\rho(H,\bar{h})} & H \\ \downarrow{\scriptstyle \mathrm{inj}} & & \downarrow{\scriptstyle \mathrm{surj}} \\ G_{\mathrm{Hdg}} & \xrightarrow{\rho(\bar{h})} & H^{\mathrm{ad}}. \end{array}$$

Here $\rho(\bar{h})$ *is the unique homomorphism such that* $\bar{h} = \rho(\bar{h})_{\mathbb{R}} \circ h_{\mathrm{Hdg}}$ (see Corollary 1.7).

PROOF. Two such homomorphisms $\rho(H, \bar{h})$ would differ by a map into $Z(H)$. Because $(G_{\mathrm{Hdg}})^{\mathrm{der}}$ is connected, any such map is constant, and so the homomorphisms will be equal.

For the existence, choose a pair (G, h) as in Proposition 1.28, and take $\rho(H, \bar{h}) = \rho(h)|(G_{\mathrm{Hdg}})^{\mathrm{der}}$. □

REMARK 1.30. Let G be a reductive group over $\mathbb{Q}$, and let h be a homomorphism $\mathbb{S} \to G_{\mathbb{R}}$ satisfying (SV1, 2*, 3). Let $H = G^{\mathrm{der}}$ and let $\bar{h} = \mathrm{ad} \circ h$. The restriction of $\rho(h)$ to $(G_{\mathrm{Hdg}})^{\mathrm{der}}$ satisfies the condition of Corollray 1.29 relative to $(H, \bar{h})$ and, hence, equals $\rho(H, \bar{h})$.

LEMMA 1.31. *The assignment* $(H, \bar{h}) \mapsto \rho(H, \bar{h})$ *is functorial: if* $\alpha\colon H \to H'$ *is a homomorphism mapping* $Z(H)$ *into* $Z(H')$ *and carrying* $\bar{h}$ *to* $\bar{h}'$, *then* $\rho(H', \bar{h}') = \alpha \circ \rho(H, \bar{h})$.

PROOF. After replacing H and H' with the subgroups generated by h and h', we may assume that α is surjective. Choose (G, h) for $(H, \bar{h})$ as in Proposition 1.28 and let $G' = G/\mathrm{Ker}(\alpha)$. Write α again for the projection $G \to G'$ and let $h' = \alpha_{\mathbb{R}} \circ h$. On restricting the maps in the equality

$$\rho(h') = \alpha \circ \rho(h)$$

to $(G_{\mathrm{Hdg}})^{\mathrm{der}}$ we obtain the equality

$$\rho(H', \bar{h}') = \alpha \circ \rho(H, \bar{h}). \quad \square$$

LEMMA 1.32. *Let* H *be a semisimple group over* $\mathbb{Q}$, *and let* $\bar{h}$ *be a homomorphism* $\mathbb{S}/\mathbb{G}_m \to H^{\mathrm{ad}}_{\mathbb{R}}$ *satisfying* (SV1, 2) *and generating* H. *If* $(H, \bar{h})$ *has a faithful family of symplectic representations, then* $\rho(H, \bar{h})$ *factors through* $(G_{\mathrm{Mab}})^{\mathrm{der}}$.

PROOF. It is clear from the definition of a symplectic representation (1.23) that $\rho(H, \bar{h})$ maps $\mathrm{Ker}(G_{\mathrm{Hdg}} \to G_{\mathrm{Mab}})$ into the kernel of any symplectic representation of H, but, by assumption, the intersection of these kernels is trivial. □

LEMMA 1.33. *Let* H *be a semisimple group over* $\mathbb{Q}$, *and let* $\bar{h}$ *be a homomorphism* $\mathbb{S}/\mathbb{G}_m \to H^{\mathrm{ad}}_{\mathbb{R}}$ *satisfying* (SV1, 2). *The homomorphism* $\rho(H, \bar{h})$ *factors through* $(G_{\mathrm{Mab}})^{\mathrm{der}}$ *if and only if* H *satisfies* (1.25).

PROOF. Suppose H satisfies (1.25). According to Theorem 1.26, there is a finite covering $\alpha\colon H' \to H$ such that $(H', \bar{h})$ has a faithful family of symplectic representations. By Lemma 1.32, $\rho(H', \bar{h})$ factors through $(G_{\mathrm{Mab}})^{\mathrm{der}}$, and therefore so also does $\rho(H, \bar{h}) = \alpha \circ \rho(H', \bar{h})$.

Conversely, suppose $\rho(H, \bar{h})$ factors through $(G_{\mathrm{Mab}})^{\mathrm{der}}$. There will be an algebraic quotient (G, h) of $(G_{\mathrm{Mab}}, h_{\mathrm{Mab}})$ such that $(H, \bar{h})$ is a quotient of $(G^{\text{der}}, \text{ad} \circ h)$. Consider the category of abelian motives M for which the action of G_{Mab} on $\omega_B(M)$ factors through G. By definition, this category is contained in the tensor category generated by $h_1(A)$ for some abelian variety A. We can replace G with the Mumford-Tate group of A. Then $(G^{\text{der}}, \text{ad} \circ h)$ has a symplectic embedding, and according to Theorem 1.26, this implies that G^{der} satisfies (1.25). Since H is a quotient of G^{der}, it also satisfies (1.25). □

We can now complete the proof of the Theorem 1.27. From Corollary 1.9, we know that $\rho(h)$ factors through G_{Mab} if and only if $\rho(G^{\text{der}}, \text{ad} \circ h)$ factors through $(G_{\text{Mab}})^{\text{der}}$, and from Lemma 1.33 we know that this is true if and only if G^{der} satisfies (1.25). □

COROLLARY 1.34. *For any pair* (G, h) *satisfying* (SV1, 2*, 3) *and such that* G^{der} *satisfies* (1.25), *there is a unique homomorphism* $\rho(h)\colon G_{\text{Mab}} \to G$ *such that* $\rho(h)_{\mathbb{R}} \circ h_{\text{Mab}} = h$.

PROOF. Let G' be the subgroup of G generated by h. The theorem implies that (G', h) is the Mumford-Tate group of an abelian motive, and so it is a quotient of $(G_{\text{Mab}}, h_{\text{Mab}})$. □

COROLLARY 1.35. *Let* (V, h) *be a special Hodge structure, and let* (G, h) *be its Mumford-Tate group. Then* (V, h) *is the Betti realization of an abelian motive if and only if* G^{der} *satisfies* (1.25).

PROOF. From the theorem we know that the action of G_{Hdg} on V factors through G_{Mab} if and only if G^{der} satisfies (1.25). □

REMARK 1.36. In order to prove Hypothesis 1.13, it suffices to prove the following: let H be a simple simply connected group over $\mathbb{Q}$, and let $\bar{h}\colon \mathbb{S}/\mathbb{G}_m \to H^{\text{ad}}_{\mathbb{R}}$ be a homomorphism satisfying (SV1, 2); then $\rho(H, \bar{h})$ factors through G_{Mot} where G_{Mot} is the group attached to **Mot**(ℂ) and the Betti fibre functor. If we knew all Hodge cycles were absolutely Hodge, this would be equivalent to showing that such a pair is of the form $(G^{\text{der}}, \text{ad} \circ h)$ where (G, h) is the Mumford-Tate group of the Betti realization of a motive. Of course, this has to be shown only for groups H not satisfying (1.25), i.e., for (simply connected) groups of type $D^{\mathbb{H}}$, mixed type D, and the exceptional types E_6 and E_7.

The extended Mumford-Tate group. Let (V, h) be a polarizable rational Hodge structure, and let (G', h') be the Mumford-Tate group of $V \oplus \mathbb{Q}(1)$. The action of G' on $\mathbb{Q}(1)$ determines a homomorphism $t\colon G' \to \mathbb{G}_m$ (defined over $\mathbb{Q}$) such that $t \circ w_{h'} = -2$, and we define the *extended Mumford-Tate group* of (V, h) to be the triple (G', h', t). We want to relate (G', h', t) to the Mumford-Tate group (G, h) of (V, h).

Suppose first that (V, h) has weight zero. Then $G' = G \times \mathbb{G}_m$, h' is the map

$$z \mapsto (h(z), |z|^{-2}), \qquad z \in \mathbb{C}^\times,$$

and t is the projection map.

When the weight is not zero, there is a smallest $m > 0$ for which $\mathbb{Q}(m)$ is in the tensor category generated by (V, h), and there is a commutative diagram

$$\begin{array}{ccccccc} 1 \to & \mathrm{Ker}(t) & \to G' & \xrightarrow{t} & \mathbb{G}_m & \to 1 \\ & \| & \downarrow & & \downarrow m & \\ 1 \to & \mathrm{Ker}(t_m) & \to G & \xrightarrow{t_m} & \mathbb{G}_m & \to 1 \end{array}$$

in which t_m is the map defined by the action of G on $\mathbb{Q}(m)$ and the right-hand square is cartesian, i.e.,

$$G' = G \underset{t_m, \mathbb{G}_m, m}{\times} \mathbb{G}_m.$$

Moreover, $G = G'/\mathrm{Ker}(m \circ t)$. Using these statements, it is possible to translate the above results in terms of extended Mumford-Tate groups.

The variation of Mumford-Tate groups in families. Let S be a complex manifold. A *holomorphic family of rational Hodge structures on* S is a triple $(\mathbb{V}, F, W)$ consisting of a local system $\mathbb{V}$ of $\mathbb{Q}$-vector spaces on S together with a decreasing (Hodge) filtration of the vector bundle $\mathscr{V} \underset{\mathrm{df}}{=} \mathscr{O}_S \otimes_{\mathbb{Q}} \mathbb{V}$ by holomorphic subbundles

$$\cdots \supset F^p\mathscr{V} \supset F^{p+1}\mathscr{V} \supset \cdots$$

and a (weight) gradation $\mathbb{V} = \bigoplus W_m\mathbb{V}$ of $\mathbb{V}$ such that, for every $s \in S$ and every $m \in \mathbb{Z}$, the $\mathbb{Q}$-vector space $W_m\mathbb{V}_s$ together with the filtration induced by F is a rational Hodge structure of weight m.

PROPOSITION 1.37. *Let $\mathbb{V} = (\mathbb{V}, F, W)$ be a holomorphic family of rational Hodge structures on a complex manifold S, and let G_s be the Mumford-Tate group of $\mathbb{V}_s$, $s \in S$. Then there exists a subset U of S with thin complement such that $s \mapsto G_s$ is locally constant on U.*

This is a consequence of the following more precise result.

PROPOSITION 1.38. *Let $(\mathbb{V}, F, W)$ be a holomorphic family of rational Hodge structures on a connected complex manifold S, and assume that $\mathbb{V}$ is the constant sheaf with stalk V. Regard G_s as a subgroup of* $\mathrm{GL}(V)$. *Then there exists a subset U of S with thin complement such that G_u is constant for $u \in U$, say $G_u = G$, and $G \supset G_s$ for $s \notin U$.*

PROOF. A tensor $t \in V^{\otimes r} \otimes V^{\vee \otimes r'}$ is a Hodge tensor for a Hodge structure h on V if and only if it has weight 0 and lies in $F_h^0(V^{\otimes r} \otimes V^{\vee\otimes r'})$. Thus the $s \in S$ where t is a Hodge tensor form a closed analytic set. Let G be the subgroup of $\mathrm{GL}(V)$ fixing all tensors $t \in V^{\otimes r} \otimes V^{\vee\otimes r}$ that are Hodge tensors for all $s \in S$. We can take U to be the set of s such that $G = G_s$. □

2. Moduli of motives

We discuss the problem of realizing a motive as a member of a universal family.

The concept of a moduli variety. Let Ω be an algebraically closed field. Suppose we have a contravariant functor $\mathcal{M}$ from the category of algebraic varieties over Ω to the category of sets, and equivalence relations $\sim$ on each of the sets $\mathcal{M}(T)$ that are compatible with morphisms in the sense that

$$m \sim m' \implies \phi^*(m) \sim \phi^*(m'), \qquad m, m' \in \mathcal{M}(S), \ \phi\colon T \to S;$$

then the pair $(\mathcal{M}, \sim)$ is called a *moduli problem over* Ω. A point t of a variety T with coordinates in Ω defines a map $m \mapsto m_t \underset{\mathrm{df}}{=} t^*m\colon \mathcal{M}(T) \to \mathcal{M}(\Omega)$. A *solution to the moduli problem* is a variety S over Ω and a bijection $\alpha\colon \mathcal{M}(\Omega)/{\sim} \to S(\Omega)$ with the properties:

(2.1.1) for every variety T over Ω and $m \in \mathcal{M}(T)$, the map $t \mapsto \alpha(m_t)\colon T(\Omega) \to S(\Omega)$ is a morphism of algebraic varieties over Ω;

(2.1.2) for some open covering $\{S_i\}$ of S, there exist elements $m_i \in \mathcal{M}(S_i)$ such that $\alpha((m_i)_s) = s$ for all $s \in S$.

These conditions determine (S, α) uniquely up to a unique isomorphism, for if (S', α') is a second pair satisfying the conditions (2.1), then (2.1.2) for S and (2.1.1) for S' show that the map $\alpha' \circ \alpha^{-1}\colon S(\Omega) \to S'(\Omega)$ becomes a morphism when restricted to the members of some open covering of S, and is therefore a morphism $S \to S'$. The same argument shows that its inverse is also a morphism.

A solution (S, α) to the moduli problem is said to be *fine* (and S is called a *fine moduli variety*) if

(2.2.1) for every variety T over Ω, the equivalence class of $m \in \mathcal{M}(T)$ is determined by the equivalence classes of the elements m_t, $t \in T(\Omega)$;

(2.2.2) there is an element $m_0 \in \mathcal{M}(S)$ such that $\alpha(m_{0s}) = s$ for all $s \in S(\Omega)$.

Choose an m_0 as in (2.2.2). Then (2.2.1) implies that for any $m \in \mathcal{M}(T)$, there is a unique morphism $\varphi\colon T \to S$ such that $\varphi^*(m_0) \sim m$. Therefore, the pair $(S, [m_0])$ represents the functor $T \mapsto \mathcal{M}(T)/{\sim}$. Conversely, if there exists an $m_0 \in \mathcal{M}(S)$ such that $(S, [m_0])$ represents $\mathcal{M}/{\sim}$, then $(\mathcal{M}, \sim)$ is a fine moduli problem and (S, α), with α the inverse of $s \mapsto [m_{0s}]$, is a solution to the moduli problem.

There are variants of these definitions that are also useful. For example, we could replace the category of varieties over Ω with that of smooth varieties over Ω or we could allow the covering in (2.1.2) to be with respect to the étale or flat topologies. Alternatively, we could replace the category of algebraic varieties with that of complex manifolds.

In §3, we shall also need the notion of a moduli variety over a nonalgebraically closed field k. To avoid problems with inseparability, we assume k to be of characteristic zero. A moduli problem over k is a pair $(\mathscr{M}, \sim)$ as before, but with $\mathscr{M}$ a functor from the category of algebraic varieties over k to that of sets. Fix an algebraically closed field Ω containing k, and define

$$\mathscr{M}(\Omega) = \varinjlim \mathscr{M}(R)$$

where the limit is over the subalgebras R of Ω that are finitely generated over k and $\mathscr{M}(R) = \mathscr{M}(\operatorname{Spec} R)$. The equivalence relations on the sets $\mathscr{M}(R)$ define an equivalence relation on $\mathscr{M}(\Omega)$. A point t of a k-variety T with coordinates in Ω, i.e., a k-morphism $\operatorname{Spec}\Omega \to T$, defines a map $m \mapsto m_t : \mathscr{M}(T) \to \mathscr{M}(\Omega)$.

A *solution to the moduli problem* $(\mathscr{M}, \sim)$ is a variety S over k together with a bijection $\alpha : \mathscr{M}(\Omega)/{\sim} \to S(\Omega)$ with the properties:

(2.3.1) for every variety T over k and $m \in \mathscr{M}(T)$, there exists a morphism $\beta_m : T \to S$ such that $\beta_m(t) = \alpha(m_t)$ for all $t \in T(\Omega)$;

(2.3.2) for some open covering (S_i) of S, there exist elements $m_i \in \mathscr{M}(S_i)$ such that $\alpha((m_i)_t) = t$ for all $t \in S(\Omega)$.

The map β_m in (2.3.1) is uniquely determined, and the conditions (2.3) determine the pair (S, α) uniquely up to a unique isomorphism.

Moduli of Hodge structures. Since our approach to the moduli of motives is via their Hodge structures, it is natural to begin by considering the problem of realizing a rational polarizable Hodge structure as a member of a universal family. Experience from abelian varieties suggests that we should study the moduli, not of Hodge structures, but of *polarized* Hodge structures. Also that, in order to obtain more general results, we should endow the Hodge structures with additional structure, for example, with an endomorphism ring or, more generally, a family of Hodge tensors.

Thus let (V, h_0) be a polarizable rational Hodge structure, which, for simplicity, we take to be of pure weight m. Let $\mathfrak{t}' = (t'_i)_{i\in I}$ be a family of Hodge tensors of (V, h_0), i.e., elements of type $(0, 0)$ in $V^{\otimes r_i} \otimes V^{\vee\otimes s_i}(m_i)$ for some r_i, s_i, m_i with $mr_i - ms_i - 2m_i = 0$. We assume that $\mathfrak{t}'$ contains a tensor $t'_0 \in V^{\vee\otimes 2}(-m)$ that is a polarization for (V, h_0), and we write $t'_0 = \psi \otimes (2\pi i)^m$. Thus ψ is a nondegenerate bilinear pairing

$$\psi : V \times V \longrightarrow \mathbb{Q}$$

such that

(2.4.1) ψ is symmetric or skew-symmetric according as m is even or odd;

(2.4.2) $\psi(V^{p,q}, V^{r,s}) = 0$ if $p + r \neq m$; equivalently, $\psi(F^p, F^{m-p+1}) = 0$ for all p, where $F = F_{h_0}$ (the Hodge filtration of h_0);

(2.4.3) $\psi(v, h_0(i)v) > 0$, if $v \in V(\mathbb{R})$, $v \neq 0$; equivalently, $i^{p-q}\psi(v, \bar{v}) > 0$ if $v \in F^p \cap \overline{F^q}$, $v \neq 0$.

Let G' be the subgroup of $\mathrm{GL}(V) \times \mathbb{G}_m$ fixing the t'_i. The action of G' on $\mathbb{Q}(1)$ defines a homomorphism $t: G' \to \mathbb{G}_m$, and we let $G = \mathrm{Ker}(t)$. Write $t'_i = t_i \otimes (2\pi i)^{m_i}$ with $t_i \in V^{\otimes r_i} \otimes V^{\vee \otimes s_i}$, and let $\mathfrak{t} = (t_i)_{i \in I}$. Then t_i is of type (m_i, m_i), and G is the subgroup of $\mathrm{GL}(V)$ fixing the t_i, i.e., for all $\mathbb{Q}$-algebras R,

$$G(R) = \{\alpha \in \mathrm{GL}(V(R)) \mid (\alpha^{\otimes r_i} \otimes \check{\alpha}^{\otimes s_i})(t_i) = t_i, \text{ all } i \in I\}.$$

Note that $h'_0 \underset{\mathrm{df}}{=} (h_0, \mathrm{Nm})$ maps into $G'_{\mathbb{R}}$ and $u_0 = h_0|U^1$ maps into $G_{\mathbb{R}}$. In particular, $C \underset{\mathrm{df}}{=} h_0(i) \in G(\mathbb{R})$, and ψ is a C-polarization of the representation of G on $V(\mathbb{R})$. Therefore (see Lemma 1.5) G is reductive.

EXAMPLE 2.5. There are two cases of particular interest.

(a) The family $\mathfrak{t}'$ contains all the Hodge tensors of (V, h_0). In this case G' is the extended Mumford-Tate group of (V, h_0), and we call G the *special Mumford-Tate group* of (V, h_0).

(b) The family $\mathfrak{t}$ consists only of $t_0 = \psi$, so that G is the subgroup of $\mathrm{GL}(V)$ fixing ψ and $G' = G \cdot \mathbb{G}_m$. This case is studied in Griffiths **[34]**, **[35]**, Schmid **[64]**, El Zein **[24,** Chapter 7].

Let $\mathscr{F}^\vee$ be the set of filtrations F on $V(\mathbb{C})$ such that

$$\dim F^p = \dim F^p_{h_0} \text{ all } p.$$

The group $\mathrm{GL}(V(\mathbb{C}))$ acts transitively on $\mathscr{F}^\vee$, and the subgroup P stabilizing F_{h_0} is parabolic.[3] Hence the bijection

$$\mathrm{GL}(V)/P \longrightarrow \mathscr{F}^\vee, \quad gP \mapsto gF_{h_0}$$

realizes $\mathscr{F}^\vee$ as a smooth projective algebraic variety over $\mathbb{C}$. The subset $\mathscr{F}$ of $\mathscr{F}^\vee$ of those filtrations F such that

$$V(\mathbb{C}) = F^p \oplus \bar{F}^{m-p+1} \quad \text{for each } p,$$

is open in $\mathscr{F}^\vee$ (for the complex topology). It parametrizes exactly the Hodge structures on $V(\mathbb{R})$ whose Hodge numbers are the same as those of (V, h_0).

Consider the set of filtrations F in $\mathscr{F}^\vee$ such that, for all $i \in I$, $t_i \in F^{m_i}$. It is a closed algebraic subvariety of $\mathscr{F}^\vee$ stable under the action of $G(\mathbb{C})$, and we let $X^\vee$ be the orbit containing F_{h_0}. Then $X^\vee$ is the quotient of

[3]Let $\xi: G \hookrightarrow \mathrm{GL}(V)$ be a faithful representation of a reductive group G over a field k of characteristic zero, and let F be a filtration on V. Suppose that there exists a cocharacter μ of G splitting the filtration, i.e., such that $F^pV = \bigoplus_{i \geq p} V^i$, $V^i = \{v \in V \mid \xi\mu(x)v = x^i v, \forall x\}$. The subgroup P of G of elements preserving F is parabolic with unipotent radical the subgroup U of elements acting as the identity map on $\bigoplus F^iV/F^{i+1}V$. The cocharacter μ defines a filtration on every representation of G, in particular, on the adjoint representation of G on $\mathfrak{g}$, and $\mathrm{Lie}(P) = F^0\mathfrak{g}$, $\mathrm{Lie}(U) = F^1\mathfrak{g}$. See Saavedra **[59,** IV.2.2.5].

$G(\mathbb{C})$ by a parabolic subgroup, and so it is a smooth projective variety over $\mathbb{C}$. An $F \in X^\vee$ satisfying (2.4.3) automatically defines a Hodge structure (see the references in Example 2.5(b)) and so lies in $\mathscr{F}$. The set of such F's is open in $X^\vee$ and stable under $G(\mathbb{R})$, and we write X for the orbit containing F_{h_0}.

We now regard X as a complex manifold rather than a set of filtrations, and we write F_x for the filtration, and h_x, u_x, and μ_x for the homomorphisms, corresponding to $x \in X$.

REMARK 2.6. (a) When $\mathfrak{t}$ consists only of t_0, $X^\vee$ contains all filtrations such that $t_0' \in F^0$, and X contains all $F \in X^\vee$ satisfying (2.4.3). Thus in this case, X contains all the filtrations on $V(\mathbb{C})$ defining Hodge structures for which $(2\pi i)^m \psi$ is a polarization (see the references in Example 2.5(b)).

(b) Let $x \in X^\vee \cap \mathscr{F}$, and let F_x be the corresponding filtration on $V(\mathbb{C})$. The cocharacter $\mu_x \colon \mathbb{G}_m \to \mathrm{GL}(V(\mathbb{C}))$ has image in $G \cdot Z$ where $Z = \mathbb{G}_m = Z(\mathrm{GL}(V(\mathbb{C}))$; so μ_x defines a filtration on $\mathfrak{g}_\mathbb{C}$. The stabilizer of F_x in G has Lie algebra $F^0\mathfrak{g}_\mathbb{C}$, and so

$$\mathfrak{g}_\mathbb{C}/F^0\mathfrak{g}_\mathbb{C} \overset{\approx}{\to} \mathrm{Tgt}_x X^\vee.$$

Now assume $x \in X$. The stabilizer B of x in $G_\mathbb{R}$ has Lie algebra $\mathfrak{g}_\mathbb{R} \cap F^0\mathfrak{g}_\mathbb{C} = \mathfrak{g}_\mathbb{R} \cap \mathfrak{g}^{0,0}$, and so

$$\mathfrak{g}_\mathbb{R}/(\mathfrak{g}_\mathbb{R} \cap \mathfrak{g}^{0,0}) \overset{\approx}{\to} \mathrm{Tgt}_x X.$$

Note that

$$\dim \mathfrak{g}_\mathbb{C}/F^0\mathfrak{g}_\mathbb{C} = \dim \mathfrak{g}_\mathbb{R}/(\mathfrak{g}_\mathbb{R} \cap \mathfrak{g}^{0,0})$$

(as real vector spaces).

Choose a lattice $V(\mathbb{Z})$ in V, and let

$$\Gamma(N) = \{g \in G(\mathbb{Q}) \mid gV(\mathbb{Z}) = V(\mathbb{Z}),\ g = \mathrm{id} \text{ on } V(\mathbb{Z})/NV(\mathbb{Z})\}.$$

Consider a triple $(W, \mathfrak{s}, \eta)$ consisting of an integral Hodge structure $W = (W(\mathbb{Z}), h)$, a family of tensors $\mathfrak{s} = (s_i)_{i \in I}$ of $W(\mathbb{Q}) \underset{\mathrm{df}}{=} W(\mathbb{Z}) \otimes \mathbb{Q}$, and an isomorphism

$$\eta \colon V(\mathbb{Z})/NV(\mathbb{Z}) \longrightarrow W(\mathbb{Z})/NW(\mathbb{Z}).$$

We call η a *level-N-structure* on W. Let $\mathscr{H}(\mathbb{C})$ be the set of such triples for which there exists an isomorphism $\beta \colon W(\mathbb{Q}) \to V$ satisfying the following conditions:

(2.7.1) for some $x \in X$, β is a morphism of rational Hodge structures

$$(W(\mathbb{Q}), h) \longrightarrow (V, h_x);$$

(2.7.2) for all $i \in I$, $\beta(s_i) = t_i$;

(2.7.3) $\beta(W(\mathbb{Z})) = V(\mathbb{Z})$ and the composite

$$V(\mathbb{Z})/NV(\mathbb{Z}) \xrightarrow{\eta} W(\mathbb{Z})/NW(\mathbb{Z}) \xrightarrow{\beta(N)} V(\mathbb{Z})/NV(\mathbb{Z})$$

is the identity map.

Note that the conditions imply that s_i is an element of type (m_i, m_i) in $W(\mathbb{Q})^{\otimes r_i} \otimes W(\mathbb{Q})^{\vee \otimes s_i}$. An *isomorphism* from one such system $(W, \mathfrak{s}, \eta)$ to a second $(W', \mathfrak{s}', \eta')$ is an isomorphism of integral Hodge structures $\gamma: (W(\mathbb{Z}), h) \to (W'(\mathbb{Z}), h')$ such that $\gamma(s_i) = s'_i$ for all $i \in I$ and $\gamma(N) \circ \eta = \eta'$.

Let $(W, \mathfrak{s}, \eta)$ be such a system, and choose an isomorphism $\beta: W(\mathbb{Q}) \to V$ satisfying the conditions (2.7). A second such isomorphism is of the form $g \circ \beta$ for some $g \in \mathrm{GL}(V)$. But (2.7.2) implies that $g(t_i) = t_i$ for all i, and so $g \in G(\mathbb{Q})$. Also (2.7.3) implies that $gV(\mathbb{Z}) = V(\mathbb{Z})$ and that g acts as the identity map on $V(\mathbb{Z})/NV(\mathbb{Z})$, and so $g \in \Gamma(N)$. Therefore, when we write $\mathrm{ad}\,\beta \circ h = h_x$, the orbit of x in $\Gamma(N)\backslash X$ is independent of the choice of β.

PROPOSITION 2.8. *The map* $(W, \mathfrak{s}, \eta) \mapsto [x]$ *just defined gives a bijection*

$$\alpha: \mathscr{H}(\mathbb{C})/\approx \longrightarrow \Gamma(N)\backslash X.$$

PROOF. Let $(W', \mathfrak{s}', \eta')$ be a second system. If

$$\gamma: (W', \mathfrak{s}', \eta') \to (W, \mathfrak{s}, \eta)$$

is an isomorphism of triples and $\beta: W(\mathbb{Q}) \to V(\mathbb{Q})$ is an isomorphism of vector spaces satisfying (2.7), then $\beta \circ \gamma$ satisfies (2.7) for $(W', \mathfrak{s}', \eta')$, and it follows that $(W', \mathfrak{s}', \eta')$ maps to the same element of $\Gamma(N)\backslash X$ as $(W, \mathfrak{s}, \eta)$. Conversely, if $(W, \mathfrak{s}, \eta)$ and $(W', \mathfrak{s}', \eta')$ map to the same class $[x]$, then we can choose the maps β and β' so that the triples map to the same element of X; now $\gamma \underset{\mathrm{df}}{=} \beta^{-1} \circ \beta'$ is an isomorphism

$$(W', \mathfrak{s}', \eta') \longrightarrow (W, \mathfrak{s}, \eta).$$

Finally, if $x \in X$, then $((V(\mathbb{Z}), h_x), \mathfrak{t}, \mathrm{id})$ maps to $[x]$. □

We next wish to endow $\Gamma(N)\backslash X$ with the structure of a complex manifold.

LEMMA 2.9. (a) *The group* $\Gamma(N)$ *acts properly discontinuously*[4] *on* X.
(b) *For* N *sufficiently divisible,* $\Gamma(N)$ *is torsion-free.*

PROOF. (a) According to a standard criterion (see Miyake **[54,** 1.5.2], it suffices to check that the stabilizers in $G(\mathbb{R})$ of the elements of X are compact and that $\Gamma(N)$ is a discrete subgroup of $G(\mathbb{R})$, but the stabilizer of $h \in X$ is compact because it fixes the positive definite symmetric form

$$(v, w) \mapsto \psi(v, Cw), \qquad C = h(i),$$

and $\Gamma(N)$ is discrete because it is a congruence subgroup.

(b) If N is sufficiently divisible, $\Gamma(N)$ will be neat, and hence torsion-free (see Borel **[4,** §17]). □

[4]Recall that this means that for any pair of points x_1 and x_2 of X, there exist neighbourhoods U_1 and U_2 of x_1 and x_2 such that $\{g \in \Gamma(N) \mid gU_1 \cap U_2 \neq \varnothing\}$ is finite.

PROPOSITION 2.10. *If* $\Gamma(N)$ *is torsion-free, then* $\Gamma(N)\backslash X$ *has a unique structure of a complex manifold such that the quotient map* $X \to \Gamma(N)\backslash X$ *is a local isomorphism.*

PROOF. If $\Gamma(N)$ is torsion-free, then the map $X \to \Gamma(N)\backslash X$ is a local homeomorphism. □

When N is sufficiently divisible that $\Gamma(N)$ is neat, we write $S(N)$ for $\Gamma(N)\backslash X$ regarded as a complex manifold.

An *integral structure* on a holomorphic family of rational Hodge structures $(\mathbb{V}, F, W)$ is a local system of $\mathbb{Z}$-modules $\mathbb{V}(\mathbb{Z}) \subset \mathbb{V}$ such that $\mathbb{V}(\mathbb{Z})\otimes\mathbb{Q} = \mathbb{V}$ and $\mathbb{V}(\mathbb{Z}) = \bigoplus \mathbb{V}(\mathbb{Z}) \cap W_m$. A holomorphic family of rational Hodge structures together with an integral structure will be referred to as a *holomorphic family of integral Hodge structures.*

Let T be a complex manifold, and consider triples $m = (\mathbb{W}, \mathfrak{s}, \eta)$ consisting of a holomorphic family of integral Hodge structures $\mathbb{W} = (\mathbb{W}(\mathbb{Z}), F, W)$ on T, a family of global tensors $\mathfrak{s}$ of $\mathbb{W}(\mathbb{Q}) \underset{\mathrm{df}}{=} \mathbb{W}(\mathbb{Z}) \otimes \mathbb{Q}$ indexed by I, and an isomorphism η from the constant sheaf $(V(\mathbb{Z})/NV(\mathbb{Z}))_T$ on T to $\mathbb{W}(\mathbb{Z})/N\mathbb{W}(\mathbb{Z})$. We define $\mathscr{H}(T)$ to be the set of such triples m with the property that $m_t \in \mathscr{H}(\mathbb{C})$ for all $t \in T$. An *isomorphism* $(\mathbb{W}, \mathfrak{s}, \eta) \xrightarrow{\approx} (\mathbb{W}', \mathfrak{s}', \eta')$ is an isomorphism of integral Hodge structures $\mathbb{W} \to \mathbb{W}'$ carrying $\mathfrak{s}$ and η into $\mathfrak{s}'$ and η'. With the obvious notion of pull-back, the pair $(\mathscr{H}, \approx)$ becomes a moduli problem on the category of complex manifolds and holomorphic maps.

EXAMPLE 2.11. On X there is a holomorphic family of integral Hodge structures of weight m whose underlying local system of $\mathbb{Z}$-modules is the constant system defined by $V(\mathbb{Z})$ and which is such that the filtration at a point x is the Hodge filtration of h_x.

If N is sufficiently divisible, then $\Gamma(N)$ is the group of covering transformations of X over $S(N)$, and we obtain a holomorphic family of integral Hodge structures $\mathbb{V}(N)$ on $S(N)$: for $o = h_0$, the local system of $\mathbb{Z}$-modules underlying $\mathbb{V}(N)$ corresponds to the representation of $\Gamma(N) = \pi_1(S(N), o)$ on $V(\mathbb{Z})$. Because t_i is fixed by $\Gamma(N)$, it defines a global tensor of $\mathbb{V}(N)$, and because $\Gamma(N)$ acts trivially on $V(\mathbb{Z})/NV(\mathbb{Z})$, there is a canonical isomorphism $\eta\colon (V(\mathbb{Z})/NV(\mathbb{Z}))_{S(N)} \to \mathbb{V}(N)/N\mathbb{V}(N)$. The system $m_0 = (\mathbb{V}(N), \mathfrak{t}, \eta) \in \mathscr{H}(S(N))$.

THEOREM 2.12. *For N sufficiently divisible, the pair* $(S(N), \alpha)$ *is a fine solution to the moduli problem* $(\mathscr{H}, \approx)$ *(in the category of complex manifolds and holomorphic maps).*

PROOF. Let $m = (\mathbb{W}, \mathfrak{s}, \eta) \in \mathscr{H}(T)$ for some complex manifold T. We have to prove that the map

$$\varphi_m\colon T \longrightarrow S(N), \qquad t \mapsto \alpha(m_t),$$

is holomorphic. Let $t_0 \in T$. Choose an open neighbourhood U of t_0 over which $\mathbb{W}$ is trivial, and fix an isomorphism $\mathbb{W}|U \approx V_U$ (constant local system on U). The map $t \mapsto F_t : T \to \mathscr{F}$ is holomorphic, and its image is contained in X. Hence it is holomorphic as a map into X, and so the composite $U \to X \to S(N)$ is holomorphic. Obviously, $\varphi_m^* m_0 \approx m$, and φ_m is the unique morphism with this property, and so $(S(N), [m_0])$ represents the functor $\mathscr{H}/\approx$. □

REMARK 2.13. In the above, we should of course allow (V, h_0) to have more than one weight, but this complicates the exposition. Once one is willing to work in that generality, it is natural to replace V with $V \oplus \mathbb{Q}(1)$ to ensure that $\mathbb{Q}(1)$ is in the tensor category generated by (V, h_0).

From now on, we assume that we are in case 2.5(a). It follows from Proposition 1.38, that there is no essential loss of generality in doing this.

Questions. If (V, h_0) is the Betti realization of a motive, then is $S(N)$ a moduli variety for motives? Four problems present themselves.

(2.14.1) Does $S(N)$ have a (unique) structure of an algebraic variety compatible with its complex structure?

(2.14.2) Is every Hodge structure (V, h), $h \in X$, motivic, i.e., the Betti realization of a motive?

(2.14.3) Assuming that (2.14.1) and (2.14.2) have positive answers, is the canonical family $\mathbb{V}(N)$ of Hodge structures on $S(N)$ defined in Example 2.11 the Betti realization of a "family of motives" (whatever that may be)?

(2.14.4) Is $S(N)$ a moduli variety for motives? Since $S(N)$ parametrizes Hodge structures with additional structure and we are asking that it parametrize motives, this implies that the motives in the family are determined (up to isomorphism) by their Betti realizations.

We shall see shortly that, in order for (2.14.3) to be true, it is necessary that (V, h_0) be a special Hodge structure. Remarkably, when we assume this and that Hypothesis 1.13 holds in families, then all statements become true.

Variations of Hodge structures. Let S be a connected complex manifold. Recall that a *connection* on a holomorphic vector bundle $\mathscr{V}$ is a $\mathbb{C}$-linear homomorphism

$$\nabla : \mathscr{V} \longrightarrow \Omega_S^1 \otimes \mathscr{V}$$

satisfying the Leibnitz identity

$$\nabla(fv) = df \otimes v + f \cdot \nabla v,$$

for f and v local sections of $\mathscr{O}$ and $\mathscr{V}$. A connection is *flat* if its curvature tensor is zero. A local section of $\mathscr{V}$ is *horizontal* if $\nabla v = 0$, and we write $\mathscr{V}^\nabla$ for the sheaf of horizontal sections. The map $(\mathscr{V}, \nabla) \mapsto \mathscr{V}^\nabla$ defines an equivalence from the category of vector bundles with flat connections to that of complex local systems on S.

DEFINITION 2.15. A holomorphic family of rational Hodge structures $(\mathbb{V}, F, W)$ on a complex manifold is a *variation of rational Hodge structures* if

(2.15.1) $(\mathbb{V}, F, W)$ admits an integral structure;
(2.15.2) (axiom of transversality): $\nabla(F^p\mathscr{V}) \subset \Omega_S^1 \otimes F^{p-1}\mathscr{V}$.

A *polarization* of a variation of Hodge structures $\mathbb{V}$ of weight m is a morphism of local systems $\mathbb{V} \otimes \mathbb{V} \to \mathbb{Q}(-m)$ that at each point s of S defines a polarization of the Hodge structure $\mathbb{V}_s$.

PROPOSITION 2.16. *The category* $\mathbf{Hdg}_{\mathbb{Q}}(S)$ *of polarizable variations of rational Hodge structures on a connected complex manifold* S *is a semisimple Tannakian category over* $\mathbb{Q}$.

PROOF. It is obvious that $\mathbf{Hdg}_{\mathbb{Q}}(S)$ is closed under the formation of direct summands, direct sums, and tensor products; moreover, it contains the constant variations $\mathbb{Q}(m)$. Therefore, we can apply Deligne **[14**, 4.2.3] to deduce that $\mathbf{Hdg}_{\mathbb{Q}}(S)$ is a semisimple abelian subcategory of the category of continuous families of Hodge structures and that it is closed under the formation of duals. For any point $o \in S$, $(\mathbb{V}, F, W) \mapsto \mathbb{V}_o$ is a fibre functor, and so $\mathbf{Hdg}_{\mathbb{Q}}(S)$ is a Tannakian category. The identity object is $\mathbb{Q}(0)$, and $\mathrm{End}(\mathbb{Q}(0)) = \mathbb{Q}$. □

THEOREM 2.17. *Let* $\pi: Y \to S$ *be a smooth projective map of algebraic varieties over* $\mathbb{C}$; *for any* r, $R^r\pi_*\mathbb{Q}$ *is a polarizable variation of rational Hodge structures on* S *of weight* r.

PROOF. This is a fundamental result in Griffiths's theory. For proofs see Griffiths **[34]** and Deligne **[16]**. □

The last two results show that the Betti realization of a family of motives must be a polarizable variation of Hodge structures. It is therefore natural to require that $\mathbb{V}(N)$ be a variation of Hodge structures on $S(N)$ or, equivalently, that (V, F_x) be a variation of Hodge structures on X. There is a simple criterion for this.

PROPOSITION 2.18. *The following statements are equivalent*:

(a) *the family* (V, F_x) *is a variation of Hodge structures on* X;
(b) *for all* $x \in X$, (G', h'_x) *satisfies* (SV1);
(c) *for all* $x \in X$, *the Hodge structure* (V, h_x) *is special*;
(d) *the Hodge structure* (V, h_0) *is special.*

PROOF. (a) $\Leftrightarrow$ (b). Consider the inclusion map $\varphi: X \hookrightarrow X^\vee$. The map on the tangent spaces at a point x of X is

$$(d\varphi)_x : \mathrm{Tgt}_x(X) = \mathfrak{g}_{\mathbb{R}}/(\mathfrak{g}_{\mathbb{R}} \cap \mathfrak{g}^{0,0}) \xrightarrow{\approx} \mathfrak{g}_{\mathbb{C}}/F^0\mathfrak{g}_{\mathbb{C}}, \quad \mathfrak{g} = \mathrm{Lie}(G),$$

(see Remark 2.6). The axiom of tranversality says that the image of $(d\varphi)_x$

is contained in $F_x^{-1}\mathfrak{g}_{\mathbb{C}}/F_x^0\mathfrak{g}_{\mathbb{C}}$ for all x, i.e., that $\mathfrak{g}_{\mathbb{C}} = F_x^{-1}\mathfrak{g}_{\mathbb{C}}$. But

$$\begin{aligned}\mathfrak{g}_{\mathbb{C}} = F_x^{-1}\mathfrak{g}_{\mathbb{C}} &\Leftrightarrow \mathfrak{g} \text{ is of type } \{(-1, 1), (0, 0), (1, -1)\} \\ &\Leftrightarrow (G', h'_x) \text{ satisfies (SV1).}\end{aligned}$$

(b) ⇒ (c). The extended Mumford-Tate group of (V, h_x) is a subgroup of G'.

(c) ⇒ (d). Obvious.

(d) ⇒ (b). Because of our assumption that we are in case 2.5(a), G' is the extended Mumford-Tate group of (V, h_0), and so (d) says that (G', h'_0) satisfies (SV1). To deduce (b), set $x = gx_0$, $g \in G(\mathbb{R})$, and note that $\mathrm{ad}(g)$ is an isomorphism $(G', h'_0) \to (G', h'_x)$. □

Before providing answers to the questions (2.14), we review some of the fundamental results concerning variations of Hodge structures.

THEOREM 2.19. *Let* $\mathbb{V}$ *be a polarizable variation of Hodge structures on a smooth quasi-projective algebraic variety* S *over* $\mathbb{C}$. *Then the vector bundle* $\mathscr{V} \underset{\mathrm{df}}{=} \mathbb{V} \otimes \mathscr{O}_S^{\mathrm{an}}$ *carries a unique algebraic structure such that the connection* ∇ *becomes algebraic and such that* ∇ *has regular singular points at infinity relative to any smooth compactification of* M. *With respect to this structure, the subbundles* $F^r \subset \mathscr{V}$ *are algebraic.*

PROOF. See Schmid [**64,** 4.13], where the result is credited to Griffiths. □

REMARK 2.20. Let $\pi\colon Y \to S$ be a smooth projective morphism of algebraic varieties over a field k. The vector bundle $R^r\pi_*\mathscr{O}_Y$ has a canonical (de Rham) filtration F_{dR} arising from the identification $R^r\pi_*\mathscr{O}_Y = \mathbb{R}^r\pi_*\Omega^{\cdot}_{Y/S}$ and a flat (Gauss-Manin) connection ∇. When $k = \mathbb{C}$ these structures agree with those defined by Theorems 2.17 and 2.19.

THEOREM 2.21 (THEOREM OF THE FIXED PART). *Let* S *be a smooth connected algebraic variety over* $\mathbb{C}$, *and let* $\mathbb{V}$ *be a polarizable variation of Hodge structures on* S. *The largest constant local system* $\mathbb{V}^f \subset \mathbb{V}$ *is a constant variation of Hodge substructures of* $\mathbb{V}$.

PROOF. When the base space is a compact complex manifold, this is proved in Griffiths [**35,** §7], and in the general case it is proved in Schmid [**64,** 7.22]. See also Deligne [**14,** 4.1.2, and footnote p. 45]. □

Consequently, a global section of $\mathbb{V}$ (in fact, even of $\mathbb{V} \otimes \mathbb{C}$) that is of type (p, q) at one point is of type (p, q) at every point. When we apply this to an endomorphism of $\mathbb{V}$, we obtain the following result.

COROLLARY 2.22. *Let* $(\mathbb{V}, F)$ *and* $(\mathbb{V}', F')$ *be two polarizable variations of Hodge structures on a smooth connected algebraic variety* S. *An isomorphism* $\alpha\colon \mathbb{V} \to \mathbb{V}'$ *of local systems such that* $\alpha(o)$ *is an isomorphism of Hodge structures for some* $o \in S(\mathbb{C})$ *is an isomorphism of variations of Hodge structures.*

Let S be a smooth connected algebraic variety over $\mathbb{C}$. The corollary implies that, given a representation of $\pi_1(S, o)$ on a finite-dimensional rational vector space V and a polarized Hodge structure on V, there is at most one way of extending these data to a polarizable variation of Hodge structures on S.

Algebraicity of $S(N)$. Recall that a *bounded domain* in $\mathbb{C}^g$ is a bounded open connected subset of $\mathbb{C}^g$, and that a *Hermitian manifold* is a complex manifold together with a holomorphic family of positive-definite Hermitian forms on its tangent spaces. A bounded domain or Hermitian manifold D is said to be *symmetric* if each point of D is an isolated fixed point of an involution of D. A complex manifold isomorphic to a symmetric bounded domain is called a *symmetric Hermitian domain.* Since the Bergmann metric provides a symmetric bounded domain with a Hermitian structure invariant under all automorphisms of the domain, every symmetric Hermitian domain has a canonical Hermitian structure with respect to which it is symmetric.

PROPOSITION 2.23. *If* (V, h_0) *is special, then every connected component of* X *is a symmetric Hermitian domain.*

PROOF. Identify a connected component X^+ of X with a $G'(\mathbb{R})^+$-conjugacy class of homomorphisms $\mathbb{S} \to G'_{\mathbb{R}}$, and apply **[18,** 1.1.17]. □

THEOREM 2.24. *Let* X^+ *be a symmetric Hermitian domain, and let* Γ *be a torsion-free arithmetic subgroup of* $\mathrm{Aut}(X^+)$. *Then* $S(\Gamma) \underset{\mathrm{df}}{=} \Gamma\backslash X^+$ *has a unique structure of an algebraic variety with the following property*:

(2.24.1) *every holomorphic map* $S \to S(\Gamma)$ *from a smooth complex algebraic variety to* $S(\Gamma)$ *is a morphism of algebraic varieties.*

PROOF. The main theorem of Baily and Borel **[1]** shows that $S(\Gamma)$ has a canonical algebraic structure compatible with its complex structure. That the structure has the property (2.24.1) is proved in Borel **[5,** 3.10]. It is obvious that this property determines the algebraic structure uniquely. □

THEOREM 2.25. *If* (V, h_0) *is special, then* $\mathbb{V}(N)$ *is a polarizable variation of Hodge structures on* $S(N)$, *and* $S(N)$ *has a unique algebraic structure compatible with its complex structure. Conversely, if* $\mathbb{V}(N)$ *is a variation of Hodge structures, then* (V, h_0) *is special.*

PROOF. Combine Propositions 2.18 and 2.23 and Theorem 2.24. □

The motivicity of (V, h_x).

PROPOSITION 2.26. *If* (V, h_0) *is the Betti realization of an abelian motive, then so also is* (V, h_x) *for all* $x \in X$.

PROOF. We saw in Proposition 2.18 that if (V, h_0) is a special Hodge structure, then so also is (V, h_x) for all $x \in X$.

Let (G_0, h_0) and (G_x, h_x) be the Mumford-Tate groups of (V, h_0) and (V, h_x) respectively (thus $G = \mathrm{Ker}(G_0 \to \mathbb{G}_m)$). If $h_x = \mathrm{ad}\, g \circ h_0$ with $g \in G_0(\mathbb{Q})$, then $G_x = G$, and it follows from Corollary 1.35 that (V, h_x) is the Betti realization of an abelian motive. The real approximation theorem states that $G(\mathbb{Q})$ is dense in $G(\mathbb{R})$, and so we can assume that $h_x = \mathrm{ad}\, g \circ h_0$ with $g \in G_0(\mathbb{R})^+$.

By assumption, (V, h_0) lies in the tensor category generated by the Betti realization of an abelian variety A. Let (G_1, h_1) be the Mumford-Tate group of A. We have a diagram

$$(G_0, h_0) \twoheadleftarrow (G_1, h_1) \hookrightarrow (G(\psi), X(\psi))$$

with $G(\psi)$ the group of symplectic similitudes defined by $H_1(A)$ and a Riemann form. Lift g to an element $g_1 \in G_1(\mathbb{R})^+$. Then we have a diagram

$$(G_0, h_x) \twoheadleftarrow (G_1, \mathrm{ad}(g) \circ h_1) \hookrightarrow (G(\psi), X(\psi)).$$

If G_{x1} denotes the inverse image of G_x in G_1, then $(G_{x1}, \mathrm{ad}(g) \circ h_1)$ is the Mumford-Tate group of an abelian variety, and it has (G_x, h_x) as a quotient. It follows that (V, h_x) is the Betti realization of an abelian motive. □

REMARK 2.27. Assume Hypothesis 1.13. If (V, h_0) is special, then for all $x \in X$, the Hodge structure (V, h_x) is the Betti realization of a special motive.

Motivic variations of Hodge structures. By a *global tensor* of a local system of $\mathbb{Q}$-vector spaces $\mathbb{V}$ on a complex manifold S, we mean an element of $\Gamma(S, \mathbb{V}^{\otimes r} \otimes \mathbb{V}^{\vee \otimes s} \otimes \mathbb{Q}(m))$ for some $r, s \in \mathbb{N}$, $m \in \mathbb{Z}$.

DEFINITION 2.28. Let $\pi: Y \to S$ be a projective smooth morphism of smooth varieties over $\mathbb{C}$. A global tensor t of $\mathscr{H}_B(Y/S) \underset{\mathrm{df}}{=} \bigoplus R^i\pi_*\mathbb{Q}$ is an *absolute Hodge tensor* of Y/S if, for all $s \in S(\mathbb{C})$, t_s is an absolute Hodge tensor of Y_s. A sum of absolute Hodge tensors will also be called an absolute Hodge tensor.

REMARK 2.29. When S is connected, a global tensor t of $\mathscr{H}_B(Y/S)$ is an absolute Hodge tensor of Y/S if t_s is an absolute Hodge tensor on Y_s for a single s [**19**, 2.12, 2.14].

DEFINITION 2.30. Let S be a smooth variety over $\mathbb{C}$.

(a) (Deligne 1971a, 4.2.4). A variation of Hodge structures $\mathbb{V}$ on S is *algebraic* if there exists a dense open subset U of S, an integer m, and a projective smooth morphism $\pi: Y \to U$ such that $\mathbb{V}|U$ is a direct summand of $\mathscr{H}_B(Y/U)(m)$.

(b) If in (a) the projector realizing $\mathbb{V}|U$ as a direct summand is an absolute Hodge tensor, then $\mathbb{V}$ will be said to be *motivic.*

(c) If in (a) Y is an abelian scheme over U, then $\mathbb{V}$ will be said to be *abelian-motivic.*

Note that the projector implicit in (a) of the definition is automatically a Hodge tensor at every point of U, and so, when Y is an abelian scheme, the main theorem of **[19]** implies that it is an absolute Hodge tensor. Thus

$$\text{abelian-motivic} \Rightarrow \text{motivic} \Rightarrow \text{algebraic}.$$

In general, when $\pi: Y \to S$ is a projective smooth morphism and $p \in \mathrm{End}(\mathscr{H}_B(Y/S))$ is both a projector and an absolute Hodge tensor, we write $\mathscr{H}_B(Y/S, p, m)$ for $\mathrm{Im}(p) \otimes \mathbb{Q}(m)$.

LEMMA 2.31. (a) *A direct sum, or tensor product of algebraic (resp. motivic, abelian-motivic) variations of Hodge structures is algebraic (resp. motivic, abelian-motivic); the constant variation of Hodge structures* $\mathbb{Q}(m)$ *is abelian-motivic.*

(b) *A direct summand of an algebraic (resp. abelian-motivic) variation of Hodge structures is algebraic (resp. abelian-motivic).*

(c) *Every algebraic variation of Hodge structures is polarizable.*

PROOF. The statements concerning algebraic variations are proved in **[14,** 4.2.5]. Similar proofs give the statements concerning abelian-motivic or motivic variations. □

PROPOSITION 2.32. *The category of abelian-motivic variations of Hodge structures on* S *is a semisimple abelian category; it is a tensor subcategory of the semisimple Tannakian category of polarizable variations of Hodge structures on* S. *If* S *is connected, then for any* $o \in S(\mathbb{C})$, $\mathbb{V} \mapsto \mathbb{V}_o$ *is a fibre functor.*

PROOF. We can again apply **[14,** 4.2.3]. □

LEMMA 2.33. *A variation of Hodge structures* $\mathbb{V}$ *on* S *is abelian-motivic if there is a smooth dominant morphism* $f: S' \to S$ *of finite-type such that* $f^*\mathbb{V}$ *is abelian-motivic.*

PROOF. The image of f is a dense open subscheme U of S, and there exists a surjective étale morphism $h: U' \to U$ and a morphism $g: U' \to S'$ such that $h = f \circ g$ (see Grothendieck **[37,** 17.16.3]. Hence $h^*\mathbb{V}$ is abelian-motivic, say $h^*\mathbb{V} = \mathscr{H}_B(Y/U', p, m)$. After replacing U with an open subset, the map $h: U' \to U$ will be finite. It is clear that $h_*h^*\mathbb{V}$ is abelian-motivic, and $\mathbb{V}|U$ is a subobject, hence direct summand, of it, and so we can apply Lemma 2.31(b). □

The motivicity of $\mathbb{V}(N)$.

THEOREM 2.34. *If* (V, h_0) *is the Betti realization of an abelian motive, then the variation of Hodge structures* $\mathbb{V}(N)$ *on* $S(N)$ *defined in* (2.11) *is abelian-motivic.*

PROOF. Suppose first that (V, h_0) is the Betti realization of an abelian variety. In this case $\mathbb{V}(N)$ is a polarizable variation of Hodge structures of type $\{(-1, 0), (0, -1)\}$, and so $\mathbb{V}(N) = \mathscr{H}_B(Y/S(N))$ for some abelian scheme Y (see [**14**, 4.4.3]. Moreover, if $G \to \mathrm{GL}(W)$ is a second representation of G and $\mathbb{W}$ is the corresponding variation of Hodge structures on $S(N)$, then, because W lies in the tensor category generated by V, $\mathbb{W}$ will lie in the tensor category generated by $\mathbb{V}$, and so (2.32) shows it to be abelian-motivic.

Now consider the general case. Because (V, h_0) is the Betti realization of an abelian motive, there is a surjective homomorphism $(G_1, u_1) \to (G, u_0)$ with (G_1, u_1) the special Mumford-Tate group of an abelian variety. Correspondingly, we have a smooth morphism

$$f\colon \Gamma_1(N)\backslash X_1 \to \Gamma(N)\backslash X.$$

The pull-back $f^*\mathbb{V}(N)$ of $\mathbb{V}(N)$ to $\Gamma_1(N)\backslash X_1$ is the variation of Hodge structures defined by the representation $G_1 \to G \hookrightarrow \mathrm{GL}(V)$, and is therefore abelian-motivic. Now apply Lemma 2.33. □

Special Hodge structures. Hypothesis 1.13 asserts that every special Hodge structure is of the form $H_B(Y, p, m)$ for some projective smooth variety Y over $\mathbb{C}$, projector p that is an absolute Hodge tensor, and integer m. The next hypothesis asserts that this holds in families.

HYPOTHESIS 2.35. *Let (V, h_0) be a special Hodge structure, and let o be the point $\Gamma(N)\cdot h_0$ in $S(N)$ (N sufficiently divisible). There exists an open neighbourhood U of o, a projective smooth morphism $\pi\colon Y \to U$, a projector p that is an absolute Hodge tensor, and an integer m such that $\mathbb{V}|U = \mathscr{H}_B(Y/U, p, m)$.*

Let M be a motive over $\mathbb{C}$. We say that *all the Hodge tensors of M are absolutely Hodge* if the functor

$$\omega_B\colon \mathbf{Mot}(\mathbb{C}) \longrightarrow \mathbf{Hdg}_{\mathbb{Q}}$$

becomes fully faithful when restricted to the tensor subcategory generated by M and $\mathbb{Q}(1)$.

PROPOSITION 2.36. *Let (V, h_0) be special. If Hypothesis 2.35 holds for (V, h_0), then (V, h_0) is the Betti realization of a motive M, and all Hodge tensors on M are absolutely Hodge.*

PROOF. A Hodge tensor t of (V, h_0) defines a global Hodge tensor of $\mathbb{V}(N)$. Let U be a neighbourhood of o as in Hypothesis 2.35. There will be a point $x \in U$ such that $\mathrm{Im}(h_x) \subset T_{\mathbb{R}}$ for some torus $T \subset G$, and the pair (V, h_x) will be the Betti realization of a CM-motive. Hence t_x is an absolute Hodge tensor, and, as we noted in Remark 2.29, this implies that t_u is an absolute Hodge tensor for all $u \in U$. In particular, $t = t_0$ is an absolute Hodge tensor. □

Hypothesis 2.35 is definitely a stronger statement than (1.13), but any proof of (1.13) is likely also to yield a proof of (2.35). Note that (2.35) implies that

$$\omega_B\colon \mathbf{Mot}^{\mathrm{sp}}(\mathbb{C}) \to \mathbf{Hdg}_{\mathbb{Q}}$$

is fully faithful where $\mathbf{Mot}^{\mathrm{sp}}(\mathbb{C})$ is the category of special motives. Hence every special Hodge structure will be the Betti realization of a *unique* special motive (unique up to a unique isomorphism).

Moduli of motives. We now define the category of motives for absolute Hodge tensors over any smooth variety S in characteristic zero. Our definition is suggested by the following theorem: let S be a smooth connected variety with generic point η over a field of characteristic zero; the functor $A \mapsto A_\eta$ from the category of abelian schemes over S to the category of abelian varieties over η is fully faithful, with essential image the category of abelian varieties B such that the action of $\pi_1(\eta, \bar{\eta})$ on $H^1(B_{\mathrm{et}}, \mathbb{Q}_\ell)$ factors through $\pi_1(S, \bar{\eta})$. Here $\bar{\eta}$ is the spectrum of an algebraically closed field containing $\kappa(\eta)$. (This theorem is a consequence of the theory of Néron models and of a theorem of Chai and Faltings—see Milne [**50,** 2.13], for a discussion of it.)

Definition 2.37. A *motive* M over a smooth connected k-variety S is a motive M_η over the generic point η of S such that the action of $\pi_1(\eta, \bar{\eta})$ on $\omega_f(M_\eta)$ factors through $\pi_1(S, \bar{\eta})$. If M_η is an abelian motive, then we call M an *abelian motive* over S. (Here ω_f is the fibre functor, defined by étale cohomology, taking values in $\mathbb{A}_f$-modules.)

Let M be a motive over S. For some m, $M_\eta(-m)$ will be an effective motive and, hence, a direct summand of a motive $h(Y)$ where Y is a smooth projective variety over the field $\kappa(\eta)$. Let p be the absolute Hodge tensor for Y projecting $h(Y)(m)$ onto M_η. For some open subset U of S, Y will extend to a smooth projective scheme Y_U over U, and p will extend to a global tensor p_U for the de Rham and étale cohomologies of Y_U/U. We then say that (Y_U, p_U, m) represents M over U, and we write $M|U = h(Y_U, p_U, m)$.

Proposition 2.38. *For any smooth connected variety S over a field k of characteristic zero, the category* $\mathbf{Mot}(S)$ *of motives over S is a Tannakian category over $\mathbb{Q}$, and the category* $\mathbf{Mot}^{\mathrm{ab}}(S)$ *of abelian motives over S is a Tannakian subcategory of* $\mathbf{Mot}(S)$. *There is an exact tensor functor from* $\mathbf{Mot}(S)$ *to the Tannakian category of local systems of $\mathbb{A}_f$-modules on $S_{\text{ét}}$.*

Proof. Since $\mathbf{Mot}(\eta)$ is a Tannakian category, it suffices to prove that $\mathbf{Mot}(S)$ is a Tannakian subcategory of $\mathbf{Mot}(\eta)$, but this is obvious. Similarly $\mathbf{Mot}^{\mathrm{ab}}(S)$ is a Tannakian subcategory of $\mathbf{Mot}(S)$ and of $\mathbf{Mot}^{\mathrm{ab}}(\eta)$. To give a local system of $\mathbb{A}_f$-modules on S is the same as to give a continuous representation of $\pi_1(S, \bar{\eta})$ on a finite-dimensional $\mathbb{A}_f$-module, and, by assumption, the representation of $\pi_1(\eta, \bar{\eta})$ on $\omega_f(M_\eta)$ factors through $\pi_1(S, \bar{\eta})$. □

DEFINITION 2.39. An *integral structure* on a motive M is the choice of a local system of torsion-free $\hat{\mathbb{Z}}$-modules $M(\hat{\mathbb{Z}})$ such that $M(\hat{\mathbb{Z}}) \otimes \mathbb{Q} = M_f$. A motive together with an integral structure is an *integral motive.*

As we have just seen, almost by definition, a motive over a smooth variety S defines a local system of $\mathbb{A}_f$-modules. Less obvious is that, when the ground field is $\mathbb{C}$, a motive defines a polarizable variation of Hodge structures on S.

THEOREM 2.40. *Let M be a motive over a smooth algebraic variety S over $\mathbb{C}$. There exists a unique polarizable variation of Hodge structures $\mathscr{H}_B(M)$ on S with the following property: if (Y, p, m) represents M over the open subset U of S, then $\mathscr{H}_B(M) = \mathscr{H}_B(Y/U, p, m)$.*

PROOF. Let (Y, p, m) represent M over U, and let $\mathbb{V} = \mathscr{H}_B(Y/U)$. Choose a point $u \in U$. The action of $\pi_1(U, u)$ on $\mathbb{V}_u$ factors through $\pi_1(S, u)$, and so $\mathbb{V}$ extends (uniquely) to a local system of $\mathbb{Q}$-vector spaces on S, which we still denote $\mathbb{V}$.

Thus we have a local system of $\mathbb{Q}$-vector spaces $\mathbb{V}$ on S, and the structure of polarizable variation of Hodge structures on $\mathbb{V}|U$. The next lemma shows that this structure extends uniquely to $\mathbb{V}$, which completes the proof.

LEMMA 2.41. *Let S be a smooth algebraic variety over $\mathbb{C}$. Let $\mathbb{V}$ be a local system of $\mathbb{Q}$-vector spaces on S, and let ψ be a bilinear form on $\mathbb{V}$. Suppose that there is a Zariski open subset U of S and a filtration F on $\mathscr{O}_U \otimes \mathbb{V}$ such that $(\mathbb{V}|U, F, \psi|U)$ is a polarized variation of Hodge structures on U of some weight m. Then F extends uniquely to a filtration on $\mathbb{V} \otimes \mathscr{O}_V$ such that $(\mathbb{V}\ , F, \psi)$ is a polarized variation of Hodge structures on S.*

PROOF. There exists a Zariski open subvariety S' of S containing U and such that $S - S'$ has codimension ≥ 2 and $S' - U$ is smooth of pure codimension 1 (i.e., a smooth divisor). Thus it suffices to consider two cases: $S - U$ is a smooth divisor; $S - U$ has codimension 2.

In the first case, the Hodge structure on $\mathbb{V}|U$ will in general degenerate into a mixed Hodge structure on the boundary, but the description of the weight filtration in terms of the action of the local monodromy group shows that it is trivial (i.e., the mixed Hodge structure is pure) when the local monodromy group acts trivially. See Schmid [**64,** 4.12] and Cattani et al [**13**].

For the second case, let D be the classifying space for polarized Hodge structures of the same type as (V, h_u, ψ), $u \in U$. Then $\mathbb{V}|U$ defines a horizontal, locally liftable holomorphic mapping $U \to \Gamma\backslash D$, which Griffiths and Schmid [**36,** 9.8] show extends to all of S (because D is "negatively curved in the horizontal direction"). From the extended map we obtain an extension of the variation of Hodge structures to S. □

PROPOSITION 2.42. *Let S be a smooth connected scheme over $\mathbb{C}$. The functor $\mathscr{H}_B$: $\mathbf{Mot}^{\mathrm{ab}}(S) \to \mathbf{Hdg}_{\mathbb{Q}}(S)$ defined in Lemma 2.41 is fully faithful with* *essential image the category of abelian-motivic variations of Hodge structures on S.*

PROOF. Recall the following result [**14**, 4.4.3]: there is an equivalence between the category of abelian schemes on S and the category of torsion-free integral polarizable variations of Hodge structures on S of type $\{(-1, 0), (0, -1)\}$. When we apply this to an open subvariety U of S, we find that there is an equivalence between the category of abelian motives on S whose restriction to U can be realized as a Tate twist of a factor of the motive of an abelian scheme over U and the category of abelian-motivic variations of Hodge structures on S whose restriction to U can be represented in the form $\mathscr{H}_B(Y/U, p, m)$ with Y an abelian scheme over U. Now take the union over all U. □

Let T be a smooth variety over $\mathbb{C}$, and consider triples $(M, \mathfrak{s}, \eta)$ consisting of an abelian motive M over T, a family $\mathfrak{s}$ of Hodge tensors of M indexed by I, and a level-N structure on $\mathscr{H}_B(M)$. We define $\mathscr{M}(T)$ to be the set of triples $(M, \mathfrak{s}, \eta)$ such that $(\mathscr{H}_B(M), \mathfrak{s}, \eta) \in \mathscr{H}(T)$. With the obvious notion of pull-back and isomorphism, we obtain a moduli problem $(\mathscr{M}, \approx)$ on the category of smooth varieties over $\mathbb{C}$.

Now assume that (V, h_0) is the Betti realization of an abelian motive. It follows from Propositions 2.26 and 2.8 that the elements of $\mathscr{H}(\mathbb{C})$ are also Betti realizations of abelian motives and, hence, that H_B defines a bijection

$$H_B : \mathscr{M}(\mathbb{C})/\!\approx\ \to \mathscr{H}(\mathbb{C})/\!\approx .$$

THEOREM 2.43. *If (V, h_0) is the Betti realization of an abelian motive, then the pair $(S(N), \alpha\circ H_B)$ is a fine solution to the moduli problem $(\mathscr{M}, \approx)$.*

PROOF. Let T be a smooth variety over $\mathbb{C}$. It follows from Proposition 2.42 that

$$\mathscr{H}_B : \mathscr{M}(T)/\!\approx\ \to \mathscr{H}(T)/\!\approx$$

is injective, and from (2.34), (2.12), and (2.42) that it is surjective. Hence the theorem follows from (2.12). □

REMARK 2.44. Theorem 2.43 realizes a vast family of varieties as moduli varieties. Except for the moduli varieties of abelian varieties, the only example I know of where this has been exploited is that in which the initial Hodge structure is the second cohomology group of a K3-surface (see for example [**17**]).

REMARK 2.45. The moduli problems for Hodge structures and motives have been defined above only for the category of smooth varieties over $\mathbb{C}$ (see Definitions 2.15 and 2.37). Ching-Li Chai has suggested to me that they can be defined for all schemes over $\mathbb{C}$ by replacing the connections with Grothendieck's notion of a stratification [**38**]. A *stratification* on an $\mathscr{O}_S$-module $\mathscr{V}$ is an isomorphism

$$\varphi : p_1^*\mathscr{V} \to p_2^*\mathscr{V},$$

where p_1, p_2 are the projections $\widehat{\Delta} \rightrightarrows S$ from the formal completion of the diagonal in $S \times S$ to S, satisfying the cocycle condition $p_{31}^*(\varphi) = p_{32}^*(\varphi)p_{21}^*(\varphi)$. Each of $p_1^*\mathscr{V}$ and $p_2^*\mathscr{V}$ has the product filtration, that on $\mathscr{O}_{\widehat{\Delta}}$ being given by the defining ideal and its powers, and in this context Griffiths transversality (2.15.2) becomes the condition that φ preserves the filtrations. It should be possible to prove that $S(N)$ solves the moduli problems on the category of all schemes by using the methods of Artin. A similar remark applies to Theorem 3.31 below.

3. Shimura varieties as moduli varieties

In the last two sections we saw that the study of Mumford-Tate groups and the moduli varieties of motives leads to the consideration of pairs (G, h) satisfying certain conditions (SV). In this section, we reverse the process: starting with a reductive group G over $\mathbb{Q}$ and a $G(\mathbb{R})$-conjugacy class X of homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$ satisfying conditions (SV), we construct a pro-variety $\mathrm{Sh}(G, X)$, the Shimura variety defined by (G, X), and show that in many cases $\mathrm{Sh}(G, X)$ can be realized as a moduli variety for motives over a number field. Then we give some applications of this result.

Review of Shimura varieties. Let G be a connected reductive group over $\mathbb{Q}$, and let X be a $G(\mathbb{R})$-conjugacy class of homomorphisms $h: \mathbb{S} \to G_{\mathbb{R}}$ satisfying the conditions (SV1,2). Note that it suffices to verify the conditions for a single h, because the map $\mathrm{ad}(g): G_{\mathbb{R}} \to G_{\mathbb{R}}$ is an isomorphism carrying h into $\mathrm{ad}(g) \circ h$. The condition (SV1) implies that the image of w_h is contained in $Z(G)$ and, hence, is independent of $h \in X$—we denote it by w_X and call it the *weight*. It is convenient to impose the following condition on the group G:

(SV0) the torus $Z(G)^0$ splits over a CM-field, and G^{ad} has no $\mathbb{Q}$-factor that is anisotropic over $\mathbb{R}$.

Recall that a *CM-field* is a finite extension E of $\mathbb{Q}$ admitting a nontrivial involution ι_E such that $\rho(\iota_E z) = \iota\rho(z)$ for all embeddings $\rho: E \hookrightarrow \mathbb{C}$. If $X^*(Z^0)$ denotes the group of characters of Z^0 defined over the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$, then Z^0 splits over a CM-field if and only if, for all $\tau \in \mathrm{Gal}(\mathbb{Q}^{\mathrm{al}}/\mathbb{Q})$, $\iota\tau$ and $\tau\iota$ have the same action on $X^*(Z^0)$. Since Z^0 and G^{ab} are isogenous tori, the condition is equivalent to G^{ab} splitting over a CM-field, and in the presence of (SV2), it implies that G itself splits over a CM-field.

The condition on G^{ad} implies that the strong approximation theorem holds in the following form: the group $G^{\mathrm{sc}}(\mathbb{Q})$ is dense in $G^{\mathrm{sc}}(\mathbb{A}_f)$. This simplifies the theory, but unfortunately eliminates zero-dimensional Shimura varieties except for those defined by tori.

Fix a pair (G, X) satisfying the conditions (SV0,1,2). Then X has a unique complex structure for which the Hodge filtrations F_h on $V \otimes \mathbb{C}$ vary holomorphically [19, 1.1.14]; cf. §2. Moreover, X has only finitely many connected components, and each is a symmetric Hermitian domain [18, 1.1.17]; cf. 2.23. As before, I write x for an element of X and h_x, μ_x for the corresponding homomorphisms.

If K is a compact open subgroup of $G(\mathbb{A}_f)$, then $\Gamma(K) \underset{\text{df}}{=} G(\mathbb{Q})\cap K$ is (by definition) a congruence subgroup of $G(\mathbb{Q})$, and hence its image $\Gamma^{\text{ad}}(K)$ in $G^{\text{ad}}(\mathbb{Q})$ is an arithmetic group. For K sufficiently small, $\Gamma(K)$ is contained in $G(\mathbb{R})^+$ [18, 2.0.14] and $\Gamma^{\text{ad}}(K)$ is torsion-free (cf. 2.9).

For K a compact open subgroup of $G(\mathbb{A}_f)$, define

$$\text{Sh}_K(G, X) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K,$$

where $G(\mathbb{Q})$ and K act on $X \times G(\mathbb{A}_f)$ according to the rule

$$q(x, a)k = (qx, qak), \qquad q \in G(\mathbb{Q}),\ x \in X,\ a \in G(\mathbb{A}_f),\ k \in K.$$

Let $G(\mathbb{Q})_+$ be the subgroup of $G(\mathbb{Q})$ of elements mapping into $G^{\text{ad}}(\mathbb{R})^+$; it is the stabilizer in $G(\mathbb{Q})$ of any connected component X^+ of X [18, 1.2.7]. Let $\mathscr{C}$ be a set of representatives for $G(\mathbb{Q})_+\backslash G(\mathbb{A}_f)/K$—the strong approximation theorem implies that it is finite. For K sufficiently small, the map

$$\coprod_{c\in\mathscr{C}} \Gamma^{\text{ad}}(cKc^{-1})\backslash X^+ \longrightarrow G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K$$

sending an element $[x] \in \Gamma^{\text{ad}}(cKc^{-1})\backslash X^+$ to $[x, c]$ is a homeomorphism. Therefore (see Lemma 2.24), $\text{Sh}_K(G, X)$ has a unique algebraic structure compatible with its complex structure; moreover, for any smooth algebraic variety T over $\mathbb{C}$, every holomorphic map $T(\mathbb{C}) \to \text{Sh}_K(G, X)$ is algebraic.

From now on, we regard $\text{Sh}_K(G, X)$ as an algebraic variety.

When we vary K among (small) compact open subgroups of $G(\mathbb{A}_f)$, we obtain a filtered projective system of algebraic varieties

$$\text{Sh}(G, X) = (\text{Sh}_K(G, X))_K.$$

The group $G(\mathbb{A}_f)$ acts continuously on this system in the sense of [18, 2.7.1], and $\text{Sh}(G, X)$, together with this action, is called the *Shimura variety* defined by (G, X).

Alternatively (and equivalently), we can define $\text{Sh}(G, X)$ to be the projective limit,

$$\text{Sh}(G, X) = \varprojlim \text{Sh}_K(G, X) \quad \text{(scheme, not of finite type, over } \mathbb{C}\text{)},$$

together with the action of $G(\mathbb{A}_f)$. The variety $\text{Sh}_K(G, X)$ can be recovered as the quotient $\text{Sh}(G, X)/K$ of $\text{Sh}(G, X)$.

Recall ([18, 2.1.10]; see also Proposition 4.11 below) that

$$\text{Sh}(G, X)(\mathbb{C}) \underset{\text{df}}{=} \varprojlim \text{Sh}_K(G, X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/Z(\mathbb{Q})^-$$

where $Z = Z(G)$ and $Z(\mathbb{Q})^-$ is the closure of $Z(\mathbb{Q})$ in $Z(\mathbb{A}_f)$. An element $g \in G(\mathbb{A}_f)$ acts on $\mathrm{Sh}(G, X)(\mathbb{C})$ as follows:

$$[x, a]g = [x, ag], \qquad x \in X, \ a \in G(\mathbb{A}_f).$$

Because Z^0 splits over a CM-field, the largest split subtorus of $Z_{\mathbb{R}}$ is defined over $\mathbb{Q}$. When this subtorus is also split over $\mathbb{Q}$, $Z(\mathbb{Q})$ is closed in $Z(\mathbb{A}_f)$ and we have

$$\mathrm{Sh}(G, X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f).$$

EXAMPLE 3.1. Let (G, h) be the Mumford-Tate group of a special Hodge structure (V, h). We saw in Proposition 1.6 that h satisfies (SV2*) and, *a fortiori*, (SV2) and that the weight w_h is defined over $\mathbb{Q}$. If H is a factor of G^{ad} such that $H_{\mathbb{R}}$ is anisotropic, then (SV2) implies that the composite $\mathbb{S} \xrightarrow{h} G_{\mathbb{R}} \longrightarrow H_{\mathbb{R}}$ is trivial, and since h generates H, H itself must be trivial. The pair $(G^{\mathrm{ab}}, h^{\mathrm{ab}})$ is the Mumford-Tate group of a Hodge structure of CM-type, and therefore G^{ab} splits over a CM-field. Thus G satisfies (SV0), and, by assumption, (G, h) satisfies (SV1). Therefore, when we define X to be the $G(\mathbb{R})$-conjugacy class of h, (G, X) satisfies the conditions to define a Shimura variety. Choose a lattice $V(\mathbb{Z})$ in V, and let

$$K(N) = \{g \in G(\mathbb{A}_f) \mid gV(\widehat{\mathbb{Z}}) = V(\widehat{\mathbb{Z}}), \quad g = \mathrm{id} \text{ on } V(\widehat{\mathbb{Z}})/NV(\widehat{\mathbb{Z}})\}.$$

Then the variety $S(N)$ attached to (V, h) in §2 is an open and closed subvariety of $\mathrm{Sh}_{K(N)}(G, X)$.

DEFINITION 3.2. If G^{der} satisfies the condition (1.25), then the Shimura variety $\mathrm{Sh}(G, X)$ is said to be of *abelian type.*[5]

REMARK 3.3. Let (G, h) be the Mumford-Tate group of an abelian motive, and let X be the $G(\mathbb{R})$-conjugacy class of h; then $\mathrm{Sh}(G, X)$ is a Shimura variety of abelian type, and w_X is defined over $\mathbb{Q}$. Conversely, let $\mathrm{Sh}(G, X)$ be a Shimura variety of abelian type whose weight is defined over $\mathbb{Q}$; if G' is the subgroup of G generated by the elements of X, then there is an $h \in X$ such that (G', h) is the Mumford-Tate group of an abelian motive.

Canonical models. We recall the notion of the canonical model of a Shimura variety. Let T be a torus over $\mathbb{Q}$ that splits over a CM-field, and let $\mu \in X_*(T)$ (group of cocharacters defined over $\mathbb{Q}^{\mathrm{al}} \subset \mathbb{C}$). The pair (T, h), $h(z) = \mu(z) \cdot \overline{\mu(z)}$, defines a Shimura variety. Let $E = E(T, h) \subset \mathbb{Q}^{\mathrm{al}}$ be the field of definition of μ, let E^{ab} be the maximal abelian extension of

[5]This is precisely the class of Shimura varieties for which Deligne proved the existence of canonical models in his Corvallis article [**18**, 2.7.21]. The name was coined by Shih and the author [**53**] in 1982 because, at the time, they seemed to be exactly the Shimura varieties that were approachable by methods involving the moduli of abelian varieties. Below we shall see a much more compelling justification for the name: among the Shimura varieties whose weight is defined over $\mathbb{Q}$, they are the varieties that are moduli varieties for abelian motives.

E (in $\mathbb{Q}^{\mathrm{al}}$), and let rec_E be the Artin reciprocity map $\mathbb{A}_E^\times \to \mathrm{Gal}(E^{\mathrm{ab}}/E)$. On applying the restriction-of-scalars functor $\mathrm{Res}_{E/\mathbb{Q}}$ to the homomorphism $\mu\colon \mathbb{G}_{mE} \to T_E$ and composing with the norm map, we obtain a homomorphism

$$N_h\colon \mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m \xrightarrow{\mathrm{Res}(\mu)} \mathrm{Res}_{E/\mathbb{Q}} T_E \xrightarrow{\mathrm{Nm}} T.$$

For any $\mathbb{Q}$-algebra R, this gives a homomorphism

$$N_h(R)\colon (E \otimes R)^\times \longrightarrow T(R).$$

Let $T(\mathbb{Q})^-$ be the closure of $T(\mathbb{Q})$ in $T(\mathbb{A}_f)$. The *reciprocity* map[6]

$$r(T, h)\colon \mathrm{Gal}(E^{\mathrm{ab}}/E) \longrightarrow T(\mathbb{A}_f)/T(\mathbb{Q})^-$$

is defined as follows: let $\tau \in \mathrm{Gal}(E^{\mathrm{ab}}/E)$, and let $s \in \mathbb{A}_E^\times$ be such that $\mathrm{rec}_E(s) = \tau$; write $s = s_\infty \cdot s_f$ with $s_\infty \in (E \otimes \mathbb{R})^\times$ and $s_f \in (E \otimes \mathbb{A}_f)^\times$; then $r(T, x)(\tau) = N_h(s_f) \pmod{T(\mathbb{Q})^-}$.

Now consider a Shimura variety $\mathrm{Sh}(G, X)$. The *reflex field* $E(G, X)$ of $\mathrm{Sh}(G, X)$ is the field of definition (in $\mathbb{C}$) of the $G(\mathbb{C})$-conjugacy class of homomorphisms $\mu\colon \mathbb{G}_m \to G_{\mathbb{C}}$ containing μ_h, $h \in X$. A *special pair* (T, h) in (G, X) is a torus $T \subset G$ together with an $h \in X$ such that h factors through $T_{\mathbb{R}}$. Clearly $E(T, h) \supset E(G, X)$.

By a *model* of $\mathrm{Sh}(G, X)$ over a subfield k of $\mathbb{C}$, we mean a scheme S over k endowed with an action of $G(\mathbb{A}_f)$ (defined over k) and a $G(\mathbb{A}_f)$-equivariant isomorphism $\mathrm{Sh}(G, X) \to S \otimes_k \mathbb{C}$. We use this isomorphism to identify $\mathrm{Sh}(G, X)(\mathbb{C})$ with $S(\mathbb{C})$.

Definition 3.4. A model of $\mathrm{Sh}(G, X)$ over a number field E, $\mathbb{C} \supset E \supset E(G, X)$, is said to be *canonical* if it has the following property: for all *special pairs $(T, h) \subset (G, X)$ and elements $a \in G(\mathbb{A}_f)$, the point $[h, a]$ is* rational over $E(h)^{\mathrm{ab}}$ and $\tau \in \mathrm{Gal}(E(h)^{\mathrm{ab}}/E(h))$ acts on $[h, a]$ according to the rule:

$$\tau[h, a] = [h, r(\tau) \cdot a] \quad \text{where } r = r(T, h); \tag{3.4.1}$$

here $E(h) = E \cdot E(T, h)$.

Proposition 3.5. *Let $f\colon G \to G'$ be a homomorphism mapping X into X', and suppose that* $\mathrm{Sh}(G, X)$ *and* $\mathrm{Sh}(G', X')$ *have canonical models over E. Then the morphism*

$$[x, g] \mapsto [f(x), f(g)]\colon \mathrm{Sh}(G, X) \to \mathrm{Sh}(G', X')$$

is defined over E.

Proof. See **[15**, 5.4]. □

[6]For an explanation of the sign, which differs from that in **[18]**, see **[50**, 1.10.]

COROLLARY 3.6. *If it exists, the canonical model of* Sh(G, X) *over* E *is uniquely determined up to a unique isomorphism.*

PROOF. Apply the proposition with f the identity map $G \to G$. □

In his report on Shimura's work, Deligne [15] proves that Shimura varieties that are moduli varieties for abelian varieties have canonical models over their reflex fields, and he deduces a similar result for one class of Shimura varieties whose members are not moduli varieties and, in fact, do not have weight defined over $\mathbb{Q}$ ("les modèles étranges", [15, §6]). In the next subsection, we prove that Shimura varieties of abelian type with rational weight are moduli varieties for abelian *motives*. This allows us in the following subsection to prove, using the methods of Deligne's article, that all Shimura varieties of abelian type have canonical models over their reflex fields.[7]

Shimura varieties as moduli varieties over $\mathbb{C}$. Throughout this subsection, $\mathrm{Sh}(G, X)$ is a Shimura variety whose weight w_X is defined over $\mathbb{Q}$. For simplicity, we assume that there is given a homomorphism $t\colon G \to \mathbb{G}_m = \mathrm{GL}(\mathbb{Q}(1))$ such that $t \circ w_X = -2$. Then $t_{\mathbb{R}} \circ h_x$ defines on $\mathbb{Q}(1)$ its usual Hodge structure for all $x \in X$.

The realization of $\mathrm{Sh}(G, X)$ as a moduli variety depends on the choice of a faithful representation $\xi\colon G \hookrightarrow \mathrm{GL}(V)$ of G. We fix such a ξ and identify G with a subgroup of $\mathrm{GL}(V)$. There will be a family of tensors $\mathfrak{t} = (t_i)_{i \in I}$, $t_i \in V^{\otimes r_i} \otimes V^{\vee \otimes s_i}$ some r_i, s_i, such that, for any $\mathbb{Q}$-algebra R,

$$G(R) = \{g \in \mathrm{GL}(V \otimes R) \mid g t_i = t_i,\ \text{all } i \in I\}.$$

For all $x \in X$, t_i will be fixed by $h_x(\mathbb{S})$, and so t_i is a Hodge tensor for the rational Hodge structure (V, h_x). For some r and s, the representation t of G on $\mathbb{Q}(1)$ will be a direct summand of the representation of G on $V^{\otimes r} \otimes V^{\vee \otimes s}$ defined by ξ. Thus it makes sense to add the requirement that there is an element $0 \in I$ such that t_0 or $-t_0$ is a polarization for (V, h_x), all $x \in X$.

Fix a (small) compact open subgroup K of $G(\mathbb{A}_f)$, and let (W, h) be a rational Hodge structure. Then K acts on the space of $\mathbb{A}_f$-linear isomorphisms $V(\mathbb{A}_f) \to W(\mathbb{A}_f)$ on the right, and an orbit for the action is called a *K-level structure* on (W, h).

EXAMPLE 3.7. Choose lattices $V(\mathbb{Z})$ and $W(\mathbb{Z})$ in V and W, and define $K(N)$ as in Example 3.1. To give a $K(N)$-level structure on W is the same as to give a level N-structure in the sense of §2.

Consider triples $(W, \mathfrak{s}, [\eta])$ consisting of a rational Hodge structure $W = (W, h)$, a family $\mathfrak{s}$ of Hodge cycles indexed by I, and a K-level structure $[\eta]$ on (W, h). We define $\mathscr{H}_K(G, X, \xi)$ to be the set of such triples satis-

[7] As noted in a previous footnote, the existence of canonical models for Shimura varieties of abelian type was proved in [18], but by less explicit methods involving connected Shimura varieties. The result was extended to all Shimura varieties in Milne [47] and in Borovoĭ [7, 8].

fying the following conditions:

(3.8.1) there exists an isomorphism of $\mathbb{Q}$-vector spaces $\beta: W \to V$ mapping each s_i to t_i and sending h to h_x, some $x \in X$;

(3.8.2) for one (hence every) η representing the level structure, η maps each t_i to s_i.

An isomorphism from one such triple $(W, \mathfrak{s}, [\eta])$ to a second $(W', \mathfrak{s}', [\eta'])$ is an isomorphism $\gamma: (W, h) \to (W', h')$ of rational Hodge structures mapping each s_i to s_i' and such that $[\alpha \circ \eta] = [\eta']$. An element $g \in G(\mathbb{A}_f)$ defines a map

$$(W, \mathfrak{s}, [\eta]) \mapsto (W, \mathfrak{s}, [\eta \circ g]): \mathscr{H}_K(G, X, \xi) \to \mathscr{H}_{g^{-1}Kg}(G, X, \xi).$$

Let $(W, \mathfrak{s}, [\eta])$ be an element of $\mathscr{H}_K(G, X, \xi)$. Choose an isomorphism $\beta: W \to V$ satisfying (3.8.1), so that β sends h to h_x some $x \in X$. The composite

$$V(\mathbb{A}_f) \xrightarrow{\eta} W(\mathbb{A}_f) \xrightarrow{\beta} V(\mathbb{A}_f), \qquad \eta \in [\eta],$$

sends each t_i to t_i, and it is therefore multiplication by an element $g \in G(\mathbb{A}_f)$, well defined up to multiplication on the right by an element of K (corresponding to a different choice of the representative η of the level structure). Since any other choice of β is of the form $q \circ \beta$ for some $q \in G(\mathbb{Q})$, $[x, g]$ is a well-defined element of $G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K = \mathrm{Sh}_K(G, X)(\mathbb{C})$.

PROPOSITION 3.9. *The above construction defines a bijection*

$$\alpha_K: \mathscr{H}_K(G, X, \xi)/\approx \;\to \mathrm{Sh}_K(G, X)(\mathbb{C}).$$

The maps α_K are compatible with the action of $G(\mathbb{A}_f)$, in the sense that, for every $g \in G(\mathbb{A}_f)$, there is a commutative diagram

$$\begin{array}{ccc} \mathscr{H}_K(G, X, \xi)/\approx & \xrightarrow{\alpha_K} & \mathrm{Sh}_K(G, X) \\ \downarrow g & & \downarrow g \\ \mathscr{H}_{g^{-1}Kg}(G, X, \xi)/\approx & \xrightarrow{\alpha_{g^{-1}Kg}} & \mathrm{Sh}_{g^{-1}Kg}(G, X). \end{array}$$

PROOF. Straightforward, and essentially the same as that of Proposition 2.8. □

We now fix (G, X) and ξ and drop them from the notation.

Let $\mathbb{V}$ be a variation of Hodge structures on a smooth algebraic variety T over $\mathbb{C}$. Because in our definition (2.15) we required $\mathbb{V}$ to admit an integal structure, there is a well-defined local system of $\mathbb{A}_f$-modules $\mathbb{W}$ for the étale topology on T such that, for any connected component T^0 of T and any $o \in T^0(\mathbb{C})$, $\mathbb{W}_o = \mathbb{V}_o \otimes_{\mathbb{Q}} \mathbb{A}_f$ as $\pi_1(T^0(\mathbb{C}), o)$-modules. We denote this étale sheaf by $\mathbb{V}(\mathbb{A}_f)$.

A *K-level structure* on a local system $\mathbb{W}(\mathbb{A}_f)$ of $\mathbb{A}_f$-modules on $T_{\text{ét}}$ is a K-equivalence class of isomorphisms $\eta: V(\mathbb{A}_f)_T \to \mathbb{W}(\mathbb{A}_f)$ on T. Here

$V(\mathbb{A}_f)_T$ is the constant local system defined by the $\mathbb{A}_f$-module $V(\mathbb{A}_f)$. The class $[\eta]$ is required to be defined on $T_{\text{ét}}$, not its individual members, which may only be defined on the universal covering of T. If T is connected and $o \in T$, then to give a K-level structure on $\mathbb{W}(\mathbb{A}_f)$ is the same as to give a K-level structure on $\mathbb{W}_o$ that is stable under the action of $\pi_1(T_{\text{ét}}, o)$ (algebraic fundamental group).

Consider triples $(\mathbb{W}, \mathfrak{s}, [\eta])$ consisting of a polarizable variation of Hodge structures $\mathbb{W}$ on T, a family of global Hodge tensors $\mathfrak{s}$ of $\mathbb{W}$ indexed by I, and a K-level structure $[\eta]$ on $\mathbb{W}(\mathbb{A}_f)$. We define $\mathscr{H}_K(T)$ to be the set of such triples having the property that, for all $t \in T$, $(\mathbb{V}_t, \mathfrak{s}_t, [\eta_t])$ lies in $\mathscr{H}_K(\mathbb{C})$. With the obvious notions of isomorphism and pull-back, $\mathscr{H}_K$ is a moduli problem on the category of smooth algebraic varieties over $\mathbb{C}$, and $\mathscr{H}_K(\text{point}) = \mathscr{H}_K(G, X, \xi)$.

PROPOSITION 3.10. *With the above notations,* $(\mathrm{Sh}_K(G, X), \alpha_K)$ *is a solution to the moduli problem* $(\mathscr{H}_K, \approx)$.

PROOF. Let $m \in \mathscr{H}_K(T)$. The same argument as in the proof of Theorem 2.12 shows that $m \mapsto \alpha(m_t)\colon T \to \mathrm{Sh}_K(G, X)(\mathbb{C})$ is holomorphic and, hence, is a morphism of algebraic varieties. As in Example 2.11, on each connected component $S_c = \Gamma^{\text{ad}}(cKc^{-1})\backslash X^+$ of $\mathrm{Sh}_K(G, X)(\mathbb{C})$, there is an $m_c \in \mathscr{H}_K(S_c)$ such that $\alpha((m_c)_s) = s$ for all $s \in S_c(\mathbb{C})$, and so $(\mathrm{Sh}_K(G, X), \alpha_K)$ satisfies the conditions (2.1). □

REMARK 3.11. If the largest $\mathbb{R}$-split subtorus of Z is already split over $\mathbb{Q}$, then $Z(\mathbb{Q})$ is closed in $Z(\mathbb{A}_f)$, and the moduli problem is fine. More precisely, there is an element $m_0 \in \mathscr{H}(\mathrm{Sh}_K(G, X))$ such that $(\mathrm{Sh}_K(G, X), [m_0])$ represents the functor $\mathscr{H}_K/\approx$.

The next proposition and theorem show that $\mathrm{Sh}(G, X)$ is a moduli variety for abelian motives if and only if it is of abelian type. (Recall that we are assuming that w_X is defined over $\mathbb{Q}$.)

PROPOSITION 3.12. *The elements of* $\mathscr{H}_K(\mathbb{C})$ *are the Betti realizations of abelian motives if and only if* $\mathrm{Sh}(G, X)$ *is of abelian type. When this is the case, then, for each connected component* S_c *of* $\mathrm{Sh}(G, X)$, *the element* $m_c \in \mathscr{H}_K(S_c)$ *defined in the proof of Proposition* 3.10 *is abelian-motivic.*

PROOF. Let G' be the $\mathbb{Q}$-subgroup of G generated by $\{h_x \mid x \in X\}$. The second part of condition (SV0) implies that $\{\mathrm{ad} \circ h_x \mid x \in X\}$ generates G^{ad}, and therefore $G'/G' \cap Z(G) = G^{\text{ad}}$. Hence G'^{der} is of finite index in G^{der}, and, being connected, the two groups are equal. Proposition 1.38 applied to the holomorphic family of Hodge structures (V, h_x) on a connected component X^+ of X shows that G' is the Mumford-Tate group of (V, h_o) for some $o \in X^+$. If (V, h_o) is the Betti realization of an abelian motive, then Theorem 1.27 shows that G'^{der} satisfies (1.25), and therefore that $\mathrm{Sh}(G, X)$ is of abelian type. Conversely, if $\mathrm{Sh}(G, X)$ is of abelian type, then Theorem

1.27 shows that (V, h_o) is the Betti realization of an abelian motive, and the same argument as in the proof of Proposition 2.26 then shows that the same is true of (V, h_x) for every $x \in X$. The proof of the last statement is the same as that of Theorem 2.34. □

For any smooth variety T over $\mathbb{C}$, define $\mathscr{M}_K(T)$ to be the set of triples $(M, \mathfrak{s}, [\eta])$ consisting of an abelian motive M over T, a family of tensors $\mathfrak{s}$ of M indexed by I, and a level K-structure $[\eta]$ on M, such that the Betti realization of the triple lies in $\mathscr{H}_K(T)$. With the obvious notions of pull-back and isomorphism, $\mathscr{M}_K$ becomes a moduli problem on the category of smooth algebraic varieties over $\mathbb{C}$.

THEOREM 3.13. *Let* $\mathrm{Sh}(G, X)$ *be a Shimura variety of abelian type whose weight is defined over* $\mathbb{Q}$ *and for which there exists a homomorphism* $t: G \to \mathbb{G}_m$ *such that* $t \circ w_X = -2$. *For any representation* $\xi: G \hookrightarrow \mathrm{GL}(V)$ *possessing a fixed tensor* t_0 *such that* $\pm t_0$ *is a polarization of* $(V, \xi_{\mathbb{R}} \circ h_x)$ *for all* $x \in X$, $(\mathrm{Sh}_K(G, X), \alpha_K)$ *is a solution of the moduli problem* $(\mathscr{M}_K, \approx)$. *When* $Z(\mathbb{Q})$ *is discrete in* $Z(\mathbb{A}_f)$, *there is an element* $m_0 \in \mathscr{M}_K(\mathrm{Sh}_K(G, X))$ *such that* $(\mathrm{Sh}(G, X), [m_0])$ *represents the functor* $\mathscr{M}_K/\approx$.

PROOF. Propositions 2.42 and 3.12 show that the map sending an element of $\mathscr{M}_K(T)$ to its Betti realization defines an isomorphism of moduli problems $(\mathscr{M}_K, \approx) \to (\mathscr{H}_K, \approx)$. Thus the theorem follows from (3.10) and (3.11). □

REMARK 3.14. When we assume Hypothesis 2.35, the same arguments show that every Shimura variety whose weight is defined over $\mathbb{Q}$ is a moduli variety for special motives.

Canonical models of Shimura varieties of abelian type. In this subsection, we prove the existence of canonical models for Shimura varieties of abelian type, and for those whose weight is defined over $\mathbb{Q}$, we realize the canonical model as a moduli variety.

Motives of CM-type. Let M be a motive of CM-type over $\mathbb{C}$, and let (T, h) be the extended Mumford-Tate group of M. Then $T(\mathbb{Q})$ acts on $H_B(M)$, and $T(\mathbb{A}_f)$ acts on $\omega_f(M)$. Associated with the pair (T, h), we have the reciprocity map

$$r(T, h): \mathrm{Gal}(E^{\mathrm{ab}}/E) \longrightarrow T(\mathbb{A}_f)/T(\mathbb{Q}).$$

As was noted in Milne [**51,** 4.7], we can regard M and any Hodge tensor on it as being defined over $\mathbb{Q}^{\mathrm{al}}$.

THEOREM 3.15. *Let* $\tau \in \mathrm{Gal}(\mathbb{Q}^{\mathrm{al}}/E(T, h))$. *For any representative* $\widetilde{r}(\tau) \in T(\mathbb{A}_f)$ *of* $r(T, h)(\tau)$, *there exists a unique isomorphism* $\gamma: M \to \tau M$ *such that*

(a) *for all Hodge tensors* s *on* M, $\gamma(s) = \tau s$;
(b) *for all* $v \in \omega_f(M)$, $\tau v = \gamma_f(\widetilde{r}(\tau) v)$.

PROOF. When M is an abelian variety and the Hodge tensors are endomorphisms or a polarization, this theorem essentially goes back to Shimura and Taniyama **[67]**. For a discussion of a stronger result (due to Deligne and Langlands) see Milne **[49**, I.5**]**. □

Generalized Siegel modular varieties. We first treat the varieties that play the same role for abelian motives that the Siegel modular varieties play for abelian varieties.

Let M be an abelian motive, and let N be the direct sum of M with the Tate motive. Let $(V, h_0) = H_B(N)$ and let t_0 be a polarization of M, which we can identify with a polarization of $H_B(M)$. Define G to be the subgroup of $\mathrm{GL}(V)$ of elements g such that

(3.16.1) g centralizes w_h;
(3.16.2) g preserves the decomposition $V = H_B(M) \oplus \mathbb{Q}(1)$;
(3.16.3) $gt_0 = t_0$.

The second condition implies that $G \subset \mathrm{GL}(H_B(M)) \times \mathbb{G}_m$, and we write t for the projection of G onto $\mathbb{G}_m$. In (3.16.3), G is to be understood as acting on $\mathbb{Q}(1)$ through t. Let X be the $G(\mathbb{R})$-conjugacy class of homomorphisms $h: \mathbb{S} \to G_{\mathbb{R}}$ containing h_0. Then (cf. Remark 2.6(a)) X consists of all Hodge structures h on V such that

(3.17.1) the weight gradation $V = \bigoplus V_m$ defined by h is the same as that of h_0;
(3.17.2) for each m, the Hodge structure on V_m defined by h has the same Hodge numbers as that of h_0;
(3.17.3) the projection of V onto $\mathbb{Q}(1)$ is a morphism of Hodge structures relative to h, and $\pm t_0$ is a polarization of $\mathrm{Ker}(V \to \mathbb{Q}(1))$.

The pair (G, X) satisfies the conditions (SV0, 1, 2*, 3), and so defines a Shimura variety $\mathrm{Sh}(G, X)$.

REMARK 3.18. Let $W = H_B(M)$, and let $W = \bigoplus W_m$ be its weight gradation. Then t_0 corresponds to a family (t_m) with t_m a polarization of W_m. Let $\psi_m = (2\pi i)^m t_m$. Then ψ_m is a pairing $V_m \times V_m \to \mathbb{Q}$ satisfying the conditions (2.4), and $\mathrm{Ker}(G \to \mathbb{G}_m)$ is a product of groups $\prod G_m$ with

$$G_m(\mathbb{Q}) = \{g \in \mathrm{GL}(W_m) \mid g\psi_m = \psi_m\}.$$

Thus $\mathrm{Ker}(G \to \mathbb{G}_m)$ is a product of symplectic and orthogonal groups.

EXAMPLE 3.19. Let A be an abelian variety over $\mathbb{C}$, and let t_0 be a polarization of A. Identify t_0 with a polarization of $W \underset{\text{df}}{=} H_B(A) \underset{\text{df}}{=} H_1(A, \mathbb{Q})$, and let $\psi = (2\pi i)^{-1} t_0$. The projection $G \to \mathrm{GL}(W)$ identifies G with the group of symplectic similitudes of the symplectic space (W, ψ), and X is the Siegel double space of all real Hodge structures of type $\{(-1, 0), (0, -1)\}$ on W for which $\pm t_0$ is a polarization. Thus $\mathrm{Sh}(G, X)$ is the Siegel modular variety.

LEMMA 3.20. *The reflex field* $E(G, X) = \mathbb{Q}$.

PROOF. Let $V = \bigoplus V_m$ be the weight gradation of h_0 and let $V(\mathbb{C}) = V^{p,q}$ be the Hodge decomposition of h_0. The $G(\mathbb{C})$-conjugacy class of μ_{h_0} can be identified with the set of gradations $V(\mathbb{C}) = \bigoplus V^p$ of $V(\mathbb{C})$ such that

$$V_m(\mathbb{C}) = \oplus(V_m(\mathbb{C}) \cap V^p),$$
$$\dim(V_m(\mathbb{C}) \cap V^p) = \dim V^{p,m-p}, \text{ and}$$
$$\psi(V_m(\mathbb{C}) \cap V^p, V_m(\mathbb{C}) \cap V^{p'}) = 0, \quad p + p' \neq m.$$

Since both the weight gradation and ψ are defined over $\mathbb{Q}$, so is this set, which shows that $E(G, X) = \mathbb{Q}$. □

THEOREM 3.21. *The Shimura variety* $\mathrm{Sh}(G, X)$ *has a canonical model over* $\mathbb{Q}$.

PROOF. The centre of G is $\mathbb{G}_m$, and $\mathbb{Q}^\times$ is discrete in $\mathbb{A}_f^\times$. Hence (see Theorem 3.13), there is an element $m_0 \in \mathscr{M}(\mathrm{Sh}_K(G, X))$ with the following property: for any $m \in \mathscr{M}_K(T)$, there is a unique morphism $\gamma\colon T \to \mathrm{Sh}_K(G, X)$ such that $\gamma^* m_0 \approx m$. For any automorphism τ of $\mathrm{Aut}(\mathbb{C})$, $\tau m_0 \in \mathscr{M}(\tau\mathrm{Sh}(G, X))$, and so there is a unique morphism

$$\gamma_\tau\colon \tau\,\mathrm{Sh}(G, X) \longrightarrow \mathrm{Sh}(G, X)$$

such that $\gamma_\tau^* m_0 \approx \tau m_0$.

LEMMA 3.22. *For any elements* $\sigma, \tau \in \mathrm{Aut}(\mathbb{C})$, $\gamma_\sigma \circ \sigma\gamma_\tau = \gamma_{\sigma\tau}$.

PROOF. The composite $\gamma_\sigma \circ \sigma\gamma_\tau$ is a map $\sigma\tau\,\mathrm{Sh}(G, X) \to \mathrm{Sh}(G, X)$ with the property that

$$(\gamma_\sigma \circ \sigma\gamma_\tau)^* m_0 = (\sigma\gamma_\tau)^* \circ \gamma_\sigma^* m_0 \approx (\sigma\gamma_\tau)^*(\sigma m_0) = \sigma(\gamma_\tau^* m_0) \approx \sigma\tau m_0.$$

It therefore equals $\gamma_{\sigma\tau}$. □

Recall the following general result from descent theory.

LEMMA 3.23. *Let* $\Omega \supset k$ *be fields of characteristic zero with* Ω *algebraically closed, and let* V *be a quasi-projective variety over* Ω. *Suppose that there are given isomorphisms* $\gamma_\tau\colon \tau V \to V$ *for all* $\tau \in \mathrm{Aut}(\Omega/k)$ *satisfying the cocycle condition*:

$$\gamma_\sigma \circ \sigma\gamma_\tau = \gamma_{\sigma\tau}, \qquad \sigma, \tau \in \mathrm{Aut}(\Omega/k).$$

Then there exists a model $(V_0, \alpha\colon V \to V_{0,\Omega})$ *of* V *over* k *such that*

$$\gamma_\tau = \alpha^{-1} \circ \tau\alpha;$$

moreover (V_0, α) *is uniquely determined (up to a unique* k*-isomorphism).*

PROOF. See Weil [72]. □

We now complete the proof of Theorem 3.21. On applying the two lemmas, we obtain a model S_K of $\mathrm{Sh}_K(G, X)$ over $\mathbb{Q}$ with the property that

$$\alpha\colon \mathscr{M}_K(\mathbb{C}) \longrightarrow S_K(\mathbb{C})$$

commutes with the action of $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$. On varying K, we obtain a model S of $\mathrm{Sh}(G, X)$ over $\mathbb{Q}$, and it remains to show that this model satisfies (3.4.1). Let x be a special point in X, and let $(M, \mathfrak{s}, [\eta])$ map to $[x, 1]$ under α. Recall that this means that there exists an isomorphism

$$\beta\colon H_B(M) \longrightarrow (V, h_x)$$

of rational Hodge structures such that $\beta(s_i) = t_i$ for all $i \in I$ and $\eta \circ \beta_f = \mathrm{id}$. Such a β defines an isomorphism of the Mumford-Tate group T of M with the Mumford-Tate group of (V, h_x), which (by definition) is commutative, and so M is of CM-type. Let $\tau \in \mathrm{Gal}(\mathbb{Q}^{\mathrm{al}}/\mathbb{Q})$ and extend it to an automorphism of $\mathbb{C}$. Then

$$\tau[x, 1] = \tau\alpha(M, \mathfrak{s}, [\eta]) = \alpha(\tau M, (\tau s_i)_i, [\tau \circ \eta]).$$

According to Theorem 3.15, there is an isomorphism $\gamma\colon M \to \tau M$ such that $\gamma(s) = \tau s$ for all Hodge cycles s on M and $\tau v = \gamma_f(\tilde{r}(\tau)v)$ for all $v \in \omega_f(M)$. Define $\beta'\colon H_B(\tau M) \to V$ to be $\beta \circ H_B(\gamma^{-1})$. Then β' is a morphism of Hodge structures $H_B(\tau M) \to (V, h_x)$ that maps τs_i to t_i for each i, and has the property that $\beta'_f \circ (\tau \circ \eta) = \tilde{r}(\tau)$. Hence

$$\alpha(\tau M, (\tau s_i)_i, [\tau \circ \eta]) = [x, \tilde{r}(\tau)]$$

where $\tilde{r}(\tau)$ represents $r(T, h_x)(\tau)$. □

Shimura varieties whose weight is defined over $\mathbb{Q}$. We next construct the canonical model of a Shimura variety of abelian type whose weight is defined over $\mathbb{Q}$. The following proposition from Deligne **[15]** will be useful.

PROPOSITION 3.24. *Let $f\colon G \hookrightarrow G'$ be an injective homomorphism sending X into X'. The map*

$$[x, g] \mapsto [f(x), f(g)]\colon \mathrm{Sh}(G, X) \longrightarrow \mathrm{Sh}(G', X')$$

is a closed immersion. If $\mathrm{Sh}(G', X')$ has a canonical model over a number field $E \supset E(G', X')$, then the image of $\mathrm{Sh}(G, X)$ in $\mathrm{Sh}(G', X')$ is defined over $E \cdot E(G, X)$ and is a canonical model.

PROOF. That the map is a closed immersion is proved in **[15,** 1.15.1**]**. Identify $\mathrm{Sh}(G, X)$ with its image. Let (T, x) be a special pair in (G, X). For any τ fixing $E{\cdot}E(T, x)$ and $a \in G(\mathbb{A}_f)$, $\tau[x, a] = [x, a{\cdot}r(T, h_x)(\tau)] \in \mathrm{Sh}(G, X)(\mathbb{C})$. We now apply the following two lemmas. The first shows that $\tau\mathrm{Sh}(G, X) = \mathrm{Sh}(G, X)$ for any τ fixing $E \cdot E(T, x)$, and the second shows that such elements generate $\mathrm{Aut}(\mathbb{C}/E \cdot E(G, X))$. □

LEMMA 3.25. *For any special point x of X,*

$$\{[x, a] \mid a \in G(\mathbb{A}_f)\}$$

is Zariski dense in $\mathrm{Sh}(G, X)$.

PROOF. See **[15,** 5.2**]**. □

Lemma 3.26. *Let (G, X) be a pair satisfying* (SV0,1,2). *For any finite extension E' of $E(G, X)$, there exists a special pair $(T, x) \subset (G, X)$ such that E' and $E(T, x)$ are linearly disjoint over $E(G, X)$.*

Proof. See [**15**, 5.1]. □

Theorem 3.27. *Every Shimura variety of abelian type whose weight is defined over $\mathbb{Q}$ has a canonical model over its reflex field.*

Proof. From Theorem 3.21 and Proposition 3.24 we obtain the following criterion: a Shimura variety $\mathrm{Sh}(G, X)$ has a canonical model over its reflex field if there exists an inclusion $(G, X) \hookrightarrow (G', X')$ with (G', X') the pair associated (as in (3.16)) to a polarized abelian motive (M, t_0).

Let $\mathrm{Sh}(G, X)$ be a Shimura variety of abelian type with weight defined over $\mathbb{Q}$. Assume that G is generated as a group over $\mathbb{Q}$ by $\{h_x \mid x \in X\}$, and choose a faithful representation $\xi\colon G \hookrightarrow \mathrm{GL}(V)$. Then it follows from Proposition 1.38 that G is the Mumford-Tate group of (V, h_o) for some $o \in X$. Since $\mathrm{Sh}(G, X)$ is of abelian type, (V, h_o) is the Betti realization of an abelian motive M (see Proposition 3.12), and it is clear that $\mathrm{Sh}(G, X)$ satisfies the above criterion relative to M and any polarization t_0 of M.

Now drop the hypothesis that G is generated by $\{h_x\}$. The composite

$$\mathbb{S} \xrightarrow{h_x} G_{\mathbb{R}} \longrightarrow G^{\mathrm{ab}}_{\mathbb{R}}$$

is independent of x. Denote it by h_X, and let H be the $\mathbb{Q}$-subtorus of G^{ab} generated by h_X. Let G' be the inverse image of H in G, and let X' be a $G'(\mathbb{R})$-conjugacy class of maps $\mathbb{S} \to G'_{\mathbb{R}}$ such that the inclusion $G' \hookrightarrow G$ maps X' into X. Then $E(G, X) = E(G', X')$, $\mathrm{Sh}(G, X) = \mathrm{Sh}(G', X') \cdot G(\mathbb{A}_f)$, and, after possibly enlarging $Z(G)$ so that $H^1(\mathbb{Q}, Z(G)) = 0$ (this is permissible by Proposition 3.24)

$$s \cdot g = s', \quad s, s' \in \mathrm{Sh}(G', X'), \quad g \in G(\mathbb{A}_f) \quad \Rightarrow \quad g \in G'(\mathbb{A}_f) \cdot Z(G)(\mathbb{Q})^{-}.$$

For $\tau \in \mathrm{Aut}(\mathbb{C}/E(G, X))$, let $\gamma_\tau\colon \tau\mathrm{Sh}(G', X') \to \mathrm{Sh}(G', X')$ be the map defined by the canonical model of $\mathrm{Sh}(G', X')$. Then γ_τ has a unique extension to a $G(\mathbb{A}_f)$-equivariant map $\gamma'_\tau\colon \tau\mathrm{Sh}(G, X) \to \mathrm{Sh}(G, X)$, namely,

$$\gamma'_\tau(s \cdot g) \underset{\mathrm{df}}{=} \gamma_\tau(s) \cdot g, \qquad s \in \mathrm{Sh}(G', X'), \ g \in G(\mathbb{A}_f),$$

Obviously the γ'_τ's satisfy the cocycle condition, and so define a model of $\mathrm{Sh}(G, X)$ over $E(G, X)$, which is canonical. □

Let M be a motive over a field k of characteristic zero. Choose an algebraic closure k^{al} of k, and let $\omega_f(M)$ be the restricted product (over ℓ) of the ℓ-adic étale cohomology groups of $M \otimes k^{\mathrm{al}}$. For any compact open subgroup K of $G(\mathbb{A}_f)$, a *level K-structure* on M is a K-orbit $[\eta]$ of isomorphisms $\eta\colon V(\mathbb{A}_f) \to \omega_f(M)$ that is stable under the action of $\mathrm{Gal}(k^{\mathrm{al}}/k)$, i.e., such that

$$\eta_0 \in [\eta], \quad \tau \in \mathrm{Gal}(k^{\mathrm{al}}/k) \implies \tau \circ \eta_0 \in [\eta].$$

Let $\mathrm{Sh}(G, X)$ be a Shimura variety of abelian type whose weight is defined over $\mathbb{Q}$. Assume that there is given a homomorphism $t\colon G \to \mathbb{G}_m$ such that $t \circ w_X = -2$, a faithful representation $\xi\colon G \hookrightarrow \mathrm{GL}(V)$ of G, and a tensor t_0 for V fixed by G and such that $\pm t_0$ is a polarization of (V, h_x) for all $x \in X$. Choose a point $o \in X$. Then $((V, h_o), t_0)$ is the Betti realization of a polarized abelian motive M, and if (G', X') is the pair associated with M (as in (3.16)), then $(G, X) \subset (G', X')$.

For any field $k \supset E(G, X)$, we define $\mathscr{M}_K(k)$ to be the set of triples $(M, \mathfrak{s}, [\eta])$ consisting of an abelian motive M over k, a set $\mathfrak{s}$ of Hodge cycles on M, and a level K-structure $[\eta]$ on M satisfying the following conditions:

(3.28.1) for every $E(G, X)$-homomorphism $\tau\colon k \hookrightarrow \mathbb{C}$, there is an isomorphism $\beta\colon H_B(\tau M) \to V$ sending each s_i to t_i and h_M to an element of X;

(3.28.2) one (hence every) representative η of the level structure maps each s_i to t_i.

LEMMA 3.29. *If* (3.28.1) *holds for one $E(G, X)$-homomorphism τ, then it holds for all.*

PROOF. Suppose (3.28.1) holds for one embedding τ. Then $\tau(M, \mathfrak{s}, [\eta])$ defines a point P of $\mathscr{M}_K(G, X)(\mathbb{C})$. There exists a compact open subgroup K' of $G'(\mathbb{A}_f)$ such that $(G, X) \hookrightarrow (G', X')$ induces a closed immersion

$$\mathrm{Sh}_K(G, X) \hookrightarrow \mathrm{Sh}_{K'}(G', X'),$$

and, by construction, this is defined over $E(G, X)$. The point P in $\mathrm{Sh}_{K'}(G', X')$ is rational over k, and lies in $\mathrm{Sh}_K(G, X)(\mathbb{C})$. If we replace τ by its composite τ' with an element of $\mathrm{Aut}(\mathbb{C}/E(G, X))$, the point remains in $\mathrm{Sh}_K(G, X)(\mathbb{C})$, and this implies that $\tau'(M, \mathfrak{s}, [\eta]) \in \mathscr{M}_K(G, X)(\mathbb{C})$. □

REMARK 3.30. In essence, Lemma 3.29 says that the moduli problem $\mathscr{M}$ is defined over $E(G, X)$. It is possible to prove this directly, i.e., without using the existence of canonical models, but the argument then is more complicated.

Lemma 3.29 allows us to define a moduli problem on smooth algebraic varieties T over $E(G, X)$. Let $\mathscr{M}_K(T)$ be the set of triples $(M, \mathfrak{s}, [\eta])$ consisting of an abelian motive M over T, a family of Hodge tensors $\mathfrak{s}$ on M indexed by I, and a level K-structure on M such that, for all $\mathbb{C}$-valued points t of T, $(M, \mathfrak{s}, [\eta])_t \in \mathscr{M}_K(\mathbb{C})$. With the obvious notions of pullback and isomorphism, this becomes a moduli problem on the category of smooth algebraic varieties over $E(G, X)$.

THEOREM 3.31. *The pair* $(\mathrm{Sh}_K(G, X), \alpha_K)$, *where* $\mathrm{Sh}(G, X)$ *here denotes the canonical model of the Shimura variety, is a solution to the moduli problem over* $E(G, X)$.

PROOF. First note that it follows from Lemma 3.29 that $\mathcal{M}_K(\mathbb{C}) = \varinjlim \mathcal{M}_K(R)$ where the limit is over the subalgebras R of Ω that are finitely generated over $E(G, X)$. From our construction of the canonical model, it is clear that the map $\beta_m\colon T_{\mathbb{C}} \to \mathrm{Sh}_K(G, X)_{\mathbb{C}}$ corresponding to an element $m \in \mathcal{M}_K(T)$ is defined over $E(G, X)$. Finally, an obvious extension of Lemma 3.23 to motives provides us with elements m_i defined on some étale covering of $\mathrm{Sh}_K(G, X)$ and satisfying (2.3.2). □

APPLICATION 3.32. In the case that the reflex field $E(G, X)$ is real, it is possible to use Theorem 3.31 to give a description of the action of complex conjugation on $\mathrm{Sh}(G, X)(\mathbb{C})$—see Milne and Shih **[52]**.

The general case.

THEOREM 3.33. *Every Shimura variety of abelian type admits a canonical model over its reflex field.*

PROOF. Consider a Shimura variety $\mathrm{Sh}(G, X)$ of abelian type. The weight map w_X is a homomorphism $\mathbb{G}_m \to Z^0 \underset{\mathrm{df}}{=} Z(G)^0$.

Write $(G^{\mathrm{ad}}, X^{\mathrm{ad}}) = \prod(G_i, X_i)$ where the G_i are the simple factors of G^{ad} and X_i is the projection of X onto G_i. Then $E(G, X)$ is the composite of the fields $E(G_i, X_i)$ and $E(G^{\mathrm{ab}}, h_X)$, and each field $E(G_i, X_i)$ is either totally real or a CM-field **[15**, 3.8]. Because of our assumption (SV0), $E(G^{\mathrm{ab}}, h_X)$ is a subfield of a CM-field, and so $E(G, X)$ is a subfield of a CM-field. The weight w_X is defined over $E(G, X)$ and is invariant under ι; it is therefore defined over a totally real subfield F of $E(G, X)$. Choose a quadratic imaginary extension E of $\mathbb{Q}$ in $\mathbb{C}$, and let $Z_* = \mathrm{Res}_{E/\mathbb{Q}}(Z^0)_E$. Define

$$\varepsilon_0 = \mathrm{Res}_{EF/F}(w_X^{-1}) : \mathrm{Res}_{EF/F}\,\mathbb{G}_m \longrightarrow \mathrm{Res}_{E/F}\,Z^0_E = (Z_*)_F.$$

Then ε_0 is defined over F, and over $\mathbb{R}$ it can be identified with a homomorphism $\varepsilon \underset{\mathrm{df}}{=} \varepsilon_{0,\mathbb{R}}\colon \mathbb{S} \to Z_{*,\mathbb{R}}$. The weight of ε is w_X.

There are natural inclusions $Z^0 \hookrightarrow G$, $Z^0 \hookrightarrow Z_*$, and we define G_* to be the quotient:

$$G_* = G \times Z_*/Z^0 \quad (\text{diagonal embedding of } Z^0).$$

Let $h_0 \in X$. The composite

$$\mathbb{S} \xrightarrow{(h_0,\varepsilon)} G_{\mathbb{R}} \times Z_{*\mathbb{R}} \longrightarrow G_{*\mathbb{R}}$$

has weight zero, and we define X_* to be its $G_*(\mathbb{R})$-conjugacy class. Clearly $\mathrm{Sh}(G_*, X_*)$ is of abelian type, and so Theorem 3.27 shows that it has a canonical model over its reflex field $E(G_*, X_*)$. Moreover,

$$E(G_*, X_*) \cdot E = E(G, X) \cdot E.$$

Let ε' be the composite

$$\mathbb{S} \xrightarrow{(1,\varepsilon)} G_{\mathbb{R}} \times Z_{*\mathbb{R}} \longrightarrow G_{*\mathbb{R}},$$

and let

$$X_* \cdot \varepsilon'^{-1} = \{h \cdot \varepsilon'^{-1} \mid h \in X_*\}.$$

On applying Lemma 3.34 below to $\mathrm{Sh}(G_*, X_*)$, we find that $\mathrm{Sh}(G_*, X_* \cdot \varepsilon'^{-1})$ has a canonical model over $E \cdot E(G_*, X_*)$. The inclusion $G \hookrightarrow G_*$ maps X into $X_* \cdot \varepsilon'^{-1}$, and so we can apply Lemma 3.24 to show that $\mathrm{Sh}(G, X)$ has a canonical model over $E \cdot E(G_*, X_*) = E \cdot E(G, X)$. Since E is an arbitrary quadratic imaginary extension of $\mathbb{Q}$, we can apply Lemma 3.35 below to show that $\mathrm{Sh}(G, X)$ has a canonical model over $E(G, X)$.

LEMMA 3.34. *Let* $\mathrm{Sh}(G, X)$ *be a Shimura variety, and let* ε *be a homomorphism* $\mathbb{S} \to Z(G)_{\mathbb{R}}$. *If* $\mathrm{Sh}(G, X)$ *admits a canonical model over a field* E *containing* $E(G, X) \cdot E(Z^0, \varepsilon)$, *then so also does* $\mathrm{Sh}(G, X \cdot \varepsilon)$.

PROOF. We have a morphism

$$[h, g], [\varepsilon, z] \mapsto [h \cdot \varepsilon, gz]: \mathrm{Sh}(G, X) \times \mathrm{Sh}(Z^0, \varepsilon) \longrightarrow \mathrm{Sh}(G, X \cdot \varepsilon).$$

Let $d: Z^0 \to G \times Z^0$ be the homomorphism $z \mapsto (z, z^{-1})$. This defines an action of $Z^0(\mathbb{A}_f)$ on $\mathrm{Sh}(G, X) \times \mathrm{Sh}(Z^0, \varepsilon)$ with quotient $\mathrm{Sh}(G, X \cdot \varepsilon)$. The quotient of the product of the canonical models of $\mathrm{Sh}(G, X)$ and $\mathrm{Sh}(Z^0, \varepsilon)$ by $Z^0(\mathbb{A}_f)$ is a canonical model for $\mathrm{Sh}(G, X \cdot \varepsilon)$. (See Deligne **[15**, 5.11].)

LEMMA 3.35. *Let* $\{E_i\}$ *be a family of finite extensions of* $E(G, X)$ *whose intersection is* $E(G, X)$. *If* $\mathrm{Sh}(G, X)$ *has a canonical model* $\mathrm{Sh}(G, X)_{E_i}$ *over each* E_i, *then it has a unique model* $\mathrm{Sh}(G, X)_E$ *over* $E(G, X)$ *such that, for all* i, *the model* $\mathrm{Sh}(G, X)_E \otimes E_i$ *over* E_i *is isomorphic to* $\mathrm{Sh}(G, X)_{E_i}$. *If, further, the family* $\{E_i\}$ *has the property that* $\bigcap E_i F = F$ *for any finite extension* F *of* $E(G, X)$, *then the model* $\mathrm{Sh}(G, X)_E$ *is canonical.*

PROOF. The first statement is proved in **[15**, 5.10]. To show that the model satisfies (3.4.1) for a specific special pair (T, h), use that $\bigcap E_i \cdot E(T, h) = E(T, h)$. □

REMARK 3.36. Let $F \subset E(G, X)$ be the field of definition of w_X. Corresponding to any quadratic imaginary extension E of $\mathbb{Q}$, we have a map

$$\mathrm{Sh}(G, X) \times \mathrm{Sh}(Z_*, \varepsilon) \longrightarrow \mathrm{Sh}(G_*, X_*)$$

rational over $E \cdot E(G, X)$. Sometimes this can be interpreted as the map sending two Hodge structures to their tensor product (see **[15**, 6.6] for an example of this), and it can be used to obtain information about the points of $\mathrm{Sh}(G, X)$.

REMARK 3.37. The existence of a canonical model over the reflex field shows that, for any automorphism τ of $\mathbb{C}$ fixing $E(G, X)$, there is a canonical isomorphism $\tau\mathrm{Sh}(G, X) \to \mathrm{Sh}(G, X)$. Langlands conjectured **[45]** that for *any* automorphism τ of $\mathbb{C}$, there is a canonical isomorphism $\tau\mathrm{Sh}(G, X) \to \mathrm{Sh}(G', X')$ for a suitable pair (G', X') defined explicitly in terms of (G, X), τ, and a special point of x. For Shimura varieties of abelian type, this conjecture can be proved using similar techniques to the above.

Applications. In many respects, Theorem 3.31 allows us to treat Shimura varieties of abelian type as easily as Shimura varieties of PEL-type, at least in characteristic zero. To illustrate this, I list some applications.

Consequences of the Tate conjecture. The Tate conjectures **[70,** T(X), E(X)] imply the following statement:

(3.38) For any motives M and N over a field k finitely generated over $\mathbb{Q}$ and for any prime ℓ, the homomorphism

$$\mathrm{Hom}(M, N) \otimes \mathbb{Q}_\ell \longrightarrow \mathrm{Hom}(\omega_\ell(M), \omega_\ell(N))^\Gamma, \qquad \Gamma = \mathrm{Gal}(k^{\mathrm{al}}/k),$$

is bijective.

Faltings has proved (3.38) for motives of the form $h_1(A)$, A an abelian variety, **[25]**. Consequently, it is also true for direct factors of such motives.[8] Silverberg has investigated the consequences of (3.38) for one class of Shimura varieties (essentially that described in Example 4.31 below)—see Silverberg **[68, 69]**. Here we explain its consequences for a Shimura variety $\mathrm{Sh}(G, X)$ of abelian type whose weight is defined over $\mathbb{Q}$ and such that $Z(\mathbb{Q})$ is closed in $Z(\mathbb{A}_f)$.

To a point $x \in X$, Shimura attaches an adèlic representation ρ_x **[66,** 7.2, 7.3, 7.6, 7.8]. When we choose a faithful representation (V, ξ) of G as above, then $\xi_f \circ \rho_x$ becomes the Galois representation on $\omega_f(M_x)$ for M_x the motive attached to $[x, 1] \in \mathrm{Sh}(G, X)$. If (3.38) holds for the motives in the family parametrized by $\mathrm{Sh}(G, X)$, then we can read off the following result:

(3.39) for $x, y \in X$ and $\alpha \in G(\mathbb{Q})$,

$$\alpha x = y \Leftrightarrow \mathrm{ad}\,\alpha \circ \rho_x = \rho_y.$$

Let $I(x) = \mathrm{Aut}(N_x)$ where $N_x = (M_x, \mathfrak{s}_x)$ is the motive together with tensor structure attached to the point $[x, 1]$ and a faithful representation of G. Assume that $I(x)$ satisfies the Hasse principle for H^1, i.e.,

$$H^1(\mathbb{Q}, I(x)) \longrightarrow \Pi_\ell H^1(\mathbb{Q}_\ell, I(x))$$

[8] Of course, if the Tate conjecture were known for abelian varieties, then the map would be bijective for all abelian motives, but that is not what Faltings proves.

is injective. Then (3.38) implies the following statement:

(3.40) if there exists an $\alpha_f \in G(\mathbb{A}_f)$ such that $\operatorname{ad}\alpha_f \circ \rho_x = \rho_y$, then there exists an $\alpha_0 \in G(\mathbb{Q})$ such that $\alpha_0 x = y$.

In geometric terms, the Tate conjecture implies that the $I(x)_{\mathbb{A}_f}$-torsor $\operatorname{Hom}(\omega_f(N_x), \omega_f(N_y))$ is obtained from the $I(x)$-torsor $\operatorname{Hom}(N_x, N_y)$ by the base change $\mathbb{Q} \to \mathbb{A}_f$. The hypothesis in (3.40) implies that the first torsor is trivial. The Hasse principal then implies that the second torsor is trivial, which implies the conclusion of (3.40).

Systems of realizations. Let S be a smooth scheme over a number field E. A *system of realizations* on S is given by the following data:

(3.41.1) For each embedding $\tau: E \hookrightarrow \mathbb{C}$, a local system of $\mathbb{Q}$-vector spaces on $(\tau S)(\mathbb{C})$.

(3.41.2) A vector bundle H_{dR} on S endowed with a flat connection ∇ and a descending "Hodge filtration" F by subbundles. The connection is required to satisfy the axiom of transversality (2.15.2) and to have regular singularities at infinity.

(3.41.3) A local system of $\mathbb{A}_f$-modules H_f on $S_{\text{ét}}$.

(3.41.4) Comparison isomorphisms relating the above data.

(3.41.5) Weight gradations on each of H_τ, H_{dR}, and H_f which are respected by the comparison isomorphisms.

(3.41.6) An involution $F_\infty: \bigoplus H_\tau \to \bigoplus H_\tau$ (*Frobenius map at infinity*) respecting the weight gradation.

The data are required to satisfy certain conditions, for example, the "Betti realization" H_τ, endowed with the weight gradation and the filtration provided, via the comparison isomorphism, by that on H_{dR}, is a variation of Hodge structures. See **[20,** §1] for more details.

Consider a Shimura variety of abelian type, and let G^c be the largest quotient of G such that, for any $h \in X$, (G^c, h) satisfies (SV2*) and has weight defined over $\mathbb{Q}$ (G^c is the quotient of G by a subgroup of its centre). Then there is a canonical tensor functor from $\mathbf{Rep}_{\mathbb{Q}}(G^c)$ to the category of systems of realizations on the canonical model of $\operatorname{Sh}(G, X)$. In fact, every rational representation of G^c defines a family of abelian motives on $\operatorname{Sh}(G, X)$, and every family of motives defines a system of realizations.

Automorphic vector bundles. Automorphic vector bundles are those vector bundles on Shimura varieties whose sections are holomorphic automorphic forms (in the classical sense). It is known that automorphic vector bundles have canonical models over number fields and, hence, that it makes sense to speak of an automorphic form being defined over such a field. In Milne **[49,** III] a heuristic explanation of this statement is given in terms of motives. For Shimura varieties of abelian type, the explanation is now a proof.

The boundaries of Shimura varieties. The study of the boundary of a moduli variety for abelian varieties (with additional structure) is equivalent to the study of the degeneration of the abelian varieties—see Namikawa **[56]** for Siegel modular varieties over $\mathbb{C}$, Faltings and Chai **[27]** for Siegel modular varieties over $\mathbb{Z}$, and Brylinski **[12]** for Shimura varieties of Hodge type over $\mathbb{Q}$. Theorem 3.31 allows us to treat Shimura varieties of abelian type in the same fashion. I hope to return to this in a future work.

4. The points on a Shimura variety modulo a prime of good reduction

Let $\mathrm{Sh}(G, X)$ be a Shimura variety of abelian type whose weight is defined over $\mathbb{Q}$. In the last section, we obtained a motivic description of the points of $\mathrm{Sh}(G, X)$ with coordinates in any field containing $E(G, X)$. In this section, we show that, if one assumes the existence of a good theory of abelian motives in mixed characteristic, then the description extends to the points of $\mathrm{Sh}(G, X)$ in finite fields, and we thereby obtain a heuristic derivation of the conjecture of Langlands and Rapoport **[46]**.

Statement of the problem. As we saw in §3, starting from a connected reductive group G over $\mathbb{Q}$, a $G(\mathbb{R})$-conjugacy class X of homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$ satisfying the conditions (SV0–2), and a compact open subgroup K of $G(\mathbb{A}_f)$, we obtain a variety $\mathrm{Sh}_K(G, X)$ over $\mathbb{C}$. The reflex field $E(G, X)$ is a number field (contained in $\mathbb{C}$) that is defined purely in terms of G and X, and $\mathrm{Sh}_K(G, X)$ has a canonical model over $E(G, X)$. Let v be a prime of E lying over a finite prime p of $\mathbb{Q}$, and let E_v be the completion of $E(G, X)$ at v. Assume that $\mathrm{Sh}_K(G, X)$ has good reduction at v, i.e., that there is a smooth scheme $\mathrm{Sh}_K(G, X)_v$ over the ring of integers $\mathcal{O}_v$ in E_v whose generic fibre is $\mathrm{Sh}_K(G, X)_{E_v}$. The problem then is to describe the sets

$$\mathrm{Sh}_K(G, X)_v(k)$$

for k a finite field containing the residue field $\kappa(v)$ at v or, equivalently, to describe the set

$$\mathrm{Sh}_K(G, X)_v(\mathbb{F}), \qquad \mathbb{F} = k^{\mathrm{al}},$$

together with the action of the Frobenius element of $\mathrm{Gal}(\mathbb{F}/\kappa(v))$. For the applications, we shall also need to know how $G(\mathbb{A}_f)$ acts on the sets. Note that the problem is well-posed only if $\mathrm{Sh}_K(G, X)$ has a *canonical* smooth model over $\mathcal{O}_v$.

The next example illustrates the fact that, unless the component of K at p is maximal, we can not expect $\mathrm{Sh}_K(G, X)$ to have good reduction at primes lying over p.

EXAMPLE 4.1. Let $G = \mathrm{GL}_2$, and let X be the conjugacy class of homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$ containing the map

$$a + bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Then $\mathrm{Sh}_{K(N)}(G, X)$ is the moduli variety for elliptic curves with level N structure, and it is known that this variety has good reduction at p if and only if p does not divide N (Deligne and Rapoport **[23]**).

Thus we should assume that K is of the form $K^p \cdot K_p$ with K_p a *maximal* compact open subgroup of $G(\mathbb{Q}_p)$ and K^p a compact open subgroup of $G(\mathbb{A}_f^p)$. However, the next example shows that even this is not sufficient to ensure that $\mathrm{Sh}_K(G, X)$ has good reduction at v.

EXAMPLE 4.2. Let B be a quaternion algebra over $\mathbb{Q}$ that splits at the real prime. Let $G = \mathrm{GL}_1(B)$. Then $G_{\mathbb{R}} \approx \mathrm{GL}_2$, and we can define X as in the last example. Let $K = K^p \cdot K_p$ where K_p is a maximal compact subgroup of $G(\mathbb{Q}_p)$. Then $\mathrm{Sh}_K(G, X)$ has good reduction at p if and only if p does not divide the discriminant of B.

In Langlands **[44**, p. 411] it is suggested that $\mathrm{Sh}_K(G, X)$ should have good reduction at $v|p$ if $K = K^p \cdot K_p$ with K_p a hyperspecial subgroup[9] of $G(\mathbb{Q}_p)$. Recall (Tits **[71**, 3.8.1] that a subgroup K_p of $G(\mathbb{Q}_p)$ is said to be *hyperspecial* if there is a smooth group scheme G_p over $\mathbb{Z}_p$ such that

(4.3.1) $G_p(\mathbb{Z}_p) = K_p$;

(4.3.2) the reduction of G_p modulo p is a connected reductive group over $\mathbb{F}_p$.

Since algebraic groups over finite fields are quasi-split, a necessary condition that there exist a hyperspecial subgroup of $G(\mathbb{Q}_p)$ is that G be quasi-split over $\mathbb{Q}_p$ and split over an unramified extension of $\mathbb{Q}_p$, i.e., that G be *unramified* at p. Conversely, Tits **[71**, p. 36] shows that this condition is sufficient. Consequently hyperspecial subgroups exist in $G(\mathbb{Q}_p)$ for almost all p.

EXAMPLE 4.4. Let (V, ψ) be a symplectic space over $\mathbb{Q}$, and let $G = G(\psi)$ be the group of symplectic similitudes. A hyperspecial subgroup of $G(\mathbb{Q}_p)$ is the stabilizer of a lattice $\Lambda \subset V(\mathbb{Q}_p)$ such that ψ (or some multiple of ψ) restricts to a $\mathbb{Z}_p$-valued form on Λ with determinant a p-adic unit.

In the following, we fix a hyperspecial subgroup K_p, and we again write G for the smooth group scheme over $\mathbb{Z}_p$ such that $G(\mathbb{Z}_p) = K_p$. Thus $G(R)$ is defined whenever R is a $\mathbb{Q}$-algebra or a $\mathbb{Z}_p$-algebra.

We may as well pass to the limit over the compact open subgroups K^p of $G(\mathbb{A}_f^p)$ and write

$$\mathrm{Sh}_p(G, X) = \varprojlim_{K^p} \mathrm{Sh}_{K^p \cdot K_p}(G, X).$$

Assume $\mathrm{Sh}_p(G, X)$ has a smooth model over $\mathcal{O}_v$ (see below), and denote it by $\mathrm{Sh}_p(G, X)_v$. By definition, the action of $G(\mathbb{A}_f^p)$ extends to $\mathrm{Sh}_p(G, X)_v$, and so we obtain a set

$$\mathrm{Sh}_p(\mathbb{F}) = \mathrm{Sh}_p(G, X)_v(\mathbb{F})$$

together with commuting actions of $G(\mathbb{A}_f^p)$ and the geometric Frobenius

[9] Conversations with Blasius, Chai, and Prasad suggest that this condition can be weakened.

element[10] $\Phi \in \mathrm{Gal}(\mathbb{F}/\kappa(v))$. The problem discussed in this section is that of determining the isomorphism class of the system $(\mathrm{Sh}_p(\mathbb{F}), \times, \Phi)$ consisting of the set $\mathrm{Sh}_p(\mathbb{F})$ together with the action " $\times$ " of $G(\mathbb{A}_f^p)$ and the action of Φ.

The building. There is a more natural definition of hyperspecial subgroups in terms of the building $\mathscr{B}(G, F)$ that Bruhat and Tits attach to a reductive group G over a local field F (Tits **[71]**). This is a set with a left action of $G(F)$, certain of whose vertices are said to be *hyperspecial*, and a subgroup K_p of $G(F)$ is hyperspecial if it is the stabilizer of such a vertex. The construction of the building commutes with the formation of unramified extensions of F.

The building $\mathscr{B}(G, F)$ is a union of apartments, and the apartments are in one-to-one correspondence with the maximal F-split subtori S of G. Let p_0 be a hyperspecial vertex of the apartment of S, and let $G(\mathscr{O}_F)$ be its stabilizer. Assume G is split over F. Because S is split, it has a canonical $\mathscr{O}_F$-structure such that $S(R) = \mathrm{Hom}(X^*(S), R^\times)$ for any $\mathscr{O}_F$-algebra R. Our hypotheses imply that

(4.5.1) $S(\mathscr{O}_F) \subset G(\mathscr{O}_F)$;
(4.5.2) $N(F) \subset G(\mathscr{O}_F) \cdot S(F)$, where N is the normalizer of S;
(4.5.3) $G(F) = G(\mathscr{O}_F) \cdot S(B) \cdot G(\mathscr{O}_F)$.

The equality (4.5.3) is the Cartan decomposition (Tits **[71,** 3.3.3]). The remaining two statements imply that the Weyl group has a set of representatives in $G(\mathscr{O}_F)$.

Now return to the situation of the previous subsection, so that G is a reductive group over $\mathbb{Q}$. For any field $F \supset \mathbb{Q}$, write $\mathscr{C}(F)$ for the set of $G(F)$-conjugacy classes of homomorphisms $\mathbb{G}_m \to G_F$. Note that a map $F \to F'$ defines a map $\mathscr{C}(F) \to \mathscr{C}(F')$; in particular, when F' is Galois over F, $\mathrm{Gal}(F'/F)$ acts on $\mathscr{C}(F')$.

PROPOSITION 4.6. (a) *For any maximal F-split torus S in G_F, with F-Weyl group Ω, the map $X_*(S)/\Omega \to \mathscr{C}(F)$ is bijective.*

(b) *If G is quasi-split over F, then $\mathscr{C}(F) = \mathscr{C}(F^{\mathrm{al}})^{\mathrm{Gal}(F^{\mathrm{al}}/F)}$.*

(c) *If F and F' are algebraically closed and $F \subset F'$, then $\mathscr{C}(F) \to \mathscr{C}(F')$ is a bijection.*

(d) *If G is split over F, then $\mathscr{C}(F) \to \mathscr{C}(F^{\mathrm{al}})$ is a bijection.*

PROOF. The first two statements are proved in Kottwitz **[40,** 1.1.3]—the hypothesis there that G^{der} is simply connected is not used in the proof of (a) or (b), and the hypothesis that G is quasi-split is not used in the proof of (a). The remaining statements follow from the first two. □

Let $\mathfrak{c}(X)$ be the $G(\mathbb{C})$-conjugacy class of cocharacters of $G_{\mathbb{C}}$ containing μ_x for $x \in X$. According to (c) of the proposition, $\mathfrak{c}(X)$ corresponds to an

[10]This is the element $x \mapsto x^{q^{-1}}$.

element $\mathfrak{c}(X)_{\mathbb{Q}^{\mathrm{al}}}$ of $\mathscr{C}(\mathbb{Q}^{\mathrm{al}})$. The group $\mathrm{Gal}(\mathbb{Q}^{\mathrm{al}}/\mathbb{Q})$ acts on $\mathscr{C}(\mathbb{Q}^{\mathrm{al}})$, and, by definition, the fixed field of the stabilizer of $\mathfrak{c}(X)_{\mathbb{Q}^{\mathrm{al}}}$ in $\mathscr{C}(\mathbb{Q}^{\mathrm{al}})$ is the reflex field $E = E(G, X)$ of $\mathrm{Sh}(G, X)$.

Consequently E_v is the subfield of E_v^{al} fixed by the stabilizer of $\mathfrak{c}(X)_{E_v^{\mathrm{al}}}$. Because G splits over $\mathbb{Q}_p^{\mathrm{un}}$, (d) of the proposition shows $E_v \subset \mathbb{Q}_p^{\mathrm{un}}$, and because G is quasi-split over $\mathbb{Q}_p$, (b) shows that there exists a cocharacter of G defined over E_v representing $\mathfrak{c}(X)_{E_v^{\mathrm{al}}}$. Hence we have the following result.

COROLLARY 4.7. (a) *The prime v is unramified over p.*

(b) *Let S be a maximal split torus of $G_{E_v^{\mathrm{un}}}$ whose apartment contains the hyperspecial vertex fixed by K_p. Then $\mathfrak{c}(X)_{E_v^{\mathrm{al}}}$ is represented by a cocharacter μ_0 of S.*

NOTATION 4.8. We write B for the completion of E_v^{un}, W for the ring of integers in B, and $\mathbb{F}$ for the residue field of W. Thus $\mathbb{F}$ is an algebraic closure of $\kappa(v)$, W is the ring of Witt vectors over $\mathbb{F}$, and B is the field of fractions of W.

Let $\mathbb{C}_p$ be the completion of an algebraic closure of B, and extend the inclusion $E \hookrightarrow B$ to an isomorphism $\mathbb{C} \to \mathbb{C}_p$. We use this isomorphism to identify B with a subfield of $\mathbb{C}$, and hence to define a fibre functor $\omega_B \colon \mathbf{Mot}(B) \to \mathbf{Vec}_{\mathbb{Q}}$.

We fix a choice of a maximal $\mathbb{Q}_p$-split torus S_0 of $G_{\mathbb{Q}_p}$ whose apartment contains the hyperspecial point fixed by K_p, a maximal B-split torus S of G_B containing S_0, and a cocharacter μ_0 of S representing $\mathfrak{c}(X)_B$.

The points with coordinates in $\mathbb{C}$. By definition

$$\mathrm{Sh}_p(\mathbb{C}) = \varprojlim_{K^p} G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K^p \cdot K_p.$$

Set $G(\mathbb{Z}_{(p)}) = G(\mathbb{Q}) \cap K_p$ and $Z(\mathbb{Z}_{(p)}) = Z(\mathbb{Q}) \cap K_p$ where Z is the centre of G.

LEMMA 4.9. *Let G be a connected reductive group over $\mathbb{Q}$. For any hyperspecial subgroup K_p of $G(\mathbb{Q}_p)$, $G(\mathbb{Q}_p) = G(\mathbb{Q}) \cdot K_p$.*

PROOF. As we noted above, the existence of a hyperspecial subgroup implies that $G_{\mathbb{Q}_p}$ splits over an unramified extension of $\mathbb{Q}_p$. Because K_p is open in $G(\mathbb{Q}_p)$, the lemma follows from the next result. □

LEMMA 4.10. *Let G be a reductive group over $\mathbb{Q}$ that splits over an unramified extension of $\mathbb{Q}_p$. Then $G(\mathbb{Q})$ is dense in $G(\mathbb{Q}_p)$.*

PROOF. The hypothesis says that G acquires a split maximal torus T over an unramified extension of $\mathbb{Q}_p$, and it is known (see [**71**, p. 36]) that T can be chosen to be defined over $\mathbb{Q}_p$. According to Harder [**39**, 5.5.], there will exist a torus $T_0 \subset G$ such that $T_{0\mathbb{Q}_p}$ is conjugate to T (by an element of $G(\mathbb{Q}_p)$). Let E be the smallest extension of $\mathbb{Q}$ over which T_0 splits. It is Galois over $\mathbb{Q}$, and the decomposition group at a prime lying over p is the Galois group of the smallest extension of $\mathbb{Q}_p$ splitting T, which is unramified. This group is therefore cyclic, and we can apply Sansuc [**60,** 3.5ii] to conclude that $G(\mathbb{Q})$ is dense in $G(\mathbb{Q}_p)$. □

PROPOSITION 4.11. *We have*

$$\mathrm{Sh}_p(\mathbb{C}) = G(\mathbb{Z}_{(p)}) \Big\backslash \left(X \times \frac{G(\mathbb{A}_f^p)}{Z^p} \right)$$

where Z^p is the closure of $Z(\mathbb{Z}_{(p)})$ in $Z(\mathbb{A}_f^p)$.

PROOF. It follows from Lemma 4.9 that the natural map

$$G(\mathbb{Z}_{(p)})\backslash X \times \frac{G(\mathbb{A}_f^p)}{K^p} \longrightarrow G(\mathbb{Q})\backslash X \times \frac{G(\mathbb{A}_f^p)}{K^p} \times \frac{G(\mathbb{Q}_p)}{K_p}$$

is a homeomorphism. For K^p open and compact in $G(\mathbb{A}_f^p)$,

$$\begin{aligned} G(\mathbb{Z}_{(p)})\backslash X \times \frac{G(\mathbb{A}_f^p)}{K^p} &= \frac{G(\mathbb{Z}_{(p)})}{Z(\mathbb{Z}_{(p)})} \Big\backslash X \times \frac{G(\mathbb{A}_f^p)}{K^p \cdot Z(\mathbb{Z}_{(p)})} \\ &= \frac{G(\mathbb{Z}_{(p)})}{Z(\mathbb{Z}_{(p)})} \Big\backslash X \times \frac{G(\mathbb{A}_f^p)}{K^p \cdot Z^p} . \end{aligned}$$

Write Γ for the discrete group $G(\mathbb{Z}_{(p)}) \cap K^p$. The image of this in $G^{\mathrm{ad}}(\mathbb{Q})$ acts properly discontinuously on X (cf. Lemma 2.9), and it follows easily that $G(\mathbb{Z}_{(p)})/Z(\mathbb{Z}_{(p)})$ acts properly on $X \times \left(G(\mathbb{A}_f^p)/Z^p\right)$ (see Bourbaki [**10,** III.4.4, Proposition 7]). Hence the quotient space $G(\mathbb{Z}_{(p)})\backslash X \times G(\mathbb{A}_f^p)/Z^p$ is separated [**10,** III.4.2, Proposition 3]. Now we can apply [**10,** III.7.2, Corollary 1] to the compact groups $K^p \cdot Z^p$ acting on the (fixed) separated space $E = G(\mathbb{Z}_{(p)})\backslash X \times G(\mathbb{A}_f^p)/Z^p$ to conclude that

$$E/ \varprojlim (K^p \cdot Z^p) = \varprojlim E/K^p \cdot Z^p .$$

But $\varprojlim (K^p \cdot Z^p) = \cap K^p \cdot Z^p = Z^p$. □

COROLLARY 4.12. *We have*

$$\mathrm{Sh}_p(\mathbb{C}) = G(\mathbb{Q}) \Big\backslash \left(X \times \frac{G(\mathbb{A}_f^p)}{Z^p} \times \frac{G(\mathbb{Q}_p)}{K_p} \right) .$$

PROOF. Again Lemma 4.9 implies that the natural map

$$G(\mathbb{Z}_{(p)})\backslash X \times \frac{G(\mathbb{A}_f^p)}{Z^p} \longrightarrow G(\mathbb{Q})\backslash X \times \frac{G(\mathbb{A}_f^p)}{Z^p} \times \frac{G(\mathbb{Q}_p)}{K_p}$$

is a homeomorphism. □

For $x \in X$, let $I(x)$ be the stabilizer of x in $G(\mathbb{Q})$. Set

$$S(x) = I(x)\backslash X^p(x) \times X_p(x)$$

with $X^p(x) = G(\mathbb{A}_f^p)/Z^p$ and $X_p(x) = G(\mathbb{Q}_p)/K_p$.

COROLLARY 4.13. *There is a canonical bijection*

$$\coprod S(x) \to \mathrm{Sh}_p(\mathbb{C})$$

where the left-hand side runs over a set of representatives for $G(\mathbb{Q})\backslash X$.

PROOF. Trivial. □

We interpret the decomposition in Corollary 4.13 motivically. For this we need to return to the situation of Theorem 3.13, namely, we suppose that $\mathrm{Sh}(G, X)$ is of abelian type, that w_X is defined over $\mathbb{Q}$, and that there is a homomorphism $t: G \to \mathbb{G}_m$ such that $t \circ w_X = -2$. Moreover, we choose a faithful representation $\xi: G \hookrightarrow \mathrm{GL}(V)$ and a family of tensors $\mathfrak{t} = (t_i)_{i \in I}$ such that G is the subgroup of $\mathrm{GL}(V)$ fixing the t_i. We assume that for some $0 \in I$, $\pm t_0$ is a polarization of $(V, \xi \circ h_x)$, for all $x \in X$. Because K_p is a maximal compact subgroup, there exists a lattice $V(\mathbb{Z}_p)$ in $V(\mathbb{Q}_p)$ whose stabilizer is K_p.

Call a pair $N = (M, \mathfrak{s})$ consisting of an abelian motive and a family of tensors *admissible* if there exists an isomorphism $\beta: \omega_B(M) \to V$ mapping each s_i to t_i and sending h_M to h_x, some $x \in X$. Given such a pair, define $X^p(N)$ to be the set of isomorphisms

$$\eta: V(\mathbb{A}_f^p) \to \omega_f^p(M)$$

modulo Z^p-equivalence mapping each t_i to s_i, and let $X_p(N)$ be the set of lattices Λ_p in $\omega_p(M)$ for which there exists an isomorphism $V(\mathbb{Q}_p) \to \omega_p(M)$ mapping each t_i to s_i and $V(\mathbb{Z}_p)$ onto Λ_p. Here ω_ℓ is the ℓ-adic étale fibre functor, and ω_f^p is the restricted product of the ω_ℓ for $\ell \neq p$. Let

$$S(N) = I(N)\backslash X^p(N) \times X_p(N), \qquad I(N) = \mathrm{Aut}(N).$$

The group $G(\mathbb{A}_f^p)$ acts on $S(N)$ according to the following rule:

$$[\eta, \Lambda_p]g = [\eta \circ g, \Lambda_p].$$

COROLLARY 4.14. *There is a canonical equivariant bijection*

$$\coprod S(N) \to \mathrm{Sh}_p(\mathbb{C})$$

where the disjoint union is over a set of representatives for the isomorphism classes of admissible pairs $N = (M, \mathfrak{s})$.

PROOF. Let N be an admissible pair, and choose a β satisfying the above condition. The $G(\mathbb{Q})$-orbit of the element $x \in X$ corresponding to h_M is independent of the choice of β, and in this way we obtain a one-to-one correspondence between the isomorphism classes of admissible pairs and the set $G(\mathbb{Q})\backslash X$ (cf. Corollary 3.13). The choice of a β determines an isomorphism $I(N) \to I(x)$ and equivariant bijections

$$X^p(N) \longrightarrow X^p(x), \qquad \eta \mapsto \beta \circ \eta,$$

$$X_p(N) \longrightarrow X_p(x), \qquad \Lambda_p \mapsto [g] \quad \text{if } \beta(\Lambda_p) = gV(\mathbb{Z}_p),$$

and hence an equivariant bijection

$$S(N) \longrightarrow S(x). \quad \square$$

We interpret the decomposition in Corollary 4.14 in terms of homomorphisms from the motivic Galois group G_{Mab} to G. (Recall that $G_{\mathrm{Mab}} = Aut^{\otimes}(\omega_B)$ where ω_B is the Betti fibre functor on the category of abelian motives.)

If H and G are algebraic groups over a field k and φ and φ' are homomorphisms $H \to G$, we set

$$\mathrm{Isom}(\varphi, \varphi') = \{g \in G(k) \mid \mathrm{ad}(g) \circ \varphi = \varphi'\},$$
$$\mathrm{Aut}(\varphi) = \mathrm{Isom}(\varphi, \varphi).$$

This notation is justified by noting that $\mathrm{Isom}(\varphi, \varphi')$ is the set of isomorphisms from φ to φ' regarded as functors of groupoids.[11]

Call a homomorphism $\varphi: G_{\mathrm{Mab}} \to G$ *admissible* if $\varphi_{\mathbb{R}} \circ h_{\mathrm{Mab}} \in X$. For such a homomorphism, set

$$I(\varphi) = \mathrm{Aut}(\varphi).$$

Let ζ_ℓ be the inclusion $e \hookrightarrow (G_{\mathrm{Mab}})_{\mathbb{Q}_\ell}$ where e is the one-element group scheme, and let ξ_ℓ be the inclusion $e \hookrightarrow G_{\mathbb{Q}_\ell}$. Let $\varphi(\ell)$ be the homomorphism obtained from φ by the base change $\mathbb{Q} \to \mathbb{Q}_\ell$. Define

$$X_\ell(\varphi) = \mathrm{Isom}(\xi_\ell, \zeta_\ell \circ \varphi(\ell)) = \{g \in G(\mathbb{Q}_\ell) \mid \mathrm{ad}\, g \circ \xi_\ell = \varphi(\ell) \circ \zeta_\ell\}.$$

Then $I(\varphi)$ acts on $X_\ell(\varphi)$ on the left, and $G(\mathbb{Q}_\ell)$ acts on it on the right and makes it into a principal homogeneous space. Choose a $\mathbb{Z}$-structure on G, and let $X'_\ell(\varphi)$ be the subset of $X_\ell(\varphi)$ of integral elements. Define $X^p(\varphi)$ to be the restricted product of the $X_\ell(\varphi)$, $\ell \neq p$, relative to the subsets $X'_\ell(\varphi)$. It is independent of the choice of the $\mathbb{Z}$-structure, and it is a principal homogeneous space for the group $G(\mathbb{A}_f^p)$. Define

$$X_p(\varphi) = G(\mathbb{Q}_p)/G(\mathbb{Z}_p),$$

and let

$$S(\varphi) = I(\varphi)\backslash(X^p(\varphi)/Z^p) \times X_p(\varphi).$$

[11] Let H and G be (abstract) groups, and regard them as groupoids in sets, i.e., as categories with a single object and with G and H as the sets of morphisms. A homomorphism $\varphi: H \to G$ of groups can be regarded as a functor $H \to G$, and if φ' is a second homomorphism (functor), then to give a morphism of functors $\varphi \to \varphi'$ is to give an element $g \in G$ such that $\varphi' = \mathrm{ad}(g) \circ \varphi$.

COROLLARY 4.15. *There is a canonical bijection*

$$\coprod S(\varphi) \to \mathrm{Sh}_p(\mathbb{C})$$

where φ *runs over the isomorphism classes of admissible homomorphisms* $G_{\mathrm{Mab}} \to G$.

PROOF. The Betti functor ω_B identifies the category of abelian motives with the category of representations of G_{Mab}. Choose a representation (V, ξ) and tensors t_i as in the discussion preceding Corollary 4.14. An admissible homomorphism φ then defines an admissible pair $N(\varphi)$, and there is a canonical isomorphism $S(\varphi) \to S(N(\varphi))$. Since the map $\varphi \mapsto N(\varphi)$ defines a bijection from the set of isomorphism classes of admissible homomorphisms to the set of isomorphism classes of admissible pairs, (4.15) follows from (4.14). □

Of course, Corollaries 4.14 and 4.15 are clumsy compared to the description (Proposition 4.11) of $\mathrm{Sh}_p(\mathbb{C})$ as a single set of double cosets. However, they are the descriptions that will persist into characteristic p.

The points of $\mathrm{Sh}_p(G, X)$ **with coordinates in** B. Since B is not algebraically closed, in order to have a good description of the points we should assume that the moduli problem is fine. In the present context, this amounts to assuming that $Z(\mathbb{Z}_{(p)})$ is closed in $Z(\mathbb{A}_f^p)$. Then (cf. (3.28) and Corollary 4.14) the points of $\mathrm{Sh}_p(G, X)$ with coordinates in $B \underset{\mathrm{df}}{=} B(\mathbb{F})$ are in one-to-one correspondence with the isomorphism classes of quadruples $(M, \mathfrak{s}, \eta^p, \Lambda_p)$ where:

(4.16.1) M is an abelian motive over B and $\mathfrak{s} = (s_i)_{i \in I}$ is a family of tensors on M for which there exists an isomorphism

$$\beta \colon \omega_B(M) \longrightarrow V(\mathbb{Q})$$

mapping each s_i to t_i and h_M to h_x, some $x \in X$. (Here h_M defines the Hodge structure on $\omega_B(M)$.)

(4.16.2) η^p is an isomorphism

$$\eta^p \colon V(\mathbb{A}_f^p) \longrightarrow \omega_f^p(M)$$

that maps each t_i to s_i and that is invariant under the action of $\mathrm{Gal}(\bar{B}/B)$.

(4.16.3) Λ_p is a $\mathbb{Z}_p$-lattice in $\omega_p(M)$, invariant under the action of $\mathrm{Gal}(\bar{B}/B)$, for which there exists an isomorphism

$$\eta_p \colon V(\mathbb{Q}_p) \longrightarrow \omega_p(M)$$

that maps each t_i to s_i and maps $V(\mathbb{Z}_p)$ onto Λ_p.

REMARK 4.17. (a) The condition (4.16.1) is independent of the choice of the isomorphism $\mathbb{C} \to \mathbb{C}_p$ (extending the embedding $E \hookrightarrow B$) because of Lemma 3.29.

(b) To say that η^p is invariant under the action of $\mathrm{Gal}(\bar{B}/B)$ simply means that $\mathrm{Gal}(\bar{B}/B)$ acts trivially on $\omega_f^p(M)$.

(c) Giving Λ_p is equivalent to giving a K_p-equivalence class of isomorphisms $\eta_p: V(\mathbb{Q}_p) \to \omega_p(M)$ such that each η_p maps each t_i to s_i and such that the class is stable under the action of $\mathrm{Gal}(\bar{B}/B)$.

REMARK 4.18. Fix a quadruple $(M, \mathfrak{s}, \eta^p, \Lambda_p)$, and an isomorphism η_p as in (4.16.3). Use η_p to transfer the action of G on $V(\mathbb{Q}_p)$ to $\omega_p(M)$. If we assume that the p-adic realizations of the s_i are fixed by the action of $\mathrm{Gal}(\bar{B}/B)$, then the action of $\mathrm{Gal}(\bar{B}/B)$ on Λ_p defines a homomorphism $\mathrm{Gal}(\bar{B}/B) \to G(\mathbb{Z}_p)$. Let Q be its image. Every Λ satisfying (4.16.3) is of the form $g\Lambda_p$ for some $g \in G(\mathbb{Q}_p)$, and there is a one-to-one correspondence

$$\{\Lambda\text{'s satisfying (4.16.3)}\} \leftrightarrow \{g \in G(\mathbb{Q}_p)/G(\mathbb{Z}_p) \mid g^{-1}Qg \subset G(\mathbb{Z}_p)\}.$$

We want to pass from the points of $\mathrm{Sh}_p(G, X)$ with coordinates in B to its points in $\mathbb{F}$, via its points in W, but the p-adic étale fibre functor does not persist into characteristic p. Thus we need to re-interpret the data in terms of the de Rham, or crystalline, fibre functor. First we review some of the theory of p-adic cohomology.

Review of p-adic cohomology. Let k be a perfect field, let $W(k)$ be the ring of Witt vectors over k, and let $B(k)$ be the field of fractions of $W(k)$. The absolute Frobenius automorphism $x \mapsto x^p$ and its liftings to $W(k)$ and $B(k)$ are denoted by σ. A *crystal* over k is a free finitely generated $W(k)$-module N together with an injective σ-linear map $\phi : N \to N$; an *isocrystal* over k is a finite-dimensional $B(k)$-vector space N together with a bijective σ-linear map $\phi: N \to N$. For example, if X is a smooth projective variety over k, then $H^i_{\mathrm{crys}}(X)/\{\text{torsion}\}$ is a crystal over k and $H^i_{\mathrm{crys}}(X) \otimes_{W(k)} B(k)$ is an isocrystal. The isocrystal with underlying space $B(k)$ and with $\varphi = p^{-1}\,\mathrm{id}$ is called the *Tate isocrystal.*

The category of k-isocrystals is a nonneutral Tannakian category over $\mathbb{Q}_p$. The dual of an isocrystal N is obtained as follows: regard ϕ as a $B(k)$-linear map ${}^{\sigma}N \to N$, where ${}^{\sigma}N = N \otimes_{B(k),\sigma} B(k)$; form the $B(k)$-linear dual ϕ^t: $N^\vee \to ({}^{\sigma}N)^\vee = {}^{\sigma}(N^\vee)$, and take ϕ on $N^\vee$ to be $\phi^\vee \underset{\mathrm{df}}{=} (\phi^t)^{-1} : {}^{\sigma}N^\vee \to N^\vee$.

Let R be a complete discrete valuation ring of characteristic zero and residue field k, and let K be the field of fractions of R. Identify $W(k)$ and $B(k)$ with subrings of R and K. A *filtered Dieudonné K-module* is an isocrystal (N, ϕ) over k together with a finite decreasing filtration Fil on $N \otimes_{B(k)} K$ such that $\mathrm{Fil}^i(N \otimes K) = N \otimes K$ for i sufficiently small and $\mathrm{Fil}^i(N \otimes K) = 0$ for i sufficiently large. We shall be mainly concerned with filtered Dieudonné $B(k)$-modules, and we usually drop the "Dieudonné". A filtered $B(k)$-module N is said to be *weakly admissible* if it contains a lattice Λ such that

$$\sum p^{-i}\phi(\mathrm{Fil}^i N \cap \Lambda) = \Lambda,$$

and a lattice Λ with this property is said to be *strongly divisible*. The weakly admissible filtered $B(k)$-modules form a neutral Tannakian category with coefficients in $\mathbb{Q}_p$ (see Fontaine **[28]** and Laffaille **[43]**).

Let X be a smooth proper scheme over R. Then (Berthelot and Ogus **[3]**) there is a canonical isomorphism

$$H^i_{\text{dR}}(X) \otimes_R K \overset{\approx}{\to} H^i_{\text{crys}}(X_k) \otimes_{W(k)} K, \qquad X_k = X \times_{\operatorname{Spec} R} \operatorname{Spec} k.$$

Therefore, the Hodge filtration on $H^i_{\text{dR}}(X)$ defines on $H^i_{\text{crys}}(X_k) \otimes B(k)$ the structure of a filtered Dieudonné K-module. For a smooth proper scheme X over W, one even has a canonical isomorphism $H^i_{\text{dR}}(X) \approx H^i_{\text{crys}}(X_k)$ (Berthelot and Ogus **[2**, 7.26]).

If X is a smooth proper scheme over $B(k)$ having good reduction, i.e., extending to a smooth proper scheme $\mathscr{X}$ over W, then the σ-linear map on $H^i_{\text{dR}}(X)$ induced by that on $H^i_{\text{crys}}(\mathscr{X}_k)$ via the isomorphisms

$$H^i_{\text{dR}}(X) \approx H^i_{\text{dR}}(\mathscr{X}) \otimes_{W(k)} B(k) \approx H^i_{\text{crys}}(\mathscr{X}_k) \otimes_{W(k)} B(k)$$

is independent of the choice of $\mathscr{X}$ (Gillet and Messing **[33]**). Therefore, $H^i_{\text{dR}}(X)$ has a canonical structure of a filtered Dieudonné $B(k)$-module.

By a *p-adic representation* of $\operatorname{Gal}(B(k)^{\text{al}}/B(k))$ we mean a continuous representation of $\operatorname{Gal}(B(k)^{\text{al}}/B(k))$ on a finite-dimensional $\mathbb{Q}_p$-vector space. Let B_{crys} be the topological $B(k)$-algebra defined in Fontaine **[29]**; it has a continuous action of $\operatorname{Gal}(B(k)^{\text{al}}/B(k))$ and a decreasing filtration. When N is a weakly admissible filtered $B(k)$-module, we set

$$\mathbb{V}(N) = \{x \in \operatorname{Fil}^0(B_{\text{crys}} \otimes_{B(k)} N) \mid \phi(x) = x\}.$$

This is a p-adic representation of $\operatorname{Gal}(B(k)^{\text{al}}/B(k))$ with dimension at most that of N; if $\dim_{\mathbb{Q}_p} \mathbb{V}(N) = \dim_{B(k)} N$, then N is said to be *admissible*.

When V is a p-adic representation, we set

$$\mathbb{D}(V) = (B(k)_{\text{crys}} \otimes_{\mathbb{Q}_p} V)^{\operatorname{Gal}(B(k)^{\text{al}}/B(k))}.$$

This is a filtered $B(k)$-module of dimension at most that of V; if $\dim_{B(k)} \mathbb{D}(V) = \dim_{\mathbb{Q}_p} V$, then V is said to be *crystalline*.

PROPOSITION 4.19. *The functor $\mathbb{D}$ defines a $\otimes$-equivalence from the category of crystalline p-adic representations to that of admissible filtered Dieudonné modules, with quasi-inverse $\mathbb{V}$.*

PROOF. See Fontaine **[28, 29]** and Fontaine and Laffaille **[31]**. □

THEOREM 4.20. *Let X be a nonsingular projective variety over $B(k)$ with good reduction; then $H^i_{\text{dR}}(X) \otimes_{W(k)} B(k)$ is an admissible filtered $B(k)$-module, $H^i_{\text{ét}}(X) \underset{\text{df}}{=} H^i_{\text{ét}}(X_{B^{\text{al}}}, \mathbb{Q}_p)$ is a crystalline p-adic representation, and there are natural isomorphisms*

$$\mathbb{V}(H^i_{\text{crys}}(X) \otimes_{W(k)} B(k)) \approx H^i_{\text{ét}}(X), \qquad \mathbb{D}(H^i_{\text{ét}}(X)) \approx H^i_{\text{crys}}(X).$$

PROOF. See Fontaine and Messing [32] and Faltings [26]. □

REMARK 4.21. (a) Let Λ be a lattice in a filtered $B(k)$-module N, and suppose $\mu\colon \mathbb{G}_m \to \mathrm{GL}(\Lambda)$ splits the filtration on Λ, i.e., if we set

$$N^i = \{n \in N \mid \mu(x)n = x^i n \text{ all } x \in B(k)^\times\},$$

then

$$\mathrm{Fil}^p N = \bigoplus_{i \geq p} N^i, \qquad \Lambda = \bigoplus_i (\Lambda \cap N^i).$$

The condition for Λ to be strongly divisible then becomes

$$\phi\Lambda = \mu(p)\Lambda. \tag{4.21.1}$$

(b) Let N be a weakly admissible filtered $B(k)$-module. In Wintenberger [73] it is shown that there is a *canonical* splitting $\mu_W\colon \mathbb{G}_m \to \mathrm{GL}(N)$ of the filtration on N. When k is algebraically closed, a lattice Λ in N is strongly divisible if and only if μ_W splits the filtration on Λ and (4.21.1) holds.

The points of $\mathrm{Sh}_p(G, X)$ with coordinates in B: Crystalline interpretation. The points of $\mathrm{Sh}_p(G, X)$ with coordinates in $B \underset{\mathrm{df}}{=} B(\mathbb{F})$ are in one-to-one correspondence with the set of isomorphism classes of quadruples $(M, \mathfrak{s}, \eta^p, \Lambda_{\mathrm{crys}})$ where $(M, \mathfrak{s})$ is as in (4.16.1), η^p is as in (4.16.2), and

(4.22) Λ_{crys} is a strongly divisible lattice in $\omega_{\mathrm{dR}}(M)$ for which there exists an isomorphism

$$\eta_{\mathrm{dR}}\colon V(B) \longrightarrow \omega_{\mathrm{dR}}(M)$$

that maps each t_i to s_i, maps $V(W)$ onto Λ_{crys}, and makes the filtration $\mathrm{Filt}(\mu_0^{-1})$ correspond to the Hodge filtration on $\omega_{\mathrm{dR}}(M)$

EXPLANATION 4.23. We give a *heuristic* explanation of how to pass from (4.16) to (4.22).

Let $\Gamma = \mathrm{Gal}(\bar{B}/B)$. We noted above that $\mathbb{D}$ and $\mathbb{V}$ define a $\otimes$-equivalence between the category of crystalline representations of Γ and that of admissible filtered Dieudonné $B(k)$-modules. In Fontaine and Laffaille [31], it is shown that a weakly admissible filtered $B(k)$-module is admissible if the length of its filtration is $< p - 1$. Hence $\mathbb{V}$ defines a $\otimes$-equivalence from the category of (weakly) admissible filtered $B(k)$-modules generated by those with length $< p - 1$ to a category of crystalline representations of Γ. It is known that this equivalence underlies an equivalence between a category of strongly divisible lattices and a category of Γ-stable lattices in crystalline representations. See Fontaine [30], 2.3.

Let $\sigma = \mu_W(p^{-1}) \circ \phi$. Then σ defines a $\mathbb{Q}_p$-structure on any weakly admissible filtered $B(\mathbb{F})$-module i.e., if we set

$$N^{\sigma=1} = \{x \in N \mid \sigma x = x\}$$

then $N^{\sigma=1} \otimes_{\mathbb{Q}_p} B(\mathbb{F}) = N$. Moreover, (4.21.1) implies that, for any strongly divisible lattice Λ in N, $\Lambda^{\sigma=1}$ is a $\mathbb{Z}_p$-structure on Λ.

Let $\mathbf{Rep}^{\mathrm{crys}}(\Gamma)$ be the (Tannakian) category of crystalline representations of Γ. There are two fibre functors on $\mathbf{Rep}^{\mathrm{crys}}(\Gamma)$ over $\mathbb{Q}_p$, namely, the forgetful functor and the functor $H \mapsto \mathbb{D}(H)^{\sigma=1}$. Assume that the torsor relating the two is trivial (cf. Wintenberger **[73,** 4.2.5]), and choose a trivialization. Then we can identify $\mathbb{D}(\mathbb{V}(H))$ with $H^{\sigma=1} \otimes B$ endowed with the filtration defined by μ_W and with ϕ acting as $x \mapsto \mu_W(p) \cdot \sigma x$.

Given Λ_p satisfying (4.16.3), define $\Lambda_{\mathrm{crys}} = \Lambda_p \otimes W$ and $\eta_{\mathrm{dR}} = \eta_p \otimes 1$. Then η_{dR} maps each t_i to s_i and it maps $V(W)$ onto Λ_{crys}. Note that these properties determine it uniquely up to conjugation by an element of $G(W)$. When we make a base change by $B \to \mathbb{C}$, we obtain a homomorphism

$$\eta_{\mathrm{dR}} \otimes 1 \colon V(\mathbb{C}) \longrightarrow \omega_{\mathrm{dR}}(M_{\mathbb{C}})$$

carrying each t_i into s_i. By (4.16.1), we already have such a map, namely, $\beta^{-1} \otimes 1$, which, moreover, has the property that it maps μ_x, some $x \in X$, to μ_M. Since the two maps differ by conjugation with an element of $G(\mathbb{C})$, this shows that $\eta_{\mathrm{dR}} \otimes 1$ maps $\mathfrak{c}(X)$ into the conjugacy class of μ_M. By definition, μ_M^{-1} splits the Hodge filtration on $\omega_{\mathrm{dR}}(M_{\mathbb{C}})$, and it follows that, after possibly replacing it with a $G(W)$-conjugate, η_{dR} will map $\mathrm{Filt}(\mu_0^{-1})$ to the Hodge filtration on $\omega_{\mathrm{dR}}(M)$.

Conversely, given Λ_{crys} we can define $\Lambda_p = (\Lambda_{\mathrm{crys}})^{\sigma=1}$.

Integral canonical models. As we saw in (4.16), in the case that $\mathrm{Sh}_p(G, X)$ is a fine moduli variety, its points with coordinates in B parametrize certain quadruples $(M, \mathfrak{s}, \eta^p, \Lambda_p)$. The existence of η^p implies that $\mathrm{Gal}(\overline{B}/B)$ acts trivially on $\omega_f^p(M)$, and this should imply that the motive has good reduction. In fact, the whole quadruple should extend over W, and so we should have $\mathrm{Sh}_p(W) = \mathrm{Sh}_p(B)$. In Milne **[50,** §2], this intuition is turned into the definition of a canonical model of $\mathrm{Sh}_p(G, X)$ over W. In order to achieve uniqueness, it is necessary to specify the points of the model, not just in W, but in very large, not necessarily Noetherian, W-algebras.

DEFINITION 4.24. A *model* of $\mathrm{Sh}_p(G, X)$ over $\mathscr{O}_v$ is a scheme S over $\mathscr{O}_v$ together with a continuous action of $G(\mathbb{A}_f^p)$ and a $G(\mathbb{A}_f^p)$-equivariant isomorphism

$$\gamma \colon S \otimes_{\mathscr{O}_v} E_v \to \mathrm{Sh}_p(G, X)_{E_v}$$

(of pro-varieties over E_v). Such a model is said to be *smooth* if there is a compact open subgroup K_0 of $G(\mathbb{A}_f^p)$ such that S_K is smooth over $\mathscr{O}_v$ for all $K \subset K_0$ and $S_{K'}$ is étale over S_K for all $K' \subset K \subset K_0$. Such a model is said to have the *extension property* if, for every regular scheme Y (not necessarily Noetherian) over $\mathscr{O}_v$ such that Y_{E_v} is dense in Y, every E_v-morphism $Y_{E_v} \to S_{E_v}$ extends uniquely to an $\mathscr{O}_v$-morphism $Y \to S$. An *integral canonical model* of $\mathrm{Sh}_p(G, X)$ is a smooth model over $\mathscr{O}_v$ with the extension property.

CONJECTURE 4.25. *The variety* $\mathrm{Sh}_p(G, X)$ *always has an integral canonical model.*

In Milne [**50**], the following results are proved:

(4.26) The integral canonical model of $\mathrm{Sh}_p(G, X)$, if it exists, is uniquely determined up to a unique isomorphism.

(4.27) The Siegel modular variety $\mathrm{Sh}_p(G(\psi), X(\psi))$ has an integral canonical model,[12] namely, the moduli scheme constructed in Mumford [**55**].

(4.28) Consider an inclusion $(G, X) \hookrightarrow (G', X')$ of pairs satisfying the axioms (SV0–2). Let K_p' be a hyperspecial subgroup of $G'(\mathbb{Q}_p)$, and assume that $K_p \underset{\mathrm{df}}{=} K_p' \cap G(\mathbb{Q}_p)$ is hyperspecial in $G'(\mathbb{Q}_p)$. Then (cf. Proposition 3.24) there is a closed immersion
$$\mathrm{Sh}_p(G, X) \hookrightarrow \mathrm{Sh}_p(G', X')$$
defined over $E(G, X)$. If $\mathrm{Sh}_p(G', X')$ has a model S' over $\mathscr{O}_v$ with the extension property, then the closure of $\mathrm{Sh}_p(G, X)$ in S' also has the extension property (and hence will be canonical if it is smooth).

We offer two further remarks.

REMARK 4.29. Let $\mathrm{Sh}(G, X)$ be a Shimura variety of PEL-type, so that $\mathrm{Sh}_p(G, X)$ solves a moduli problem over $E(G, X)$ classifying isomorphism classes of triples consisting of a polarized abelian variety, an identification of the endomorphism algebra of the abelian variety with a fixed algebra, and a level structure, all satisfying certain conditions. (See Milne [**50**, 1.1] for a precise definition.) There is a homomorphism $G \hookrightarrow G(\psi)$ sending X into $X(\psi)$, and hence a closed immersion $\mathrm{Sh}_p(G, X) \hookrightarrow \mathrm{Sh}_p(G(\psi), X(\psi))$ into the Siegel modular variety for any hyperspecial subgroup K_p' of $G(\psi)(\mathbb{Q}_p)$ such that $K_p = G(\mathbb{Q}_p) \cap K_p'$. It is more-or-less known that the closure of $\mathrm{Sh}_p(G, X)$ in the integral canonical model of $\mathrm{Sh}_p(G(\psi), X(\psi))$ represents a smooth functor, and therefore is itself smooth (see for example, Langlands and Rapoport [**46**, 6.2]; corrected in Kottwitz [**42**, §5]). Together with (4.28), this proves that $\mathrm{Sh}_p(G, X)$ has an integral canonical model.

REMARK 4.30. The validity of Conjecture 4.25 does not depend on $Z(G)$, at least if G^{der} is simply connected. To prove this, we need to make use of connected Shimura varieties (Deligne [**18**]; Milne [**49**, II.1]), and a descent theorem (Bosch et al. [**9**, 6.2, Proposition C.1]).

The descent theorem says the following: The functor that associates with an $\mathscr{O}_v$-scheme S the triple (S_1, S_2, θ) consisting of the E_v-scheme $S_1 \underset{\mathrm{df}}{=}$

[12] The proof of this uses, in an essential way, Theorem V.6.8 of [27]. Recently I have learned that Gabber has a counterexample to this theorem, but the example appears to require ramification. Thus it is possible that the theorem of Faltings and Chai remains valid over the Witt vectors, at least for $p \neq 2$. If not, the definition of an integral canonical model will have to be modified.

$S \otimes_{\mathscr{O}_v} E_v$, the W-scheme $S_2 \underset{\text{df}}{=} S \otimes_{\mathscr{O}_v} W$, and the canonical isomorphism $\theta\colon S_1 \otimes_{E_v} B \to S_2 \otimes_W B$ is fully faithful. Its essential image consists of all triples (S_1, S_2, θ) that admit a quasi-affine open covering (in an obvious sense).

Now consider a system (G, X, K_p) as before, and suppose that G^{der} is simply connected. The composite of $h \in X$ with $G \to G^{\text{ab}}$ is independent of h—we write it h_X. The map $G \to G^{\text{ab}}$ defines a surjection

$$\text{Sh}(G, X) \to \text{Sh}(G^{\text{ab}}, h_X),$$

which, for simplicity, we assume identifies $\pi_0(\text{Sh}(G, X))$ with $\text{Sh}(G^{\text{ab}}, h_X)$ (in general, $\pi_0(\text{Sh}(G, X))$ will be a finite covering of $\text{Sh}(G^{\text{ab}}, h_X)$; see Deligne [**15**, 2.7.1]). The inverse image of $e = [h^{\text{ab}}, 1]$ in $\text{Sh}(G, X)$ can be identified with the connected Shimura variety $\text{Sh}^0(G^{\text{der}}, X^+)$ for a suitable connected component X^+ of X.

On passing to the quotient by K_p we obtain a surjection

$$\text{Sh}_p(G, X) \to \text{Sh}_p(G^{\text{ab}}, h_X)$$

whose fibre over $[h_X, 1]$ we denote by $\text{Sh}_p^0(G^{\text{der}}, X^+)$. Because G is unramified at p, G^{ab} splits over B, and all the points of $\text{Sh}_p(G^{\text{ab}}, h_X)$ are rational over B (cf. Milne [**50**, 2.16]).

Now suppose $\text{Sh}_p(G, X)$ has an integral canonical model S over $\mathscr{O}_v$. From S_2 we obtain a model of the connected Shimura variety $\text{Sh}_p^0(G^{\text{der}}, X^+)$ over W that is smooth and has the extension property (in the sense of Definition 4.24), and S_2 can be recovered from this integral canonical model by "induction", i.e., by translating it by elements of $G(\mathbb{A}_f^p) \times Z(\mathbb{Q}_p)$. Thus, by using the descent theorem, we see that S can be recovered from the integral canonical model of $\text{Sh}_p^0(G^{\text{der}}, X^+)$ over W.

Suppose now that $\text{Sh}_p^0(G^{\text{der}}, X^+)$ arises in the same way from a second variety $\text{Sh}_p(G_1, X_1)$. From the above discussion we see that, if $\text{Sh}_p(G_1, X_1)$ has an integral canonical model, then so also does $\text{Sh}_p(G, X)$.

EXAMPLE 4.31. Let B be a quaternion algebra over a totally real field F, split at the real primes v_i for $1 \le i \le r$ and nonsplit at the real primes v_i for $i > r$. Denote its canonical involution by $z \mapsto \bar{z}$. Let V be a free B-module endowed with a nondegenerate symmetric F-bilinear form Φ such that

$$\Phi(bx, y) = \Phi(x, \bar{b}y), \qquad b \in B,\ x, y \in V.$$

Assume that for $i > r$, the form defined by Φ on $V \otimes_{F, v_i} \mathbb{R}$ is positive-definite. Define G to be the reductive group over $\mathbb{Q}$ such that

$$G(\mathbb{Q}) = \{g \in \text{GL}_B(V) \mid \Phi(gx, gy) = \mu(g)\Phi(x, y),\ \text{some } \mu(g) \in F^\times\}.$$

There is a unique $G(\mathbb{R})$-conjugacy class X of homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$ satisfying (SV1,2). The Shimura variety $\mathrm{Sh}(G, X)$ is not a moduli variety if $r \neq [F:\mathbb{Q}]$ because its weight is not defined over $\mathbb{Q}$. However, there is a Shimura variety of PEL-type $\mathrm{Sh}(G_*, X_*)$ such that $(G, X)^+ = (G_*, X_*)^+$ (see Deligne [**15,** §6]). Therefore, we can apply the preceding remarks to show that $\mathrm{Sh}_p(G, X)$ has an integral canonical model.

Henceforth, we assume that $\mathrm{Sh}_p(G, X)$ has an integral canonical model, which we again denote $\mathrm{Sh}_p(G, X)$.

The points of $\mathrm{Sh}_p(G, X)$ with coordinates in W. From the definition of the integral canonical model, the points of $\mathrm{Sh}_p(G, X)$ with coordinates in W are the same as those with coordinates in B, but we shall need a second interpretation.

The following conditions on a motive M over B should be equivalent:

(4.32.1) the action of $\mathrm{Gal}(B^{\mathrm{al}}/B)$ on $\omega_f^p(M)$ is trivial;
(4.32.2) the filtered Dieudonné module $\omega_{\mathrm{dR}}(M)$ is admissible;
(4.32.3) M has good reduction.

We assume the equivalence of (4.32.1) and (4.32.2) for abelian motives, and use (4.32.3) to define good reduction. It is probably too optimistic to expect that the conditions imply that M extends to a motive over W constructed from smooth schemes over W. Rather, one expects that M will extend to a "log-smooth" motive over W whose reduction is smooth. If M is a motive over B with a $\mathbb{Z}_p$-integral structure, so that $\omega_{\mathrm{dR}}(M)$ is a W-module, then one expects it to be a strongly divisible lattice in $\omega_{\mathrm{dR}}(M) \otimes \mathbb{Q}$ when $M \otimes \mathbb{Q}$ has good reduction and the length of the filtration on $\omega_{\mathrm{dR}}(M) \otimes \mathbb{Q}$ is $< p$.

Define $\mathbf{Mot}^{\mathrm{ab}}(W)$ to be the subcategory of $\mathbf{Mot}^{\mathrm{ab}}(B)$ satisfying the conditions (4.32), and let $\mathbf{Mot}(\mathbb{F})$ be the category of motives over $\mathbb{F}$, defined as in (Milne 1993, §1). We shall need to assume that the étale and crystalline fibre functors are defined on $\mathbf{Mot}(\mathbb{F})$ and that there is a "reduction" functor of tensor categories

$$M \mapsto \bar{M} \colon \mathbf{Mot}^{\mathrm{ab}}(W) \longrightarrow \mathbf{Mot}(\mathbb{F})$$

such that

(4.33.1) $\omega_f^p(\bar{M}) = \omega_f^p(M)$; $\omega_{\mathrm{crys}}(\bar{M}) = \omega_{\mathrm{dR}}(M)$;
(4.33.2) the functor $M \mapsto (\bar{M}, \omega_{\mathrm{dR}}(M))$ is fully faithful (here $\omega_{\mathrm{dR}}(M)$ is regarded as a filtered Dieudonné B-module).

Of course, we also expect that the functor $M \mapsto \bar{M}$ and the isomorphisms implicit in (4.33.1) are compatible with those on abelian varieties.

With the above assumptions, we see that the points of $\mathrm{Sh}_p(G, X)$ with coordinates in W are in natural one-to-one correspondence with the set of isomorphism classes of quintuples $(M, \mathfrak{s}, F, \eta^p, \Lambda_{\mathrm{crys}})$ consisting of a motive M over $\mathbb{F}$, a family $\mathfrak{s}$ of tensors on M, a filtration F on $\omega_{\mathrm{crys}}(M)$, an isomorphism $\eta^p\colon V(\mathbb{A}_f^p) \to \omega_f^p(M)$, and a lattice $\Lambda_{\text{crys}} \subset \omega_{\text{crys}}(M)$ satisfying the following condition:

(4.34) there is a quadruple $(\widetilde{M}, \tilde{\mathfrak{s}}, \tilde{\eta}^p, \widetilde{\Lambda}_{\text{crys}})$ satisfying (4.16.1), (4.16.2), and (4.22) that maps to $(M, \mathfrak{s}, F, \eta^p, \Lambda_{\text{crys}})$ under the reduction functor.

The points of $\text{Sh}_p(G, X)$ **with coordinates in** $\mathbb{F}$**.** Because $\text{Sh}_p(G, X)$ is smooth over W, the reduction map

$$\text{Sh}_p(G, X)(W) \longrightarrow \text{Sh}_p(G, X)(\mathbb{F})$$

is surjective. When we interpret $\text{Sh}_p(G, X)(W)$ as in (4.34), this map should correspond to the reduction of motives. Assume this. Then the points of $\text{Sh}_p(G, X)$ with coordinates in $\mathbb{F}$ are in natural one-to-one correspondence with the set of isomorphism classes of quadruples $(M, \mathfrak{s}, \eta^p, \Lambda_{\text{crys}})$ consisting of a motive M over $\mathbb{F}$, a family $\mathfrak{s}$ of tensors on M, an isomorphism $\eta^p\colon V(\mathbb{A}_f^p) \to \omega_f^p(M)$, and a lattice $\Lambda_{\text{crys}} \subset \omega_{\text{crys}}(M)$ satisfying the following condition:

(4.35) there exists a filtration F on $\omega_{\text{crys}}(M)$ and a quadruple $(\widetilde{M}, \tilde{\mathfrak{s}}, \tilde{\eta}^p, \widetilde{\Lambda}_{\text{crys}})$ satisfying (4.16.1), (4.16.2), and (4.22) that maps to $(M, \mathfrak{s}, F, \eta^p, \Lambda_{\text{crys}})$ under the reduction functor.

Call a pair $N = (M, \mathfrak{s})$ *admissible* if there exists a pair $(\eta^p, \Lambda_{\text{crys}})$ such that $(M, \mathfrak{s}, \eta^p, \Lambda_{\text{crys}})$ satisfies the condition (4.35).

Fix N, and let $S(N)$ be the set of quadruples $(M, \mathfrak{s}, \eta^p, \Lambda_{\text{crys}})$ such that $(M, \mathfrak{s}) \approx N$. Then

$$\text{Sh}_p(\mathbb{F}) = \coprod S(N)$$

where the disjoint union is over a set of representatives for the isomorphism classes of the admissible pairs. The actions of $G(\mathbb{A}_f^p)$ and Φ preserve $S(N)$.

Let $I(N) = \text{Aut}(M, \mathfrak{s})$. Define $X^p(N)$ to be the set of isomorphisms $\eta^p\colon V(\mathbb{A}_f^p) \to \omega_f^p(M)$ carrying each t_i to s_i, and define $X_p(N)$ to be the set of lattices Λ_{crys} in $\omega_{\text{crys}}(M)$ for which there exists a filtration F on $\omega_{\text{crys}}(M)$ such that $(M, \mathfrak{s}, F, \Lambda_{\text{crys}})$ is the reduction of a triple $(\widetilde{M}, \tilde{\mathfrak{s}}, \Lambda_{\text{crys}})$ satisfying (4.22). There is an action of $I(N)$ on $X^p(N) \times X_p(N)$ on the left, an action of $G(\mathbb{A}_f^p)$ on $X^p(N)$, an action of $Z(\mathbb{Q}_p)$ on $X_p(N)$, and an action of Φ on $X_p(N)$.

PROPOSITION 4.36. *With the above assumptions, there is a canonical bijection*

$$S(N) \approx I(N)\backslash X^p(N) \times X_p(N)$$

compatible with the actions of $G(\mathbb{A}_f^p)$ *and* Φ.

PROOF. Obvious. □

Let $N = (M, \mathfrak{s})$ be admissible. Choose an isomorphism

$$\beta\colon \omega_{\text{crys}}(M) \longrightarrow V(B)$$

sending s_i to t_i for all i. The map

$$x \mapsto \beta\phi\beta^{-1}(\sigma^{-1}x)\colon V(B) \to V(B)$$

is linear, and it maps t_i to t_i for all i. Therefore, it is multiplication by an element $b \in G(B)$. This b is the unique element of $G(B)$ such that

$$\beta\phi(y) = b\sigma\beta(y), \quad \text{all } y \in \omega_{\text{crys}}(M).$$

If we replace β with $g \circ \beta$, $g \in G(B)$, then b is replaced by its σ-conjugate $gb(\sigma g)^{-1}$.

Let $\Lambda_{\text{crys}} \in X_p(N)$. According to our assumption, there exists an isomorphism

$$\eta_p\colon V(B) \longrightarrow \omega_{\text{crys}}(M)$$

mapping each t_i to s_i, mapping $V(W)$ onto Λ_{crys}, and such that Λ_{crys} is strongly divisible for the filtration defined by μ_0^{-1}. The composite $\beta \circ \eta_p$ fixes each t_i and, hence, is multiplication by $g \in G(B)$. The situation is summarized by the following diagram:

$$\begin{array}{ccccc}
V(B) & \xrightarrow{\eta_p} & \omega_{\text{crys}}(M) & \xrightarrow{\beta} & V(B) \\
t_i & \leftrightarrow & s_i & \leftrightarrow & t_i \\
V(W) & \xrightarrow{\text{onto}} & \Lambda_{\text{crys}} & \xrightarrow{\text{onto}} & gV(W) \\
g^{-1}b\sigma g & \leftrightarrow & \phi & \leftrightarrow & b\sigma \\
\text{Filt}(\mu_0^{-1}) & \leftrightarrow & F & \leftrightarrow & \text{Filt}(g\mu_0^{-1}).
\end{array}$$

The vector space $V(B)$ endowed with the σ-linear map $x \mapsto g^{-1}b\sigma(gx)$ and the filtration $\text{Filt}(\mu_0^{-1})$ is a weakly admissible filtered module, with $V(W)$ as a strongly divisible module. Since μ_0^{-1} splits the filtration on $V(W)$, according to (4.21.1) this last condition means that

$$(g^{-1}b\sigma g)V(W) = \mu_0(p^{-1})V(W).$$

But the stabilizer in $G(B)$ of $V(W)$ is $G(W)$, and so

$$g^{-1}b\sigma g \in G(W) \cdot \mu_0(p^{-1}) \cdot G(W).$$

PROPOSITION 4.37. *The map* $\Lambda_{\text{crys}} \mapsto g \cdot G(W)$ *defines a bijection*

$$X_p(N) \longrightarrow \{g \cdot G(W) \in G(B)/G(W) \mid g^{-1} \cdot b \cdot \sigma g \in G(W) \cdot \mu_0(p^{-1}) \cdot G(W)\}.$$

PROOF. Straightforward. □

We now restate our results in terms of the groupoid attached to the category of motives over $\mathbb{F}$. Let w be the extension of v to $\mathbb{Q}^{\mathrm{al}}$ induced by the inclusion $\mathbb{Q}^{\mathrm{al}} \hookrightarrow \mathbb{C}_p$, and let $\mathbb{Q}^w = \mathbb{Q}^{\mathrm{al}} \cap B$. The category $\mathbf{Mot}(\mathbb{F})$ has a fibre functor over $\mathbb{Q}^{\mathrm{al}}$ (Milne **[51]**), and the obstruction to it having a fibre functor over $\mathbb{Q}^w$ lies in $H^2(\mathbb{Q}^w, P)$ where P is the Weil number pro-torus **[51,** §2]. But a theorem of Lang shows that $\mathbb{Q}^w$ is a C_1-field (Shatz **[65,** p. 116, Theorem 27], and so $H^2(\mathbb{Q}^w, P) = 0$. Thus $\mathbf{Mot}(\mathbb{F})$ has a fibre functor over $\mathbb{Q}^w$. We choose one, and let $\mathfrak{M}$ be the corresponding $\mathbb{Q}^w/\mathbb{Q}$-groupoid. For each $\ell \neq p, \infty$ étale cohomology provides a fibre functor ω_ℓ over $\mathbb{Q}_\ell$, and correspondingly we obtain a morphism of groupoids

$$\zeta_\ell : \mathfrak{G}_\ell \longrightarrow \mathfrak{M}(\ell)$$

where $\mathfrak{G}_\ell$ is the trivial $\mathbb{Q}_\ell^{\mathrm{al}}/\mathbb{Q}_\ell$-groupoid and $\mathfrak{M}(\ell)$ is the $\mathbb{Q}_\ell^{\mathrm{al}}/\mathbb{Q}_\ell$-groupoid obtained from $\mathfrak{M}$ by base change. Let $\mathfrak{G}_p$ be the $B/\mathbb{Q}_p$-groupoid attached to the category of isocrystals over $\mathbb{F}$ and the forgetful functor. Then ω_{crys} defines a morphism of groupoids

$$\zeta_p : \mathfrak{G}_p \longrightarrow \mathfrak{M}(p).$$

Finally, there is a morphism of groupoids over $\mathbb{R}$,

$$\zeta_\infty : \mathfrak{G}_\infty \longrightarrow \mathfrak{M}(\infty)$$

(Milne **[51,** 3.29]). Let $\mathfrak{G}_G$ be the $\mathbb{Q}^w/\mathbb{Q}$-groupoid defined by G. With each homomorphism $\varphi : \mathfrak{M} \to \mathfrak{G}_G$ and representation of G, there is associated a motive $N(\varphi)$ over $\mathbb{F}$ endowed with a family of tensors, and we say that φ is *admissible* if $N(\varphi)$ is admissible in the above sense for one (hence every) faithful representation of G. For an admissible φ, define

$$I(\varphi) = \mathrm{Aut}(\varphi) \underset{\mathrm{df}}{=} \{g \in G(\mathbb{Q}^w) \mid \mathrm{ad}(g) \circ \varphi = \varphi\}$$

$X^p(\varphi) = \prod_{\ell \neq p} X_\ell(\varphi)$ (restricted product), where $X_\ell(\varphi) = \mathrm{Isom}(\xi_\ell, \varphi(\ell) \circ \zeta_\ell)$. Here $\varphi(\ell)$ is obtained from φ by base change, and ξ_ℓ is the obvious morphism $\mathfrak{G}_\ell \to \mathfrak{G}_G(\ell)$ (see Milne **[50,** p. 186], for more details). Choose an $s(\sigma)$ in $\mathfrak{G}_p$ mapping to $\sigma \in \mathrm{Gal}(\mathbb{F}/\mathbb{F}_p)$, and let $(\varphi(p) \circ \zeta_p)(s(\sigma)) = (b, \sigma)$. Define

$$X_p(\varphi) = \{g \cdot G(W) \in G(B)/G(W) \mid g^{-1} \cdot b \cdot \sigma g \in G(W) \cdot \mu_0(p^{-1}) \cdot G(W)\}.$$

As in **[50,** p. 188] we can define an operator Φ on $X_p(\varphi)$. Let

$$S(\varphi) = I(\varphi)\backslash X^p(\varphi) \times X_p(\varphi),$$

with $G(\mathbb{A}_f^p)$ acting through its action on $X^p(\varphi)$ and Φ acting through its action on $X_p(\varphi)$.

THEOREM 4.38. *With the above assumptions, there is an isomorphism of sets with actions*

$$(\mathrm{Sh}_p(\mathbb{F}), \times, \Phi) \longrightarrow \coprod_{\varphi} (S(\varphi), \times(\varphi), \Phi(\varphi))$$

where the disjoint union is over a set of representatives for the isomorphism classes of admissible homomorphisms $\mathfrak{M} \to \mathfrak{G}_G$.

PROOF. The choice of a faithful representation of G determines a bijection $\varphi \mapsto N(\varphi)$ from the set of isomorphism classes of admissible φ's to the set of isomorphism classes of admissible pairs. Moreover, using (4.35), one sees that

$$(S(\varphi), \times(\varphi), \Phi(\varphi)) = (S(N), \times(N), \Phi(N)),$$

and so it follows from Proposition 4.36 and the discussion preceding it that there is a bijection $\mathrm{Sh}_p(\mathbb{F}) \to \coprod S(\varphi)$ compatible with the action of $G(\mathbb{A}_f^p)$. Checking that the actions of Φ agree is straightforward (it is similar to the proof of the last step of the proof of Theorem 3.21). □

Statement of the conjecture. We now drop all unproven assumptions, and state a conjecture. Let $(\mathfrak{P}, (\zeta_\ell))$ be a system as in Milne **[51,** 3.31], except that now $\mathfrak{P}$ is a $\mathbb{Q}^w/\mathbb{Q}$-groupoid (rather than a $\mathbb{Q}^{\mathrm{al}}/\mathbb{Q}$-groupoid).

Given a homomorphism $\varphi: \mathfrak{P} \to \mathfrak{G}_G$, we can define a set $S(\varphi)$ with an action of $G(\mathbb{A}_f^p)$ and a commuting action of a Frobenius element Φ exactly as in the last section (see also Milne **[51,** §4]). The conjecture will then state that there is an isomorphism of sets with operators

$$(\mathrm{Sh}_p(\mathbb{F}), \times, \Phi) \longrightarrow \coprod_{\varphi} (S(\varphi), \times(\varphi), \Phi(\varphi))$$

where the disjoint union is over a set of representatives for the isomorphism classes of "admissible" homomorphisms $\varphi: \mathfrak{P} \to \mathfrak{G}_G$. The only remaining problem in stating the conjecture is to define "admissible".

Necessary local conditions. There are some obvious necessary conditions for φ to be admissible.

$(4.39)_\ell$ The set $X_\ell(\varphi)$ is nonempty.

$(4.39)_p$ The set $X_p(\varphi)$ is nonempty.

$(4.39)_\infty$ The homomorphism $\varphi(\infty)\circ\zeta_\infty: \mathfrak{G}_\infty \to \mathfrak{G}_G(\infty)$ of $\mathbb{C}/\mathbb{R}$-groupoids is isomorphic to that defined by X (see Milne **[50,** 4.5]).

The case of a torus. Consider a pair (T, h) satisfying the conditions (SV0, 1, 2) with T a torus. There is a unique homomorphism $\rho(h): S \to T$ such that $\rho(h)_{\mathbb{R}} \circ h_{\mathrm{can}} = h$. As is explained in Milne **[51,** §4], there is a canonical homomorphism $\mathfrak{P}' \to \mathfrak{G}_S$, and we write φ_h for its composite with the map $\mathfrak{G}_S \to \mathfrak{G}_T$ defined by $\rho(h)$. Here $\mathfrak{P}'$ is the $\mathbb{Q}^{\mathrm{al}}/\mathbb{Q}$-groupoid obtained from $\mathfrak{P}$ by base change. Assume $T(\mathbb{Q}_p)$ has a hyperspecial subgroup $T(\mathbb{Z}_p)$. Then φ_h arises from a homomorphism $\mathfrak{P} \to \mathfrak{G}_T$.

PROPOSITION 4.40. *Let* $\mathrm{Sh}_p = \mathrm{Sh}_p(T, x)$; *then the sets with operators* $(\mathrm{Sh}_p(\mathbb{F}), \times, \Phi)$ *and* $(S(\varphi_h), \times(\varphi_h), \Phi(\varphi_h))$ *are isomorphic.*

PROOF. See Milne [**50,** 4.2]. □

The homomorphism φ_h satisfies the conditions (4.39), but unless T satisfies the Hasse principle for H^1 there will be other such homomorphisms. This suggests adding another condition.

$(4.39)_0$ The composite of φ with the projection $\mathfrak{G}_G \to \mathfrak{G}_{G^{\mathrm{ab}}}$ is equal to φ_{h_X}.

Statement of the conjecture. In the above, we have found conditions that are surely necessary for φ to be admissible. We now provide one that is surely sufficient.

Let $(T, x) \subset (G, X)$ be a special pair. As above, we obtain a homomorphism $\varphi_x\colon \mathfrak{P} \to \mathfrak{G}_T \subset \mathfrak{G}_G$. Define $\varphi\colon \mathfrak{P} \to \mathfrak{G}_G$ to be *special* if it becomes isomorphic to φ_x for some special pair (T, x) when extended to $\mathfrak{P}'$. We should have the following implications:

$$\varphi \text{ special } \Rightarrow \varphi \text{ admissible } \Rightarrow \varphi \text{ satisfies (4.39).}$$

THEOREM 4.41. *If* G^{der} *is simply connected, then* $\varphi\colon \mathfrak{P} \to \mathfrak{G}_G$ *is special if and only if it satisfies the conditions* (4.39).

PROOF. See Langlands and Rapoport [**46,** 5.3]. □

In other words, when G^{der} is simply connected, the obvious necessary condition agrees with the obvious sufficient condition. Thus, when G^{der} is simply connected we can define φ to be admissible if it satisfies (4.39) or (equivalently) if it is special.

Unfortunately, the two notions diverge when G^{der} is not simply connected. For reasons that we shall presently explain, we choose the second condition.

CONJECTURE 4.42. *There is an isomorphism of sets with operators*

$$(\mathrm{Sh}_p(\mathbb{F}), \times, \Phi) \longrightarrow \coprod_{\varphi} (S(\varphi), \times(\varphi), \Phi(\varphi))$$

where the φ *runs over a set of representatives for the isomorphism classes of special homomorphisms* $\varphi\colon \mathfrak{P} \to \mathfrak{G}_G$.

THEOREM 4.43. *If Conjecture* 4.42 *is true for all Shimura varieties defined by groups with simply connected derived groups, then it is true for all Shimura varieties.*

PROOF. See Milne [**50,** 4.19]. □

An example of Langlands and Rapoport shows that Theorem 4.43 is false if the disjoint union in Conjecture 4.42 is taken over isomorphism classes of homomorphisms satisfying (4.39). It is for this reason that we use special homomorphisms in the statement of (4.42).

In the case that G^{der} is simply connected, Conjecture 4.42 is essentially Conjecture 5.e of Langlands and Rapoport [**46**].

References

1. W. Baily and A. Borel, *Compactification of arithmetic quotients of bounded symmetric domains*, Ann. of Math. (2) **84** (1966), 442–528 (= A. Borel, Collected Papers. II, no. 69).
2. P. Berthelot and A. Ogus, *Notes on Crystalline Cohomology*, Princeton Univ. Press, Princeton, NJ, 1978.
3. ——, *F-isocrystals and de Rham cohomology*, Invent. Math. **72** (1983), 159–199.
4. A. Borel, *Introduction aux Groupes Arithmétiques*, Hermann, Paris, 1969.
5. ——, *Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem*, J. Differential Geom. **6** (1972), 543–560 (= Collected Papers. III, no. 95).
6. ——, *Linear Algebraic Groups* (second edition), Springer, Heidelberg, 1991.
7. M. Borovoĭ, *Langlands's conjecture concerning conjugation of Shimura varieties*, Selecta Math. Soviet **3** (1983/4), 3–39.
8. ——, *On the groups of points of a semisimple group over a totally real field*, Problems in Group Theory and Homological Algebra, Yaroslavl, 1987, pp. 142–149. English translation: Selecta Math. Soviet **9** (1990), 331–338.
9. S. Bosch, W. Lütkebohmert, and M. Raynaud, Néron Models, Springer-Verlag, Berlin-Heidelberg and New York, 1990.
10. N. Bourbaki, *Topologie Générale*, Livre III, Hermann, Paris, 1960.
11. ——, *Groupes et Algèbres de Lie*, Chapitres 4–6, Masson, 1981.
12. J. L. Brylinski, "1-*motifs*" *et formes automorphes* (*Théorie arithmétique des domaines de Siegel*), Journées Automorphes, Publ. Math. Univ. Paris VII, vol. 15, Univ. Paris VII, Paris, 1983, pp. 43–106.
13. E. Cattani, A. Kaplan, and W. Schmid, *Degeneration of Hodge structures*, Ann. of Math. (2) **123** (1986), 457–535.
14. P. Deligne, *Théorie de Hodge*. II, Inst. Hautes Études Sci. Publ. Math. **40** (1971), 5–58.
15. ——, *Travaux de Shimura*, Sém. Bourbaki, Exposé 389, Lecture Notes in Math., vol. 244, Springer, Heidelberg, 1971.
16. ——, *Travaux de Griffiths*, Sém. Bourbaki, Exposé 376, Lecture Notes in Math., vol. 180, Springer, Heidelberg, 1971.
17. ——, *La conjecture de Weil pour les surfaces* K3, Invent. Math. **15** (1972), 206–226.
18. ——, *Variétés de Shimura: Interprétation modulaire, et techniques de construction de modèles canoniques*, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 247–290.
19. ——, *Hodge cycles on abelian varieties* (*Notes by J. S. Milne*), Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer, New York, 1982, pp. 9–100.
20. ——, *Le groupe fondamental de la droite projective moins trois points*, Galois Groups over $\mathbb{Q}$, Springer, Heidelberg, 1989, pp. 79–297.
21. ——, *Catégories Tannakiennes*, The Grothendieck Festschrift, vol. II, Progr. Math., vol. 87, Birkhäuser, Boston, MA, 1990, pp. 111–195.
22. P. Deligne and J. Milne, *Tannakian categories*, Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer-Verlag, New York, 1982, pp. 101–228.
23. P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Modular Functions of One Variable II, Lecture Notes in Math., vol. 349, Springer-Verlag, Heidelberg, 1973, pp. 143–316.
24. F. El Zein, *Introduction à la Théorie de Hodge Mixte*, Hermann, Paris, 1991.
25. G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), 349–366; erratum **75** (1984), 381.
26. ——, *Crystalline cohomology and p-adic Galois-representations*, Algebraic Analysis, Geometry, and Number Theory (Igusa, ed.), Johns Hopkins, 1989, pp. 25–80.
27. G. Faltings and C.-L. Chai, *Degeneration of Abelian Varieties*, Springer, Heidelberg, 1990.
28. J.-M. Fontaine, *Modules galoisiens, modules filtrés et anneaux de Barsotti-Tate*, Journées de Géométrie Algébrique de Rennes, Astérisque **65** (1979), 3–80.
29. ——, *Cohomologie de de Rham, cohomologie cristalline et représentations p-adiques*, Lecture Notes in Math., vol. 1016, Springer, New York, 1983, pp. 86–108.

30. ——, *Représentations p-adiques des corps locaux*, The Grothendieck Festschrift, vol. II, Progr. Math., vol. 87, Birkhäuser, Boston, MA, 1990, pp. 249–309.
31. J.-M. Fontaine and G. Laffaille, *Construction de représentations p-adiques*, Ann. Sci. École Norm. Sup. (4) **15** (1982), 547–608.
32. J.-M. Fontaine and J.-M. W. Messing, *p-adic periods and p-adic étale cohomology*, Current Trends in Arithmetical Algebraic Geometry (Arcata, Calif. 1985), Contemp. Math., vol. 67, Amer. Math. Soc., Providence, RI, 1987.
33. H. Gillet and W. Messing, *Cycle classes and Riemann-Roch for crystalline cohomology*, Duke Math. J. **55** (1987), 501–538.
34. P. Griffiths, *Periods of integrals on algebraic manifolds*. I, II, Amer. J. Math. **90** (1968) 568–626, 805–865.
35. ——, *Periods of integrals on algebraic manifolds: summary of main results and discussion of open problems*, Bull. Amer. Math. Soc. (N. S.) **76** (1970) 228–296.
36. P. Griffiths and W. Schmid, *Locally homogeneous complex manifolds*, Acta Math. **123** (1969), 253–302.
37. A. Grothendieck, *Eléments de Géométrie Algébrique*, Chapitres IV, Inst. Hautes Études Sci. Publ. Math. **20, 24, 28, 32** (1964–67).
38. ——, *Crystals and the De Rham cohomology of schemes (notes by I. Coates and O. Jussila)*, Dix Exposés sur la Cohomologie des Schémas, North-Holland, Amsterdam, 1968, pp. 306–358.
39. G. Harder, *Über die Galoiskohomologie halbeinfacher Matrizengruppen*. II, Math. Z. **92** (1966), 396–415.
40. R. Kottwitz, *Shimura varieties and twisted orbital integrals*, Math. Ann. **269** (1984), 287–300.
41. ——, *Shimura varieties and λ-adic representations*, Automorphic Forms, Shimura Varieties, and *L*-functions, Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, pp. 161–209.
42. ——, *Points on some Shimura varieties over finite fields*, J. Amer. Math. Soc. **5** (1992), 373–444.
43. G. Laffaille, *Groupes p-divisibles et modules filtrés: le cas peu ramifié*, Bull. Soc. Math. France **108** (1980), 187–206.
44. R. Langlands, *Some contemporary problems with origin in the Jugendtraum*, Mathematical developments arising from Hilbert's problems, Amer. Math. Soc., Providence, RI, 1976, pp. 401–418.
45. ——, *Automorphic representations, Shimura varieties, and motives*, Ein Märchen, in Automorphic Forms, Representations, and *L*-functions, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 205–246.
46. R. Langlands and M. Rapoport, *Shimuravarietäten und Gerben*, J. Reine Angew. Math. **378** (1987), 113–220.
47. J. Milne, *The action of* $\mathbb{C}$ *on a Shimura variety and its special points*, Prog. Math., vol. 35, Birkhäuser, Boston, MA, 1983, pp. 239–265.
48. ——, *Automorphic vector bundles on connected Shimura varieties*, Invent. Math. **92** (1988), 91–128.
49. ——, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, Automorphic Forms, Shimura Varieties, and *L*-functions, Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, pp. 283–414.
50. ——, *The points on a Shimura variety modulo a prime of good reduction*, The Zeta Function of Picard Modular Surfaces (Langlands and Ramakrishnan, eds.), Les Publications CRM, Montréal, 1992, pp. 153–255.
51. ——, *Motives over finite fields*, 1993, these Proceedings, vol. 1, pp. 401–459.
52. J. Milne and K-y. Shih, *The action of complex conjugation on a Shimura variety*, Ann. of Math. (2) **113** (1981), 569–599.
53. ——, *Conjugates of Shimura varieties*, Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Math., vol. 900, Springer, Heidelberg, 1982, pp. 280–356.
54. T. Miyake, *Modular Forms*, Springer, Heidelberg, 1989.
55. D. Mumford, *Geometric Invariant Theory*, Springer, Heidelberg, 1965.

56. Y. Namikawa, *Toroidal Compactification of Siegel Spaces*, Lecture Notes in Math., vol. 812, Springer-Verlag, New York, 1980.
57. M. Rapoport, *Complément à l'article de P. Deligne "La conjecture de Weil pour les surfaces* K3", Invent. Math. **15** (1972), 227–236.
58. ——, *Compactifications de l'espace de modules de Hilbert-Blumenthal*, Compositio Math. **36** (1978), 255–335.
59. N. Saavedra Rivano, *Catégories Tannakiennes*, Lecture Notes in Math., vol. 265, Springer, Heidelberg, 1972.
60. J.-J. Sansuc, *Groupe de Brauer et arithmétique des groupes algébriques linéaires sur un corps de nombres*, J. Reine Angew. Math. **327** (1981), 12–80.
61. I. Satake, *Holomorphic imbeddings of symmetric domains into Siegel space*, Amer. J. Math. **87** (1965), 425–461.
62. ——, *Symplectic representations of algebraic groups satisfying certain analyticity conditions*, Acta Math. **117** (1967), 215–279.
63. ——, *Algebraic Structures of Symmetric Domains*, Princeton Univ. Press, Princeton, NJ, 1980.
64. W. Schmid, *Variation of Hodge structure: The singularities of the period mapping*, Invent. Math. **22** (1973), 211–319.
65. S. Shatz, *Profinite Groups, Arithmetic, and Geometry*, Ann. of Math. Stud., vol. 67, Princeton Univ. Press, Princeton, NJ, 1972.
66. G. Shimura, *On canonical models of arithmetic quotients of bounded symmetric domains*, Ann. of Math. (2) **89** (1970), 144–222.
67. G. Shimura and Y. Taniyama, *Complex Multiplication of Abelian Varieties and its Applications to Number Theory*, Publ. Math. Soc. Japan, vol. 6, Tokyo, 1961.
68. A. Silverberg, *Canonical models and adelic representations*, Amer. J. Math. **114** (1992), 1221–1241.
69. ——, *Galois representations attached to points on Shimura varieties*, preprint 1992.
70. J. Tate, *Conjectures on algebraic cycles in ℓ-adic cohomology*, these Proceedings, vol. 1, pp. 71–83.
71. J. Tits, *Reductive groups over local fields*, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 29–69.
72. A. Weil, *The field of definition of a variety*, Amer. J. Math. **78** (1956), 509–526 (= *Collected Papers*. II, 1956).
73. J.-P. Wintenberger, *Un scindage de la filtration de Hodge pour certaines variétés algébriques sur les corps locaux*, Ann. of Math. (2) **119** (1984), 511–548.

UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN
E-mail address: `jmilne@um.cc.umich.edu`

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Zeta Functions of Shimura Varieties

DON BLASIUS AND JONATHAN D. ROGAWSKI

Introduction

The first link between zeta functions of Shimura varieties and automorphic L-functions was discovered by M. Eichler. In 1954 Eichler proved that the zeta function of a suitable model over $\mathbf{Q}$ of the algebraic curve $\Gamma_0(N)\backslash\mathscr{H}$ is equal to a product of L-series of the form $L(s, f)$ where f is a modular form of weight two with respect to $\Gamma_0(N)$ **[E]**. In the 1960s, Shimura developed his fundamental theory of canonical models and, in particular, applied it to compute the zeta functions of all quaternionic curves **[S1, S2]**. The work of both Eichler and Shimura was based on the congruence relation. Ihara treated higher-weight modular forms using a method based on the Eichler-Selberg trace formula. More precisely, in **[I]** the Eichler-Selberg trace formula is compared with an expression for the number of points modulo p on fiber systems of elliptic curves over modular curves, thus relating the zeta functions of the fiber systems to the L-series of higher-weight modular forms. In the early 70s Langlands initiated a program to compute the Hasse-Weil zeta functions attached to all Shimura varieties. The purpose of this article is to review the general conjecture which has emerged.

The congruence relation of Eichler and Shimura has a generalization to symplectic groups (cf. §6) which establishes that there is a close relation between the eigenvalues of Frobenius and the eigenvalues of the Hecke operators. However, to make this connection precise for higher-dimensional Shimura varieties, the congruence relation is not enough, and it is necessary to employ basic results and conjectures from the theory of harmonic analysis on reductive groups. In fact, the search for this generalization led to the development of a substantial speculative framework which now guides much

1991 *Mathematics Subject Classification*. Primary 11F32, 11F46, 11F55, 14G35, 11G40, 11S37, 11S40.

The first author was partially supported by NSF grant DMS-90-01878. The second author was partially supported by NSF grant DMS-91-06194.

This paper is in final form and no version of it will be submitted for publication elsewhere.

current research in the theory of automorphic forms. Some of the catchwords of the subject are:

(1) L-packets and A-packets,
(2) multiplicity formula,
(3) tempered vs. nontempered spectrum,
(4) endoscopy,
(5) stable trace formula.

We discuss the first three topics in §3. However, we do not discuss endoscopy or the stable trace formula, since these are not strictly necessary for the formulation of the conjecture in §5. Of course, these two notions are crucial to the general program of proof, and their introduction by Langlands was motivated by the problem of relating Hasse-Weil zeta functions to automorphic forms **[L1, L3, L7]**. We refer to the basic articles of Kottwitz **[K3]**, Kottwitz-Shelstad **[KS]**, Arthur **[A1, A2]**, and the references therein for a detailed discussion of the theory and the problems it raises.

Nevertheless, the above topics cover only the "analytic part" of the program. The expressions for the zeta functions are obtained by comparing an expression obtained from the stable trace formula with another expression that comes from a description of points modulo p on the Shimura variety. The papers **[K5]**, **[LR]**, and **[Mi3]** deal with this latter problem in detail, and the survey articles of Milne **[Mi2]** and Rapoport **[Ra]** provide an overview of the extensive work on this subject. In §6, we state the general form of the congruence relation, which is still conjectural for groups other than $\mathrm{Sp}(n)$ but may well be accessible in other special cases.

The Shimura varieties associated to $\mathrm{GL}(2)$ and unit groups of quaternion algebras over $\mathbf{Q}$ and, more generally, over totally real fields were the first to be studied thoroughly. After $\mathrm{GL}(2)$, it is natural to consider the Siegel modular three-folds associated to $\mathrm{GSp}(4)$ and the Picard modular surfaces associated to unitary groups in three variables. However, to understand even these special cases turns out to require all of the components of the theory alluded to above in an essential way. The case of Picard modular surfaces is studied in detail in **[M]**. These results are reviewed in §7. The theory for Siegel modular three-folds remains to be worked out. However, we discuss some implications of the general conjectures for this case and also for unitary groups. To illustrate the conjectural structure as predicted by the theory of A-parameters (cf. §§6 and 7), we also prove some new results which can be deduced from the congruence relation. In particular, we prove the existence of global A-parameters for discrete representations of a certain type on groups $U(n, 1)$ assuming the congruence relation (cf. Proposition 7.2). We also examine the unique stable nontempered A-parameter on $\mathrm{GSp}(4)$ and prove that the associated ℓ-adic representations satisfy a nontrivial constraint which cannot be detected at the level of L-parameters (cf. Proposition 7.1).

It is a pleasure to thank Laurent Clozel, Michael Harris, Jean-Pierre Labesse, Joachim Schwermer, as well as the referee for their careful read-

ing of a preliminary manuscript and many helpful comments.

Some notation. The letter F will denote a local field or a number field and G will denote a connected reductive algebraic group over F. Let $\overline{\mathbf{Q}}$ be the algebraic closure of $\mathbf{Q}$ in $\mathbf{C}$. Number fields will be considered as subfields of $\overline{\mathbf{Q}}$. If F is a number field, let $\Gamma_F = \mathrm{Gal}(\overline{\mathbf{Q}}/F)$ and let $\mathbf{A}_F$ be the adèle ring of F. Locally and globally, W_F will denote the the Weil group of F. If v is a finite place of F, Φ_v will denote a Frobenius element for v in Γ_F.

1. L-groups and functoriality

This chapter contains a review of the material contained in the articles **[Bo, L6]**.

1.1. Review of the L-group and Satake isomorphism. The L-group of G is a semidirect product

$$^LG = \widehat{G} \rtimes W_F$$

where $\widehat{G}$ is the dual group. The Weil group acts on $\widehat{G}$ through its projection onto Γ_F. The definition of $\widehat{G}$ and the action of Γ_F on it is recalled below.

REMARK. The L-group is constructed with the Weil group rather than the Galois group in order to allow for more maps between L-groups. Such maps are needed to formulate the functoriality principle in sufficient generality (cf. Example 5.4).

1.2. Duality for tori. Let T be an F-torus. Associated to T are the character and co-character groups

$$X_*(T) = \mathrm{Hom}(GL_1, T) \quad \text{and} \quad \mathrm{Hom}(T, GL_1) = X^*(T).$$

These are free $\mathbf{Z}$-modules, dual relative to the pairing

$$\langle\, , \rangle : X_*(T) \times X^*(T) \to \mathbf{Z} = \mathrm{Hom}(GL_1, GL_1).$$

The dual group of T is the complex torus

$$\widehat{T} = \mathrm{Hom}(X_*(T), \mathbf{C}^*).$$

There is an obvious identification $X^*(\widehat{T}) = X_*(T)$ and by duality, $X_*(\widehat{T}) = X^*(T)$. In other words, passage to the dual group interchanges characters and co-characters.

The Galois group Γ_F acts on $\widehat{T}$ through its action on $X_*(T)$. The L-group ^LT determines T up to isomorphism as an F-torus, since the F-isomorphism class of T is determined by the module $X^*(T)$ with its Γ_F-action. In particular, the action of Γ_F is trivial if and only if T is a split torus.

1.3. Root data. Let (X^*, R^*, X_*, R_*) be a quadruple consisting of free abelian groups X^*, X_* of finite rank in duality relative to a pairing $\langle\, , \rangle$ and subsets

$$R^* \subset X^*, \qquad R_* \subset X_*$$

with a given bijection $\alpha \to \alpha^\vee$ such that $\langle \alpha, \alpha^\vee \rangle = 2$. We call (X^*, R^*, X_*, R_*) a **root datum** if the reflections

$$s_\alpha(x) = x - \langle x, \alpha^\vee \rangle \alpha, \qquad s_{\alpha^\vee}(y) = y - \langle \alpha, y \rangle \alpha^\vee$$

preserve R^* and R_*, respectively. In this case, R^* and R_* are called the sets of roots and co-roots, respectively. The root datum is called **reduced** if for all $\alpha \in R^*$, no multiple $n\alpha$ belongs to R^* with $n > 1$.

The set R defines a root system in $X^* \otimes \mathbf{R}$, and one has the notion of positive roots and simple roots. Let $\Delta^* = \{\alpha\}$ be a set of simple roots, and let $\Delta^* = \{\alpha^\vee\}$. The quadruple $\Psi = (X^*, \Delta^*, X_*, \Delta_*)$ is called a **based root datum**. It determines the original root datum. We associate to Ψ the **dual** based root datum

$$\Psi^\vee = (X_*, \Delta_*, \Delta^*, \Delta^*).$$

1.4. Root datum of a reductive group. Let (B, T) be a **Borel pair**, i.e., a pair consisting of a maximal torus T of $G_{\overline{F}}$ and a Borel subgroup B containing T. If there exists a Borel pair defined over F, then G is said to be **quasi-split**. If T itself is split over F, then G is said to be split. The pair (B, T) gives rise to a reduced root datum

$$\psi(G, B) = (X^*(T), \Delta^*, X_*(T), \Delta_*)$$

where $\Delta^* \subset X^*(T)$ is the subset of simple roots of T that are positive with respect to B, and Δ_* is the set of co-roots associated to the roots in Δ^*. Recall that a root α is positive if it occurs in the action of T on $\mathrm{Lie}(N)$, where N is the unipotent radical of B.

All Borel pairs are conjugate under the adjoint group G_{ad}. If $\mathrm{ad}(g)$ takes (B', T') to (B, T), it induces an isomorphism on the associated root data which is independent of the choice of g. This allows us to identify canonically all root data. We denote the datum so obtained by $\Psi(G)$. With this convention, the full automorphism group $\mathrm{Aut}(G)$ acts on $\Psi(G)$, with G_{ad} acting trivially.

Duality for reductive groups is based on the following classification theorem.

THEOREM 1.4.1. *Let k be an algebraically closed field. The map $G \to \Psi(G)$ induces a bijection between the set of k-isomorphism classes of connected reductive groups G and the set of isomorphism classes of reduced based root data. An isomorphism between reductive groups is uniquely determined up to inner automorphism by the induced map on root data.*

1.5. Dual group. A **dual group** for G is a complex reductive Lie group $\widehat{G}$ together with an isomorphism

$$i : \Psi(\widehat{G}) \xrightarrow{\sim} \Psi(G)^{\vee}.$$

A dual group exists and is unique up to inner automorphism by Theorem 1.4.1. It depends only on the isomorphism class of G over $\overline{F}$.

Let T be a maximal torus in G. There is a unique $\widehat{G}$-conjugacy class of embeddings of $\widehat{T}$ in $\widehat{G}$ which is compatible with i. More precisely, if (T, B) and $(\widehat{S}, \widehat{B})$ are Borel pairs in G and $\widehat{G}$, then there is a unique isomorphism $\widehat{T} \to \widehat{S}$ inducing i on the corresponding root data.

Duality obeys the rules:

(1) The dual of a simple group is a simple group of the same type in the Killing-Cartan classification, unless G is of type B_n or C_n.
(2) Duality interchanges the groups of type B_n and C_n.
(3) Duality interchanges simply connected simple groups and simple groups of adjoint type.
(4) The classical simply connected simple groups and their duals:

$$\begin{aligned} A_n &: \ \mathrm{SL}(n) \leftrightarrow \mathrm{PSL}(n), \\ B_n &: \ \mathrm{Spin}(2n+1) \leftrightarrow \mathrm{PSp}(2n), \\ C_n &: \ \mathrm{Sp}(2n) \leftrightarrow \mathrm{SO}(2n+1), \\ D_n &: \ \mathrm{Spin}(2n) \leftrightarrow \mathrm{PO}(2n). \end{aligned}$$

1.6. Galois action. A **splitting** for $\widehat{G}$ is a triple $\Sigma = (\widehat{B}, \widehat{T}, \{X_\alpha\})$, where, for each simple root α of $\widehat{T}$, X_α is an α-root vector in Lie$(\widehat{G})$. The set of all splittings is a principal homogeneous space for G_{ad}. By Theorem 1.4.1, the following sequence is split exact:

$$1 \to G_{\mathrm{ad}} \to \mathrm{Aut}(G) \to \mathrm{Aut}(\Psi(\widehat{G})) \to 1.$$

A choice of Σ determines a splitting since each element of $\mathrm{Aut}(\Psi(\widehat{G}))$ lifts to a unique automorphism of G fixing Σ.

The action of Γ_F on $\Psi(G)^{\vee} = \Psi(\widehat{G})$ lifts to an action on $\widehat{G}$ which fixes a given splitting Σ. Let the Weil group act on $\widehat{G}$ through its projection to Γ_F. The L-group is the semidirect product relative to this action

$${}^L G = \widehat{G} \rtimes W_F.$$

The L-group depends on a choice of splitting, but different choices give rise to isomorphic L-groups.

Let F be a number field, and let v be a place of F. There is a distinguished W_F-conjugacy class of embeddings $\eta_v : W_{F_v} \to W_F$ [T] and hence a class of embeddings ${}^L G_v \to {}^L G$ inducing the identity on $\widehat{G}$.

1.7. Inner forms. Let $f : G^* \to G$ be an $\overline{F}$-isomorphism of F-algebraic groups. Then f induces an isomorphism of root data $\overline{f} : \Psi(G^*) \xrightarrow{\sim} \Psi(G)$.

The map f is called an **inner twisting** if for all $\sigma \in \Gamma_F$, there exists $g_\sigma \in G(\overline{F})$ such that $f(\sigma(g)) = \mathrm{ad}(g_\sigma)(\sigma(f(g)))$. This is the case if and only if $\overline{f}$ commutes with the Galois actions. If an inner twisting exists, then G^* is said to be an **inner form** of G.

The L-group ${}^LG$ depends only on the class of inner forms to which G belongs. It follows from the conjugacy of Borel pairs that in each such class there is a unique quasi-split form. The L-group determines the F-isomorphism class of this quasi-split form. In particular, the Galois action on $\widehat{G}$ is trivial if and only if G is an inner form of a split group.

1.8. Examples of L-groups. (a) Let $G = \mathrm{GL}(n)$. In this case, $\widehat{\mathrm{GL}(n)} = \mathrm{GL}_n(\mathbf{C})$. To verify this, identify X^* and X_* with $\mathbf{Z}^n$ under the standard pairing $\langle e_i, e_j \rangle = \delta_{ij}$ and let

$$\Delta^* = \Delta_* = \{e_i - e_{i+1} : 1 \leq i \leq n - 1\}.$$

Let $i : \Psi(G) \to \Psi(G)^\vee$ be the identity bijection. The Galois action on $\widehat{G}$ is trivial since G is a split group; hence, ${}^LG = \mathrm{GL}_n(\mathbf{C}) \times W_F$. More generally, ${}^LG' = \mathrm{GL}_n(\mathbf{C}) \times W_F$ for any inner form G' of $\mathrm{GL}(n)$. These are obtained as the multiplicative groups of central simple algebras over F of dimension n^2.

(b) **Unitary groups**. Let E/F be a quadratic extension, and let σ be the nontrivial element of $\mathrm{Gal}(E/F)$. If $\Phi \in \mathrm{GL}_n(E)$ is a Hermitian matrix, let $G = U(\Phi)$ denote the unitary group in n variables with respect to Φ and E/F, viewed as an algebraic group over F. Then

$$G(F) = \{g \in \mathrm{GL}_n(E) : \Phi\, {}^t\overline{g}^{-1} \Phi^{-1} = g\}$$

where the bar denotes conjugation with respect to E/F. Over E, G is isomorphic to $\mathrm{GL}(n)_{/E}$, and G is an outer F-form of $\mathrm{GL}(n)_{/E}$. It follows that $\widehat{G} = \mathrm{GL}_n(\mathbf{C})$ and the Galois action on $\widehat{G}$ factors through $\mathrm{Gal}(E/F)$. There is a standard splitting for $\widehat{G}$:

$$\begin{aligned} \widehat{T} &= \text{the diagonal subgroup}, \\ \widehat{B} &= \text{the group of upper-triangular matrices}, \\ \{X_\alpha\} &= \text{set of matrices } (a_{ij}) = (\delta_{i,k}\delta_{j,k+1}) \text{ for } k = 1, 2, \dots, n-1. \end{aligned}$$

Let Φ_n be the matrix $(\Phi_n)_{ij} = (-1)^{i-1}\delta_{i,n-j+1}$. Then $g \to \Phi_n\, {}^tg^{-1}\Phi_n^{-1}$ is the unique noninner automorphism that preserves the standard splitting. Hence the nontrivial element $\sigma \in \mathrm{Gal}(E/F)$ acts on $\widehat{G}$ by this automorphism and the L-group is the semidirect product ${}^LU(n) = \mathrm{GL}_n(\mathbf{C}) \rtimes W_F$.

(c) **Unitary similitudes**. Let $G = \mathrm{GU}(\Phi)$ be the group of unitary similitudes of the Hermitian form Φ. Then

$$G(F) = \{g \in \mathrm{GL}_n(E) : \Phi\, {}^t\overline{g}^{-1}\Phi^{-1} = \lambda^{-1}g \text{ where } \lambda \in F^*\}.$$

As an algebraic group, G is the F-form of the group $\mathrm{GL}_n \times \mathrm{GL}_1$ over E defined relative to the action $(g, \lambda) \to (\bar{\lambda}\Phi\, {}^t\bar{g}^{-1}\Phi^{-1}, \bar{\lambda})$. Its dual group is $\mathrm{GL}_n(\mathbf{C}) \times \mathrm{GL}_1(\mathbf{C})$. The action of W_F again factors through the projection to $\mathrm{Gal}(E/F)$, and σ acts by

$$(g, \lambda) \to (\Phi_n\, {}^t g^{-1}\, \Phi_n, \lambda \det(g)).$$

(d) **Symplectic similitudes**. Let $G = \mathrm{GSp}(2n)$ be the group of symplectic similitudes. Its dual group is $\widehat{G} = \mathrm{GSpin}(2n+1)$, defined by the exact sequence below. Consider the exact sequence

$$1 \to \{1, -1\} \to \mathrm{GSp}(2n) \to \mathrm{GL}(1) \times \mathrm{PSp}(2n) \to 1.$$

The third arrow takes g to $(\nu(g), g')$ where $\nu(g)$ is the multiplier of G and g' is the image of $\pm\nu(g)^{-\frac{1}{2}}g$ in $\mathrm{PSp}(2n)$. Duality yields

$$1 \to \{1, \varepsilon\} \to \mathrm{GL}(1) \times \mathrm{Spin}(2n+1) \to \mathrm{GSpin}(2n+1) \to 1,$$

where ε is the central element of order two in $\mathrm{GL}(1) \times \mathrm{Spin}(2n+1)$ both of whose projections are nontrivial. The Galois action on $\widehat{G}$ is trivial.

(e) If $G = \mathrm{GSp}(4)$ then $\widehat{G} = \mathrm{GSp}(4)$. This follows from the above and the fact that $\mathrm{Spin}(5)$ is isomorphic to $\mathrm{Sp}(4)$.

1.9. Admissible representations. We briefly recall the notion of admissible representation of a reductive group over a local field. It is necessary to treat the Archimedean and p-adic cases separately.

Let F be a p-adic field. A homomorphism

$$\pi : G(F) \to \mathrm{Aut}(V),$$

where V is a complex vector space, is called an **admissible representation** if it satisfies the conditions

(a) Every vector $v \in V$ is fixed by an open subgroup of $G(F)$;
(b) For every open subgroup $K \subset G(F)$, the space of K-fixed vectors in V is finite dimensional.

Admissible representations (π_1, V_1) and (π_2, V_2) are *equivalent* if there exists a bijective linear map $T : V_1 \to V_2$ that intertwines π_1 and π_2. Suppose that $G(F)$ acts on a Hilbert space H by a unitary representation, and let H_0 be the subspace of vectors in H that are fixed by some open compact subgroup of $G(F)$. Then H_0 is stable under the action of $G(F)$. According to a theorem of J. Bernstein, if H is irreducible, then H_0 is an admissible representation of G. Two irreducible unitary representations are unitarily equivalent if and only if the representations on their associated spaces of smooth vectors are equivalent.

If F is Archimedean, it is more appropriate to consider admissible $(\mathfrak{G}, K)$-modules where K is a maximal compact subgroup of $G(F)$ and $\mathfrak{G}$ is the

Lie algebra of $G(F)$. This is a pair (π, V) where V is a complex vector space on which $\mathfrak{G}$ and K act such that:

(a) V decomposes as a vector space direct sum of irreducible representations of K, each occurring with finite multiplicity.
(b) The two actions of $\mathrm{Lie}(K)$ obtained from the differential of the K action and the restriction of the $\mathfrak{G}$ action coincide.
(c) $\pi(\mathrm{Ad}(k)X) = \pi(\mathrm{Ad}(k))\pi(X)\pi(\mathrm{Ad}(k)X)$.

Two $(\mathfrak{G}, K)$-modules are said to be equivalent if there is a bijective linear map between them that intertwines the actions of $\mathfrak{G}$ and K.

We denote the set of equivalence classes of irreducible admissible representations (or $(\mathfrak{G}, K)$-modules, in the Archimedean case) by $\Pi_F(G)$ or $\Pi(G)$.

1.10. L-Parameters and L-maps. Suppose that F is a local field, and set

$$\mathscr{L}_F = \begin{cases} W_F \times \mathrm{SU}_2(\mathbf{R}) & \text{if } F \text{ is } p\text{-adic}, \\ W_F & \text{if } F \text{ is Archimedean.} \end{cases}$$

Here $\mathrm{SU}_2(\mathbf{R})$ is the compact special unitary group. An L-parameter is a homomorphism

$$\phi : \mathscr{L}_F \to {}^LG$$

such that

(i) composition with the projection to W_F induces the identity on W_F,
(ii) $\phi(w)$ is semisimple for all $w \in W_F$.

Two parameters ϕ_1 and ϕ_2 are said to be equivalent if they are conjugate under the action of $\widehat{G}$. Let $\Phi_F(G)$ or $\Phi(G)$ denote the set of equivalence classes of L-parameters.

Let $r : {}^LG \to \mathrm{GL}(V)$ be a linear representation. To each equivalence class of parameters ϕ is associated an L-function $L(s, \phi, r)$ as follows. If F is p-adic, extend $r \circ \phi$ to a homomorphism $\mathscr{L}_F \times \mathrm{SU}_2(\mathbf{C}) \to \mathrm{GL}(V)$. Choose a nonzero nilpotent element $N \in \mathrm{Lie}(\mathrm{SU}_2(\mathbf{C}))$ and a nontrivial homomorphism $w : \mathbf{G}_m \to \mathrm{SU}_2(\mathbf{C})$ such that $\mathrm{ad}(w(t))(N) = t^2N$. Let $V_N = \ker(\phi(N))$, and let V_N^I be the space of invariants in V_N under the action of the inertia subgroup of W_F. Let Φ be a Frobenius element in W_F. The restriction of $\phi(\Phi)$ to V_N^I is well defined. We set

$$L(s, \phi, r) = \det(1 - q^{-s}r(\phi(\Phi, w(q^{-1/2})))|V_N^I)^{-1},$$

where q is the cardinality of the residue field of F. This is independent of the choices. See [T] for a formulation in terms of the Weil-Deligne group and for the definition of the local factors in the Archimedean case.

LOCAL LANGLANDS CONJECTURE. *The set $\Pi_F(G)$ is partitioned into finite subsets called L-packets in such a way that there is a "natural" bijection between $\Phi_F(G)$ and the set of L-packets.*

See **[Ku]** for a discussion of the desired naturality properties of the correspondence. If $\phi \in \Phi_F(G)$, let $\Pi(\phi)$ denote the L-packet conjecturally associated to ϕ.

REMARKS. (1) The action of $G_{\mathrm{ad}}(F)$ on $\Pi_F(G)$ may be nontrivial. Representations in the same $G_{\mathrm{ad}}(F)$-orbit will lie in the same L-packet. However, this does not suffice to define L-packets in general. In particular, an L-packet for a group of adjoint type may contain more than one element.

(2) L-packets for unramified representations can be defined in general. See below.

(3) Some cases in which L-packets have been defined are as follows. For $G = \mathrm{GL}(n)$, an L-packet is a singleton. For $G = \mathrm{SL}_n(F)$, an L-packet is an orbit of representations under the action of $\mathrm{PGL}_n(F)$ **[LL]**. They have been defined for unitary groups in three variables **[R]**. For real groups, L-packets were defined in **[L5]**. The real theory is developed in **[Sh2]**.

A homomorphism between L-groups

$$\psi : {}^LH \to {}^LG$$

is called an **L-map** if it commutes with the projections onto W_F. Composition $\varphi \to \psi \circ \varphi$ gives rise to a map

$$\psi : \Phi_F(H) \to \Phi_F(G).$$

If the local Langlands conjecture is true, this corresponds to a **transfer of L-packets** on H to L-packets on G.

1.11. Unramified parameters. In this section and the next, we review the local Langlands correspondence in the unramified case **[L6]**. This is essential to define functorial transfers of automorphic representations in the global case.

Assume that F is p-adic, and let q be the cardinality of the residue field k_F of F. Then W_F is the subgroup of elements in Γ_F that induce an integral power of the Frobenius automorphism Φ. Let $\mathrm{val} : W_F \to \mathbf{Z}$ be the valuation map: $\mathrm{val}(w) = n$ if w induces Φ^n on k. The group G is called **unramified** if

(1) G is quasi-split,
(2) G splits over an unramified extension of F.

Suppose that G is unramified. The action of W_F on ${}^LG$ factors through $\mathrm{val} : W_F \to \mathbf{Z}$ and the semidirect product $\widehat{G} \rtimes \mathbf{Z}$ is defined. An L-parameter ϕ is called **unramified** if it is trivial on the $\mathrm{SU}_2(\mathbf{R})$-factor and the composite

$$W_F \to {}^LG \to \widehat{G} \rtimes \mathbf{Z}$$

factors through val. Let $\Phi_{\mathrm{un}}(G)$ be the set of equivalence classes of unramified L-parameters.

An unramified parameter ϕ is determined by the semisimple element $\phi(\Phi) = g \rtimes \Phi$. Unramified parameters ϕ_1 and ϕ_2 are equivalent if the elements $\phi_j(\Phi)$ are $\widehat{G}$-conjugate. It follows that $\Phi_{\text{un}}(G)$ can be identified with the set of $\widehat{G}$-conjugacy classes in ${}^LG$ of semisimple elements of the form $g \rtimes \Phi$.

1.12. Unramified representations. We continue to assume that F is p-adic and that G is unramified. Fix a hyperspecial maximal compact subgroup K of $G(F)$ [Ti]. An irreducible admissible representation (π, V) of $G(F)$ is said to be **unramified** if V^K is nonzero. Let $\Pi_{\text{un}}(G)$ denote the set of equivalence classes of unramified representations of $G(F)$.

PROPOSITION 1.12.1. *There is a natural bijection between* $\Phi_{\text{un}}(G)$ *and* $\Pi_{\text{un}}(G)$.

PROOF. Assume first that G is an F-torus T. For $t \in T(F)$, the map $\chi \to \text{val}_F(\chi(t))$ defines an element $\text{Hom}(X^*(T), \mathbf{Z})$ which factors through the Galois co-invariants and hence defines a homomorphism

$$v_t : X^*(T)_{\Gamma_F} \to \mathbf{Z}.$$

The map v_t is trivial if and only if t belongs to the maximal compact subgroup T_c of $T(F)$. By duality, $X_*(T)^{\Gamma_F} = \text{Hom}(X^*(T)_{\Gamma_F}, \mathbf{Z})$, and we obtain a map

$$v : T(F)/T_c \to X_*(T)^{\Gamma_F}.$$

The surjectivity of v is obtained by passing to an unramified splitting field K of T and then descending using that $H^1(\text{Gal}(K/F), \mathscr{O}_K^*)$ is trivial.

Let S be the maximal F-split torus in T. Then $X_*(S) = X_*(T)^{\Gamma_F}$ and

$$\widehat{S} = \text{Hom}(X_*(T)^{\Gamma_F}, \mathbf{C}^*) = \text{Hom}(T(F)/T_c, \mathbf{C}^*).$$

The group on the right is $\Pi_{\text{un}}(T)$. This yields an identification

$$\widehat{S} = \Pi_{\text{un}}(T). \tag{1.12.1}$$

To finish, observe that there is an exact sequence

$$1 \to \widehat{T}^{1-\Phi} \to \widehat{T} \to \widehat{S} \to 1.$$

corresponding to the inclusion $X_*(S) \to X_*(T)$. This yields a bijection

$$\widehat{S} \leftrightarrow \widehat{T} \rtimes \Phi \text{ modulo the adjoint action of } \widehat{T}$$

since the set on the right is in bijection with the elements of $\widehat{T}$ modulo $\widehat{T}^{1-\Phi}$.

Now assume that G is an unramified reductive group, and let S be a maximal F-split torus in G. Let T be the maximal F-torus containing S. Choose an embedding of $\widehat{T}$ into $\widehat{G}$ in the canonical conjugacy class (cf. 1.5) with image contained in a Borel subgroup $\widehat{B}$ such that the pair $(\widehat{T}, \widehat{B})$ is fixed by the Galois action. This extends to an embedding of ${}^LT$ in ${}^LG$ and yields a map

$$\widehat{S} = \Phi_{\text{un}}(T) \to \Phi_{\text{un}}(G).$$

Let $\Omega_F(T)$ be the Weyl group T in G. It preserves S and acts on $\widehat{S}$ by duality. The above map factors to an isomorphism

$$\Phi_{\mathrm{un}}(T)/\Omega_F(T) = \Phi_{\mathrm{un}}(G) \tag{1.12.2}$$

[**Bo**, p. 37]. On the representation-theoretic side, a bijection

$$\Pi_{\mathrm{un}}(T)/\Omega_F(T) \to \Pi_{\mathrm{un}}(G),$$
$$\chi \to \pi(\chi)$$

is constructed by means of the unramified principal series as follows. Let N be the unipotent radical of B. Then $B(F) = T(F)N(F)$, and the Iwasawa decomposition $G(F) = B(F)K$ holds. Let $\chi \in \Pi_{\mathrm{un}}(T)$ and extend χ to a character of $B(F)$ trivial on $N(F)$. Let $I(\chi)$ be the representation of $G(F)$ by right translation on the space of locally constant functions $f : G(F) \to \mathbf{C}^*$ such that

$$f(bg) = \chi(b)\delta^{1/2}(b)f(g) \quad \text{for all } b \in B(F).$$

Here, $\delta(b) = |\det(\mathrm{ad}(b)|\,\mathrm{Lie}(N))|$ is the modulus function for $B(F)$ acting on $N(F)$ via conjugation. By the Iwasawa decomposition, the action of $G(F)$ on $I(\chi)$ is admissible and $I(\chi)$ contains a one-dimensional subspace which is invariant under K. Therefore, $I(\chi)$ contains a unique irreducible subquotient V such that $\dim(V^K) = 1$. We denote the representation of $G(F)$ on V by $\pi(\chi)$. The unramified correspondence now follows from the following result in the theory of spherical functions on p-adic groups.

THEOREM 1.12.2. *Every unramified representation of $G(F)$ is of the form $\pi(\chi)$ for some unramified character χ. Furthermore, $\pi(\chi)$ is equivalent to $\pi(\chi')$ precisely when χ is conjugate to χ' under $\Omega_F(T)$.*

REMARK. The notion of unramified representation depends on the choice of hyperspecial maximal compact subgroup K. However, the hyperspecial maximal compact subgroups form a single orbit under the action of $G_{\mathrm{ad}}(F)$ [**Ti**, p. 47]. An L-packet is called **unramified** if it contains an unramified representation. This notion is independent of the choice of K.

1.13. The Satake isomorphism. Let K be a hyperspecial maximal compact subgroup of G. The **Hecke algebra** of G with respect to K is the algebra $\mathscr{H}$ of compactly supported, **Z**-valued functions f on $G(F)$ such that $f(k_1 g k_2) = f(g)$ for all $k_1, k_2 \in K$. Multiplication in $\mathscr{H}$ is convolution with respect to the Haar measure on $G(F)$ normalized so that $\mathrm{meas}(K) = 1$. Let $\mathscr{H}_{\mathbf{C}} = \mathscr{H} \otimes \mathbf{C}$. Define the Satake transform $f \to \hat{f}$, where $\hat{f}$ is the function $\pi \to \mathrm{Trace}(\pi(\hat{f}))$ on $\Pi_{\mathrm{un}}(G)$. Equivalently, we may view $\hat{f}$ as a function on $\widehat{S}/\Omega_F$. The space $\widehat{S}/\Omega_F$ is an affine variety whose coordinate ring is the ring $\mathbf{C}[X_*(S)]^{\Omega_F(T)}$ of $\Omega_F(T)$-invariants in the group ring of $X_*(S)$. The following reformulates a result of Satake [**Bo**].

THEOREM (SATAKE). *For all $f \in \mathscr{H}_{\mathbf{C}}$, the Satake transform $\hat{f}$ belongs to $\mathbf{C}[X_*(S)]^{\Omega_F(T)}$, and the Satake transform defines an algebra isomorphism of $\mathscr{H}_{\mathbf{C}}$ with $\mathbf{C}[X_*(S)]^{\Omega_F(T)}$.*

In **[L6]**, it is shown that $\mathbf{C}[X_*(S)]^{\Omega_F}$ coincides with the ring of functions on $\widehat{S}$ generated by the restrictions to the space of conjugacy classes of characters of finite-dimensional algebraic representations

$$r : {}^L G \to \mathrm{GL}(V)$$

(view $\widehat{S}$ as a quotient of $\widehat{T} \rtimes \Phi$ as above). Given r, consider the characteristic polynomial

$$P_r(X) = \det(1 - r(g \rtimes \Phi)X)$$

as a function on $\widehat{G} \rtimes \Phi$. The coefficients of $P_r(X)$ belong to the ring of restrictions of characters of algebraic representations. In view of the above theorem, the coefficients of $P_r(X)$ are Satake transforms of functions in $\mathscr{H}_{\mathbf{C}}$.

EXAMPLE. $G = \mathrm{GL}(2)_{\mathbf{Q}_p}$, $r =$ the two-dimensional representation of ${}^L G$ which is the identity on the $\mathrm{GL}_2(\mathbf{C})$-factor and trivial on the W_F-factor. Then

$$P_r(X) = 1 - p^{-1/2} T_p X + T(p, p) X^2.$$

Here T_p and $T(p, p)$ are $\mathrm{meas}(K)^{-1}$ times the characteristic functions of $K\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)K$ and $K\left(\begin{smallmatrix} p & \\ & p \end{smallmatrix}\right)$, respectively, where $K = \mathrm{GL}_2(\mathbf{Z}_p)$.

1.14. Automorphic L-functions. Let F be a number field. An admissible representation π of $G(\mathbf{A}_F)$ is isomorphic to a restricted tensor product $\bigotimes \pi_v$[F]. For almost all places v, the local component π_v is unramified.

If $r : {}^L G \to \mathrm{GL}(V)$ is a linear representation of the L-group, the local L-factor for π_v is defined by

$$L(s, \pi_v, r) = L(s, \phi_v, r)$$

where ϕ_v is the L-parameter attached to π_v. This definition is valid only if the Langlands corrspondence for π_v is known.

If π_v is unramified, then ϕ_v is defined. The $\widehat{G}$-conjugacy class of the element

$$g(\pi_v) = \phi_v(\Phi_v) \in {}^L G$$

is determined by π_v and is called the **Langlands class** of π_v. We have

$$L(s, \pi_v, r) = \det(1 - Nv^{-s} r(g(\pi_v)))^{-1}.$$

The partial global L-function is defined as an Euler product

$$L_S(s, \pi, r) = \prod_{v \notin S} L(s, \phi_v, r)$$

for any finite set S of places outside of which the local factors are defined. If π is unitary, then the product converges absolutely in a half-plane of the form $\mathrm{Re}(s) > c$ **[Bo]**.

2. Shimura varieties

2.1. Definition. Let G be a connected, reductive group defined over $\mathbf{Q}$, and let $\mathscr{R} = \mathrm{Res}_{\mathbf{C}/\mathbf{R}}(\mathbf{G}_m)$. Suppose that there exists a homomorphism of algebraic groups over $\mathbf{R}$

$$h : \mathscr{R} \to G_{\mathbf{R}}$$

which satisfies the axioms of Deligne **[Mi]**. Then G gives rise to a Shimura variety.

Let K_h be the stabilizer of h in $G(\mathbf{R})$. The orbit $\mathfrak{X}$ of h under the adjoint action of $G(\mathbf{R})$ is isomorphic to $G(\mathbf{R})/K_h$ and carries a natural structure of a Hermitian symmetric domain. Let K be an open compact subgroup of $G(\mathbf{A}_f)$ of the form $K = \prod_{v<\infty} K_v$, where K_v is open compact in G_v and $K_v = G(\mathscr{O}_v)$ for almost all finite places v. Set

$$S_K(\mathbf{C}) = G(\mathbf{Q})\backslash\mathfrak{X} \times G(\mathbf{A}_f)/K.$$

We always assume that K is sufficiently small so that $S_K(\mathbf{C})$ is a smooth complex manifold. As is well known, this manifold is the set of complex points of a quasi-projective variety S_K which is defined over an explicitly given number field, the **reflex field** E. The definition of E is recalled in §5. The Baily-Borel-Satake compactification of S_K is also defined over E. We denote it by $\overline{S}_K$. Set

$$IH^i = IH^i(\overline{S}_K(\mathbf{C}), \mathbf{Q}),$$

the **intersection cohomology** of $\overline{S}_K(\mathbf{C})$. For any field L containing $\mathbf{Q}$, set $IH^i(L) = IH^i \otimes L$.

For any field L, let $\mathscr{H}_L(K)$ be the Hecke algebra of L-valued, bi-K-invariant functions on $G(\mathbf{A}_f)$. It is spanned over L by the characteristic functions of double cosets KgK for $g \in G(\mathbf{A}_f)$. The Hecke algebra $\mathscr{H}_{\mathbf{C}}(K)$ acts on $IH^i(\mathbf{C})$. We now review the description of $IH^i(\mathbf{C})$ as a Hecke module in terms of automorphic forms.

2.2. Automorphic representations. Let ξ be a character of $Z(\mathbf{Q})\backslash Z(\mathbf{A})$. Then there exists an $\mathbf{R}_+^*$-valued character ξ_0 of $G(\mathbf{Q})\backslash G(\mathbf{A})$ whose restriction to $Z(\mathbf{A})$ is $|\xi|$. Let $L(G, \xi)$ be the Hilbert space of measurable functions f on $G(\mathbf{Q})\backslash G(\mathbf{A})$ such that $\xi_0(g)^{-1}|f(g)|$ is square-integrable on $G(\mathbf{Q})Z(\mathbf{A})\backslash G(\mathbf{A})$. Let $L_d(G, \xi)$ be the part of $L(G, \xi)$ that decomposes discretely under the action by right translation, i.e., the closed sum of all irreducible $G(\mathbf{A})$-invariant subspaces of $L(G, \xi)$. It is known that $L_d(G, \xi) = L(G, \xi)$ if and only if G is anisotropic. Furthermore, $L_d(G)$ decomposes as a direct sum of irreducible unitary representations with finite multiplicities:

$$L_d(G) = \bigoplus m(\pi) \cdot \pi.$$

A representation that occurs in this direct sum is called a **discrete automorphic representation**. All such representations are factorizable as a *restricted* tensor product of admissible representations: $\pi = \bigotimes \pi_v$ (see **[F]** for the precise definitions). We shall write

$$\pi = \pi_\infty \otimes \pi_f$$

where π_∞ is a unitary representation of $G(\mathbf{R})$ and $\pi_f = \bigotimes_{v<\infty} \pi_v$ is a representation of $G(\mathbf{A}_f)$.

Let π_f^K be the subspace of K-invariants in the space of π_f. Admissibility implies that π_f^K is a finite-dimensional space. The space π_f^K is an irreducible module over $\mathscr{H}_{\mathbf{C}}(K)$ **[Ca]**.

2.3. Representations with cohomology. Let Coh_∞ be the set of unitary representations π_∞ of $G(\mathbf{R})$ (up to equivalence) such that the relative Lie algebra cohomology $H^*(\mathfrak{G}, K_\infty, \pi_\infty)$ is nontrivial, where $\mathfrak{G}$ is the Lie algebra of G_∞. The set Coh_∞ is finite and has been determined in **[VZ]** (see §4). For any π_f, set

$$\mathrm{Inf}(\pi_f) = \{\pi_\infty \in \mathrm{Coh}_\infty : m(\pi_\infty \otimes \pi_f) \neq 0\}.$$

Let Coh_f be the set of π_f such that $\mathrm{Inf}(\pi_f)$ is nonempty.

The L^2 cohomology $H^i_{(2)}(S_K(\mathbf{C}), \mathbf{C})$ can be decomposed in terms of discrete automorphic representations **[BC]**. More precisely, $\mathscr{H}_{\mathbf{C}}(K)$ acts on the L^2-cohomology and the decomposition with respect to this action is

$$H^i_{(2)}(S_K(\mathbf{C}), \mathbf{C}) = \bigoplus m(\pi) H^i(\mathfrak{G}, K_\infty, \pi_\infty) \otimes \pi_f^K$$

where π ranges over the discrete automorphic representations. For anisotropic G, the L^2-cohomology is ordinary de Rham cohomology, and this is Matsushima's formula. There is an intrinsic Hodge (p, q)-decomposition on the spaces $H^i(\mathfrak{G}, K_\infty, \pi_\infty)$ **[BW]**. In the anisotropic case (i.e., $S_K(\mathbf{C})$ compact) the above equality respects the (p, q)-decomposition. One may expect that this will continue to hold in the noncompact case.

According to the Zucker conjecture, proved by Looijenga and Saper-Stern, the L^2-cohomology of S_K is isomorphic to the intersection cohomology of $\overline{S}_K$ **[Z]**. The isomorphism commutes with the action of $\mathscr{H}_{\mathbf{C}}(K)$. If L is a sufficiently large number field, then the action of $\mathscr{H}_L(K)$ on $IH^i(L)$ decomposes completely. We have $IH^i(\mathbf{C}) = IH^i(L) \otimes \mathbf{C}$, and this yields an isotypic decomposition with respect to the Hecke algebra

$$IH^i(L) = \bigoplus_{\pi_f \in \mathrm{Coh}_f} H^i(\pi_f) \otimes \pi_f^K(L). \tag{2.3.1}$$

Here $\pi_f^K(L)$ is a realization of π_f^K over L and $H^i(\pi_f)$ is a finite-dimensional L-vector space of dimension

$$\dim(H^i(\pi_f)) = \sum_{\pi_\infty \in \mathrm{Inf}(\pi_f)} m(\pi_\infty \otimes \pi_f) \dim(H^i(\mathfrak{G}, K_\infty, \pi_\infty)).$$

2.4. λ-Adic representations. Let λ be a finite place of L, and set $IH^i(L_\lambda) = IH^i(L) \otimes L_\lambda$. The comparison theorem in étale cohomology allows us to identify $IH^i(L_\lambda)$ with the étale cohomology group $IH^i_{\text{ét}}(\overline{S_K} \times \overline{\mathbf{Q}}, \mathbf{Q}_\ell) \otimes L_\lambda$. This equips $IH^i(L_\lambda)$ with a continuous action of the Galois group Γ_E. This action commutes with the action of $\mathscr{H}_{L_\lambda}(K)$. By (2.3.1), we obtain a representation $\rho^i_\lambda(\pi_f)$ of Γ_E on

$$H^i_\lambda(\pi_f) = H^i(\pi_f) \otimes L_\lambda.$$

The collection $\rho^i(\pi_f) = \{\rho^i_\lambda(\pi_f)\}$ forms a system of λ-adic representations which is compatible at almost all finite places of F. Since the field L is given as a subfield of $\mathbf{C}$, the L-function of $\rho^i(\pi_f)$ is defined:

$$L(s, \rho^i(\pi_f)) = \prod_v L_v(s, \rho^i_\lambda(\pi_f)).$$

General problem. Describe the representations $\rho^i_\lambda(\pi_f)$ in terms of π_f. In particular, determine their L-functions as automorphic L-functions.

The planned generalization of Eichler-Shimura theory is a description of $L(s, \rho(\pi_f))$ as an automorphic L-function $L(s, \sigma, r')$ where σ is an automorphic representation of a reductive group H and r' is a representation of ${}^LH$. There is an embedding of L-groups $\varphi: {}^LH \to {}^LG$ such that π is the functorial transfer of σ and r' is a factor of $r \circ \varphi$, where r is the representation defined in §5.1. The definition of the groups H, which are now called **elliptic endoscopic groups**, was found by Langlands [L7]. Since we are not stressing the role of endoscopic groups, we shall describe $L(s, \rho(\pi_f))$ directly as a factor of an L-function of π itself.

3. Parameters and packets

The purpose of this section is to introduce more of the ingredients that go into the statement of the conjectural description of the L-functions of the systems $\{\rho^i_\lambda(\pi_f)\}$.

3.1. Global parameters. To understand the structure of the automorphic spectrum, it is useful and perhaps necessary to assume the existence of the Langlands group $\mathscr{L}_F$. This is a conjectural extension of W_F by a compact group whose basic property is that its isomorphism classes of continuous n-dimensional irreducible representations are in natural bijection with the set of equivalence classes of cuspidal representations of $\mathrm{GL}_n(\mathbf{A}_F)$. In this article we shall assume the existence, not only of the Langlands group, but also of the associated theory of packets outlined below. To avoid overuse of the word "conjecture", we shall take it as understood that all statements involving $\mathcal{L}_F$ are necessarily conjectural, as are the various properties of packets. Note, however, that since W_F is a quotient of $\mathcal{L}_F$, examples can be constructed which give the flavor of the formalism.

The local Langlands groups $\mathcal{L}_{F_v}$ have been defined in §1.10. The formalism requires that there exist a distinguished conjugacy class of embeddings $i_v : \mathcal{L}_{F_v} \to \mathcal{L}_F$. This allows us to pull back an irreducible representation $\varphi : \mathcal{L}_F \to \mathrm{GL}_n(\mathbf{C})$ to a representation φ_v of $\mathcal{L}_{F_v}$ for all v such that the equivalence class of φ_v is well defined. The local L-factors $L(s, \varphi_v)$ have been defined in §1.10. The bijection between irreducible representations of $\mathcal{L}_F$ and cuspidal representations is normalized by the condition

$$\varphi \leftrightarrow \pi \quad \text{if} \quad L(s, \varphi_v) = L(s, \pi_v) \quad \text{for almost all } v.$$

This suffices to determine π uniquely since π is determined by its local components π_v for $v \notin S$ for any finite set S by the strong multiplicity-one theorem for $\mathrm{GL}(n)$.

A map

$$\phi : \mathcal{L}_F \to {}^LG$$

is called an **L-parameter** if it commutes with the projections to W_F and if $\phi(w)$ is semisimple for all $w \in \mathcal{L}_F$. If $\phi(L_F)$ is bounded modulo $Z(\widehat{G})$ in ${}^LG$, then ϕ is said to be **tempered** (or "essentially tempered"). An A-**parameter** is a map

$$\phi : \mathcal{L}_F \times \mathrm{SL}_2(\mathbf{C}) \to {}^LG$$

such that the restriction of ϕ to $\mathcal{L}_F$ is a tempered L-parameter. In particular, a tempered L-parameter is also an A-parameter:

$$\{L\text{-parameters}\} \cap \{A\text{-parameters}\} = \{\text{tempered parameters}\}.$$

By restriction, the local parameter ϕ_v is defined up to equivalence for all v.

Two A-parameters ϕ_1 and ϕ_2 are said to be **equivalent** if there exists $g \in \widehat{G}$ such that $g\phi_1(w, x)g^{-1} = z_w\phi_2(w, x)$, where $z_w \in Z(\widehat{G})$ and the class of the 1-cocycle $\{z_w\}$ in $H^1(\mathcal{L}_F, Z(\widehat{G}))$ is locally trivial. Note that if Γ_F acts trivially on $Z(\widehat{G})$, then $H^1(\mathcal{L}_F, Z(\widehat{G})) = \mathrm{Hom}(\mathcal{L}_F, Z(\widehat{G}))$; in this case, Cebotarev density for $L_F^{\mathrm{ab}} = W_F^{\mathrm{ab}}$ implies that there are no nontrivial elements of $\mathrm{Hom}(\mathcal{L}_F, Z(\widehat{G}))$ that are locally trivial.

Let $C(\phi)$ be the centralizer of $\mathrm{im}(\phi)$ in $\widehat{G}$, and let $S(\phi)$ be the group of $g \in \widehat{G}$ such that $\mathrm{ad}(g)$ induces a self-equivalence of ϕ. Then $S(\phi)$ contains $Z(\widehat{G})C(\phi)$, but it may be larger. For all places v, $S(\phi) \subset S(\phi_v)$. Observe that the image of $S(\phi)$ under an irreducible representation r of ${}^LG$ centralizes $r({}^LG)$. Indeed, r maps $Z(\widehat{G})$ to the scalars. Hence the image of a locally trivial cocycle $\{z_w\}$ under r defines a locally trivial element of

$\mathrm{Hom}(\mathscr{L}_F, \mathbf{C}^*)$, which is trivial. The elements of $S(\phi)$ therefore can be used to decompose the representations of $\mathscr{L}_F$ of the form $r \circ \phi$.

DEFINITION. An A-parameter ϕ is said to be **discrete** if $S(\phi)^0 \subset Z(\widehat{G})$.

Discreteness can be defined by the equivalent condition $C(\phi)^0 \subset Z(\widehat{G})$ **[K1]**.

3.2. Global packets. Assume the local correspondence (§1.10). If ϕ is a global L-parameter, then ϕ_v is unramified for almost all v and $\Pi(\phi_v)$ contains a unique unramified representation π_v^0 for almost all finite v. The global packet

$$\Pi(\phi) = \bigotimes \Pi(\phi_v)$$

is the set of restricted tensor products $\bigotimes \pi_v$ where $\pi_v \in \Pi(\phi_v)$ for all v and $\pi_v = \pi_v^0$ for almost all finite v.

If ϕ is an A-parameter, we define an L-parameter

$$\phi'(w) = \phi\left(w, \begin{pmatrix} |w|^{1/2} & 0 \\ 0 & |w|^{-1/2} \end{pmatrix}\right).$$

Here $|\cdot| : \mathscr{L}_F \to \mathbf{R}_+^*$ is the pull-back of the absolute value map of W_F. Arthur conjectures the existence of an enlarged L-packet or "A-packet" of representations $\prod(\phi_v)$ which would contain $\prod(\phi'_v)$ but could be larger if ϕ' is nontempered. For almost all v, π_v^0 would be the only unramified representation element in $\Pi(\phi_v)$ and the global A-packet $\Pi(\phi) = \otimes \prod(\phi_v)$ would be defined as above.

There are important reasons for introducing packets in the global theory beyond the formal considerations involving the Langlands group. First of all, they are necessary to stabilize the trace formula **[L3, A2]**. In addition, it appears that packets rather than representations are the objects that can be transferred functorially from one group to another. As observed above, there is a distinction between A-packets and L-packets only in the nontempered case. Indeed, the motivation for introducing A-packets is to specify which nontempered L-packets will intervene in the discrete spectrum, and to enlarge those L-packets so that their formalism resembles that of tempered L-packets. We refer to **[A1, A2, K1]** for further discussion as well as for the desiderata to be satisfied by all of these objects.

Two A-parameters ϕ_1 and ϕ_2 determine the same A-packet precisely when the local parameters ϕ_{1v} and ϕ_{2v} are locally equivalent at all places. In general, the map from global equivalence classes of parameters to packets is not one-to-one, even in the tempered case (see remarks below).

DEFINITION. Let Π be an A-packet. Then Π is called

(a) **discrete** if $\Pi = \Pi(\phi)$ for some discrete parameter ϕ,

(b) **tempered** if $\Pi = \Pi(\phi)$ for some tempered parameter ϕ.

An A-packet Π is a set of equivalence classes of representations of $G(\mathbf{A}_F)$, some of whose members may be automorphic and others not. Conjecturally, every discrete representation belongs to at least one discrete A-packet Π. Although local A-packets are certainly not disjoint, one may hope that in many cases global A-packets are disjoint. If so, we obtain a decomposition

$$L_d = \bigoplus L_d(\Pi)$$

where Π ranges over the set of global A-packets and $L_d(\Pi)$ is the direct sum of the discrete representations that are isomorphic to a member of Π.

EXAMPLES. (a) The trivial representation π_{triv} of $G(\mathbf{A})$ is discrete for all groups G. The associated A-parameter ϕ_{triv} is trivial on $\mathcal{L}_F$ and the image

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbf{C})$$

is a Γ_F-invariant regular unipotent element in $\widehat{G}$. Furthermore $S(\phi_{\text{triv}}) = Z(\widehat{G})$.

(b) For $G = \mathrm{GL}(n)$, A-packets and L-packets are defined to be singletons. Since ${}^LG = \mathrm{GL}_n(\mathbf{C}) \times W_F$, with the trivial Galois action, an A-parameter amounts to a homomorphism $\phi : \mathcal{L}_F \times \mathrm{SL}_2(\mathbf{C}) \to \mathrm{GL}(n)$ and $S(\phi) = C(\phi)$. It is discrete if and only if ϕ is an irreducible representation. The conjecture therefore asserts that the discrete automorphic representations of $\mathrm{GL}(n)$ correspond bijectively to equivalence classes of irreducible n-dimensional representations of $\mathcal{L}_F \times \mathrm{SL}_2(\mathbf{C})$. This would follow from the cuspidal correspondence and the results of **[MW]**. Cebotarev density for $\mathcal{L}_F$ would imply that everywhere local and global equivalence are the same.

(c) $G = \mathrm{SL}(n)$. Representations of $G(\mathbf{A})$ (resp. G_v) are said to belong to the same A- or L-packet if they are conjugate under the action of $\mathrm{PGL}_n(\mathbf{A})$ (resp. $\mathrm{PGL}_n(F_v)$) **[LL]**.

(d) The theory of A-packets on the group $G = U(3)$ is developed in **[R1, R2]**.

3.3. Weak transfers. Let $\varphi : {}^LH \to {}^LG$ be an L-map. If the above formalism were available, one would obtain a map $\phi \to \varphi \circ \phi$ from parameters on H to parameters on G and hence a transfer $\Pi(\phi) \to \Pi(\varphi)$ at the level of packets. Lacking the formalism, one can define a notion of weak transfer. For almost all finite v, φ yields a map from unramified parameters for H to unramified parameters for G and, hence, a map $\Pi_{\text{un}}(H_v) \to \Pi_{\text{un}}(G_v)$. An automorphic representation π of G is said to be a **weak transfer** of an automorphic representation σ of H if π_v is the transfer of σ_v with respect to φ for almost all unramified places. A weak form of the general functoriality conjecture asserts that weak transfers exist.

3.4. Multiplicity formulas. A key ingredient in the theory of automorphic forms is a remarkable multiplicity formula for members of a discrete A-packet Π. For each A-parameter ϕ, define the finite group

$$\mathfrak{S}(\phi) = \pi_0(S(\phi)/Z(\widehat{G})).$$

One postulates that for all global A-parameters ϕ such that $\Pi = \Pi(\phi)$ there exists

(1) a pairing
$$\langle\, ,\,\rangle : \Pi \times \mathfrak{S}(\phi) \to \mathbf{C}^*$$
such that $s \to \langle s, \pi\rangle$ is a character of a finite-dimensional representation of $\mathfrak{S}(\phi)$ for all $\pi \in \Pi$;

(2) a homomorphism $\varepsilon_\phi : \mathfrak{S}(\phi) \longrightarrow \{\pm 1\}$.

The character $\varepsilon_\phi(s)$ will be trivial if ϕ is tempered. In general, one decomposes $\mathrm{Lie}(\widehat{G})$ into a direct sum of irreducible representations σ_j under the joint action of $S(\phi) \times \mathscr{L}_F \times \mathrm{SL}_2(\mathbf{C})$. Write σ_j as a tensor product $\lambda_j \otimes \mu_j \otimes \nu_j$, and let $\varepsilon_j(s) = \det(\lambda_j(s))$. Assume that the L-function $L(s, \mu)$ of a representation of $\mathscr{L}_F$ satisfies a functional equation of the form $L(s, \mu) = \varepsilon(s, \mu_j)L(1-s, \widetilde{\mu})$. If μ is a self-dual representation, then the root number $\varepsilon(\frac{1}{2}, \mu)$ is ± 1. Define
$$\varepsilon_\phi(s) = \prod_{j \in T} \varepsilon_j(s)$$
where T is the set of indices such that σ_j is a self-dual representation and the root number $\varepsilon(\frac{1}{2}, \mu_j)$ is -1 [A2]. Observe that ε_ϕ will take values in $\{\pm 1\}$.

Conjecture. *Let π be a discrete automorphic representation. For each A-parameter ϕ such that $\pi \in \Pi(\phi)$, define*
$$m_\phi(\pi) = |\mathfrak{S}(\phi)|^{-1} \sum_{s \in \mathfrak{S}(\phi)} \varepsilon_\phi(s)\langle s, \pi\rangle.$$
The multiplicity $m(\pi)$ is equal to
$$m(\pi) = \sum_\phi m_\phi(\pi)$$
where ϕ ranges over the set of equivalence classes of discrete A-parameters such that $\pi \in \Pi(\phi)$.

In known examples, the global pairing is obtained from the local pairings which, in turn, are defined via local harmonic analysis. The local-global relation is given by a product formula
$$\langle s, \pi\rangle = \prod_v \langle s, \pi_v\rangle.$$
In general, one expects that the local signs $\langle s, \pi_v\rangle$ will be defined relative to local groups $S(\phi_v) = C(\phi_v)Z(\widehat{G})$ and
$$\mathfrak{S}(\phi_v) = \pi_0(S(\phi_v)/Z(\widehat{G})).$$

Since there will exist a homomorphism from $\mathfrak{S}(\phi)$ to $\mathfrak{S}(\phi_v)$, $\langle s, \pi_v\rangle$ will be defined for $s \in \mathfrak{S}(\phi)$. A modification of this procedure will be necessary if G is not quasi-split.

In the known cases, the local pairings depend on certain choices (e.g., a choice of local additive character ψ_v for all v). Their product is independent of the choices, provided that the local data are chosen compatibly (e.g., the ψ_v are obtained from a global additive character of $\mathbf{A}_F/F$).

REMARKS. (1) Since we hypothesize that $\langle \bullet, \pi\rangle$ is the character of a finite-dimensional representation of $\mathfrak{S}(\phi)$, the numbers $m_\phi(\pi)$ are all integers.

(2) If the global pairing is a product of local pairings as above, then the numbers $m_\phi(\pi)$ will presumably depend only on Π and not on the choice of ϕ. If $\langle \bullet, \pi\rangle$ is an irreducible character, then $m_\phi(\pi)$ is 0 or 1 and the multiplicity of π will be equal to the number of globally inequivalent parameters ϕ such that $\pi \in \Pi(\phi)$.

(3) If global A-packets are disjoint and $\pi \in \Pi(\phi)$, then $m(\pi) = n(\phi)m_\phi(\pi)$ where $n(\phi)$ is the number of equivalence classes of global parameters that are everywhere locally equivalent to ϕ.

(4) Multiplicity one holds for the discrete spectrum of $\mathrm{GL}(n)$ (see example (b) above). The above conjecture would imply that it holds whenever everywhere local equivalence implies global equivalence of parameters and the characters $\langle \bullet, \pi\rangle$ are irreducible. Examples of cuspidal representations having multiplicity greater than one have been constructed by Labesse and Langlands **[LL]** for the norm one subgroups of nonsplit quaternion algebras. In these cases, the characters $\langle \bullet, \pi\rangle$ fail to be irreducible. In **[B]** one constructs cuspidal representations having multiplicity greater on $\mathrm{SL}(n)$ for $n \geq 3$. The higher multiplicity in this case is due to the existence of globally inequivalent parameters that are locally everywhere equivalent. Multiplicity one is conjectured to hold for $\mathrm{SL}(2)$.

(5) We call an A-packet Π **stable** if $\mathfrak{S}(\phi) = \{1\}$. In this case, $m_\phi(\pi)$ is constant for $\pi \in \Pi(\phi)$.

(6) The multiplicity formula for tempered L-packets was known to Langlands in the 1970s, who defined endoscopic groups and the group $\mathscr{L}_F$ **[L4]** and worked out the contribution of the discrete series L-packets to the zeta function (cf. **[L1, Sh1]**). However, it was not clear how to take account of the nontempered spectrum before Arthur introduced A-parameters which included the $\mathrm{SL}_2(\mathbf{C})$-factor and the characters ε_Π.

3.5. Restriction of scalars. Let $G' = \mathrm{Res}_{F/\mathbf{Q}}(G)$. The formalism of Shimura varieties is usually stated in terms of groups over $\mathbf{Q}$, i.e., in terms of G' rather than G. Of course, the groups $G(\mathbf{A}_F)$ and $G'(\mathbf{A}_\mathbf{Q})$ are canonically isomorphic and their automorphic representations are the same. Since it is often convenient to shift between the two points of view, it is useful to make explicit the bijection between equivalence classes of A-parameters for the two groups. To define the correspondence, which is essentially a case of

Shapiro's lemma, observe that $\widehat{G'} = \widehat{G}^d$ where $d = [F : \mathbf{Q}]$. Fix representatives $\sigma_1, \dots, \sigma_d$ for $\mathscr{L}_{\mathbf{Q}}/\mathscr{L}_F$, and let $\overline{\sigma}_1, \dots, \overline{\sigma}_d$ be their images in $\Gamma_{\mathbf{Q}}$. For $\gamma \in \Gamma_{\mathbf{Q}}$, let $\gamma(j)$ denote the permutation of $\{1, \dots, d\}$ such that

$$\gamma\overline{\sigma}_{\gamma^{-1}(j)}\Gamma_F = \overline{\sigma}_j\Gamma_F$$

and set

$$\gamma_j = \overline{\sigma}_j^{-1}\gamma\overline{\sigma}_{\gamma^{-1}(j)} \in \Gamma_F.$$

Then γ acts by

$$(a_1, \dots, a_d) \to (\gamma_1(a_{\gamma^{-1}(1)}), \dots, \gamma_d(a_{\gamma^{-1}(d)})).$$

Define $\gamma(j)$ and γ_j for $\gamma \in \mathscr{L}_F$ similarly. If $\phi : \mathscr{L}_F \times \mathrm{SL}_2(\mathbf{C}) \to {}^LG$ is an A-parameter for G, then

$$\phi' : \mathscr{L}_{\mathbf{Q}} \times \mathrm{SL}_2(\mathbf{C}) \to {}^LG'$$

is defined as follows. For $\gamma \times g \in \mathscr{L}_{\mathbf{Q}} \times \mathrm{SL}_2(\mathbf{C})$, set

$$\phi'(\gamma \times g) = (\phi(\gamma_1 \times g), \dots, \phi(\gamma_d \times g)) \times \overline{\gamma}.$$

The diagonal embedding of $\widehat{G}$ into $\widehat{G'}$ induces the top row of the commutative diagram

$$\begin{array}{ccc} S(\phi) & \longrightarrow & S(\phi') \\ \downarrow & & \downarrow \\ \ker^1(\mathscr{L}_F, Z(\widehat{G})) & \longrightarrow & \ker^1(\mathscr{L}_{\mathbf{Q}}, Z(\widehat{G'})). \end{array}$$

Here $\ker^1$ denotes the set of classes in H^1 that are locally trivial. The bottom row is obtained from Shapiro's Lemma. Hence $S(\phi')$ is $Z(\widehat{G'})$ times the image of $S(\phi)$. This proves that the diagonal map yields an isomorphism

$$\mathfrak{S}(\phi) \xrightarrow{\sim} \mathfrak{S}(\widehat{\phi}).$$

3.6. Tempered spectrum. A unitary representation π of G_v is called **tempered** if all of its matrix coefficients $\phi(g)$ satisfy $|\phi| \in L^{2+\varepsilon}(Z_v\backslash G_v)$ for all $\varepsilon > 0$. It is known that the support of the Plancherel measure of G_v is precisely the set of tempered representations. In the global setting, an automorphic representation $\pi = \bigotimes \pi_v$ is called **tempered** if π_v is tempered for all v. Conjecturally, the tempered discrete representations are precisely those that belong to tempered packets. The hypothesis that $\mathscr{L}_F$ is an extension of a compact group by $\mathbf{R}_+^*$ equipped with embeddings of the groups L_{F_v} is essentially equivalent to the following:

RAMANUJAN CONJECTURE. *Every unitary cuspidal representation of* $\mathrm{GL}_n(\mathbf{A}_F)$ *is tempered.*

This is, of course, a generalization of the classical Ramanujan conjecture on the size of the Fourier coefficients of the Ramanujan Δ-function. The classical conjecture and its extension to holomorphic modular forms of arbitrary weight and level on GL(2) was proved by Deligne by reducing to the Riemann hypothesis over finite fields. The same method evidently applies whenever the cuspidal representation can be related via L-series to a Galois representation for which the Riemann hypothesis is known to hold. In the general setting, it would follow from the existence of functorial transfers for cuspidal representations of general linear groups **[L6]**. However, there is no known method for proving the existence of such transfers.

The discrete representations that occur in the residual spectrum are known to be nontempered; hence, all such would belong to nontempered A-packets. For a general reductive group G, the cuspidal spectrum may also contain nontempered representations. For example, if $G(\mathbf{R})$ is a group of Hermitian symmetric type and π is a cuspidal representation that has cohomology outside the middle dimension, then π_∞ is nontempered. Numerous examples of nontempered representations have been constructed using the oscillator representation. See, for example, **[HPS]**, **[Li]**. Recently, another method has been found **[BSL]**.

Arthur's conjecture puts a serious limitation on the kind of nontempered representations that can occur discretely. A simple consequence of the conjecture is that a discrete representation π that is nontempered at one place v is necessarily nontempered at almost all places. We verify this in a special case below (Proposition. 6.2).

The general functoriality conjecture implies that if π is a **unitary** discrete representation of a general connected reductive group G, then the automorphic L-functions of π can be factored:

$$L(s, \pi, r) = \prod_j L(s + s_j, \pi_j)$$

where π_j is a unitary cuspidal representation of $\mathrm{GL}_{n_j}(\mathbf{A}_F)$ and $s_j \in \mathbf{R}$. The Ramanujan conjecture says that each $L(s, \pi_j)$ is pure (in the sense that each local factor is tempered) and hence that $L(s, \pi, r)$ is a product of twists of "pure" L-functions. Arthur's conjecture has as a further important consequence that the twists s_j always occur in "Lefschetz orbits", by which we mean that $L(s, \pi, r)$ is a product of factors of the form

$$L(s - \tfrac{m-1}{2}, \pi')L(s - \tfrac{m-3}{2}, \pi') \cdots L(s - \tfrac{1-m}{2}, \pi')$$

where π' is a unitary cuspidal representation of GL_d for some d.

4. Cohomological parameters

An Archimedean A-parameter

$$\phi : W_{\mathbf{R}} \times \mathrm{SL}_2(\mathbf{C}) \to {}^L G$$

is called **cohomological** if the associated A-packet $\Pi(\phi)$ consists of cohomological representations. The classification of unitary cohomological representations is due to Vogan-Zuckerman **[VZ]**. Their partition into A-packets is due to Adams and Johnson **[AJ, A1]**.

4.1. We assume that $G_{\mathbf{R}}$ has an $\mathbf{R}$-elliptic maximal torus, i.e., a maximal torus T defined over $\mathbf{R}$ such that $T(\mathbf{R})$ is compact modulo the center of $G(\mathbf{R})$. Let $(\widehat{S}, \widehat{B})$ be a Borel pair in $\widehat{G}$ fixed by the Galois action. The choice of a Borel subgroup B containing T determines an isomorphism

$$\widehat{T} \xrightarrow{\sim} \widehat{S} \tag{4.1.1}$$

as in §1.5. This also determines a bijection $P \leftrightarrow \widehat{P}$ between parabolic subgroups containing T and parabolic subgroups containing $\widehat{S}$. Let M_P and $\widehat{M}$ denote the unique Levi factors of P and $\widehat{P}$ containing T and $\widehat{S}$, respectively. The above isomorphism extends to an isomorphism $\widehat{M}_P \xrightarrow{\sim} \widehat{M}$.

Up to equivalence, the cohomological A-parameters are indexed by parabolic subgroups P containing B. We now review this indexing. Denote the parameter associated to P by ϕ_P, and set $\Pi_P = \Pi(\phi_P)$. Note that M is defined over $\mathbf{R}$, although P is not defined over $\mathbf{R}$, since complex conjugation sends each root of T to its negative.

We shall define an embedding

$$\xi_P : {}^L M \to {}^L G$$

as follows. The restriction of ξ_P to $\widehat{M}$ is the given inclusion. For any Levi subgroup R containing T, let δ_R be the half-sum of the positive co-roots of $\widehat{S}$ contained in the corresponding Levi factor $\widehat{R}$. Let n_R be an element of $\widehat{R}_{\mathrm{ad}} \subset \widehat{R}$ that maps the positive roots of $\widehat{S}$ in $\widehat{R}$ to the negative roots. Such an element exists since M contains an $\mathbf{R}$-elliptic maximal torus. The real Weil group $W_{\mathbf{R}}$ is a union $\mathbf{C}^* \cup \mathbf{C}^* w_{\sigma_\infty}$ where w_{σ_∞} is an element of square -1 that maps to complex conjugation $\sigma_\infty \in \mathrm{Gal}(\mathbf{C}/\mathbf{R})$. Set $\widehat{\delta}_P = \delta_G - \delta_M$ and define

$$\xi_P(z) = \left(\frac{z}{\overline{z}}\right)^{\widehat{\delta}_P} = z^{2\widehat{\delta}_P}(z\overline{z})^{-\widehat{\delta}_P},$$

$$\xi_P(w_{\sigma_\infty}) = n_G n_M^{-1} \rtimes w_{\sigma_\infty}.$$

Let

$$\phi_{M,triv} : W_{\mathbf{R}} \times \mathrm{SL}_2(\mathbf{C}) \to {}^L M$$

be the A-parameter corresponding to the trivial representation of M, and set

$$\phi_P = \xi_P \circ \phi_{M,\mathrm{triv}}.$$

Thus Π_P is the functorial transfer relative to ξ_P in $G(\mathbf{R})$ of the trivial representation of $M(\mathbf{R})$.

Remarks. (1) The A-packet ϕ_P is tempered if and only if $P = B$. In this case, Π_B is the set of discrete series representations whose infinitesimal character (resp., central character) coincides with the infinitesimal character (resp., central character) of the trivial representation.

(2) We have

$$S(\phi_P) = Z(\widehat{G})Z(\widehat{M})^{\sigma_\infty} \subset \widehat{S}$$

where $Z(\widehat{M})^{\sigma_\infty}$ denotes the invariants under the action $\mathrm{ad}(\xi_P(w_{\sigma_\infty}))$. In particular, $S(\phi)$ is abelian. To verify this, observe that $\widehat{M}$ is the centralizer of $\phi_P(\mathbf{C}^*)$ since $\langle \alpha, \hat{\delta}_P \rangle \neq 0$ for all roots that do not belong to M. The centralizer of $\phi_P(\mathbf{C}^* \times \mathrm{SL}_2(\mathbf{C}))$ is $Z(\widehat{M})$ and hence $C(\phi_P) = Z(\widehat{M})^{\sigma_\infty}$. Furthermore, $\mathfrak{S}(\phi_P) = Z(\widehat{M})^{\sigma}/Z(\widehat{G})^{\sigma}$ is a finite abelian 2-group.

(3) A global parameter ϕ is called cohomological if ϕ_v is cohomological for all $v \in S_\infty$. If ϕ is cohomological, then $S(\phi)$ is abelian since $S(\phi)$ is a subgroup of $S(\phi_v)$ for $v \in S_\infty$. For the same reason, $\mathfrak{S}(\phi)$ is a finite group and it follows that a global cohomological A-parameter must be discrete.

4.2. Parametrization of Π_P. The members of Π_P are parametrized by the double coset space

$$\Omega(T, M)\backslash\Omega_{\mathbf{C}}(T)/\Omega_{\mathbf{R}}(T)$$

where $\Omega_{\mathbf{C}}(T)$ is the Weyl group of T in G, $\Omega_{\mathbf{R}}(T)$ is the subgroup of elements in $\Omega(T)$ with a representative in $G(\mathbf{R})$, and $\Omega(T, M)$ is the Weyl group of T in M. The element π_w of Π_P corresponding to the double coset with representative w is the derived functor module $A_{w^{-1}Pw}$ [**A1**].

4.3. Hodge types. The relative Lie algebra cohomology of π_w has a Hodge decomposition that can be described explicitly. Let K be the centralizer of μ in $G(\mathbf{R})$. Let $\mathscr{P}^+$ and $\mathscr{P}^-$ be the subspaces of $\mathfrak{G}$ on which $\mathrm{ad}(\mu(t))$ acts by t^{-1} and t, respectively. Then

$$H^{p,q}(\mathfrak{G}, K, \pi_w) = \mathrm{Hom}_K(\Lambda^p\mathscr{P}^+ \otimes \Lambda^q\mathscr{P}^-, \pi_w). \tag{4.3.1}$$

The dimension of $H^{p,q}(\mathfrak{G}, K, \pi_w)$ is computed as follows. Let $G(\mathbf{R})' = G(\mathbf{R})^0K$. The set of restrictions of elements of Π_P to $G(\mathbf{R})'$ is parametrized by

$$\Sigma' = \Omega(T, M)\backslash\Omega_{\mathbf{C}}(T)/\Omega(T, K),$$

where $\Omega(T, K)$ is the Weyl group of T in K. Choose a set of representatives $\{w\}$ of minimal length for Σ' and let π'_w be the representation corresponding to w. Let $\{r\}$ be a set of representatives of minimal length for the cosets $\Omega(T, M)/\Omega(T, M \cap wKw^{-1})$.

Let $\mathscr{N}_w$ be the Lie algebra of the unipotent radical of $w^{-1}Pw$ and set

$$p_w = \dim(\mathscr{N}_w \cap \mathscr{P}^+), \qquad q_w = \dim(\mathscr{N}_w \cap \mathscr{P}^-).$$

According to [**A1**, §9], $\dim(H^{p,q}(\mathfrak{G}, K, \pi'_w))$ is equal to the number of r such that $(p, q) = (p_w + \ell(r), q_w + \ell(r))$.

4.4. Generic representations. The local pairings that appear in the multiplicity formula are not canonically defined in general. However, if G_v is quasi-split, then one expects that Whittaker models will provide a means for choosing base points in an A-packet and that the pairing can be uniquely determined relative to such a choice. Let F be a local field and assume that G_F is quasi-split. Let B be a Borel subgroup defined over F and let N be its unipotent radical. A character $\psi : N(F) \to \mathbf{C}^*$ is called nondegenerate if it has a nontrivial restriction to all one-parameter subgroups of N corresponding to simple roots. A ψ-Whittaker functional for an irreducible admissible representation (π, V) of $G(F)$ is a linear functional $\lambda : V \to \mathbf{C}$ such that $\lambda(\pi(n)v) = \psi(n)\lambda(v)$ for all $v \in V$, $n \in N(F)$. The dimension of the space of ψ-Whittaker functionals is at most one. When the dimension is one, π is called **generic** or ψ-generic. We assume the following hypothesis.

HYPOTHESIS. *For a given choice of nondegenerate ψ, a tempered L-packet contains a unique ψ-generic representation π^{gen}.*

This is known to be true in the Archimedean case **[ABV]**.

A distinguished element π^{gen} can also be defined for A-packets whose restriction to $\mathrm{SL}_2(\mathbf{C})$ is nontrivial. However, π^{gen} may not be generic since $\Pi(\phi)$ need not contain a generic representation. An example is provided by the A-packet consisting of the trivial representation alone. However, recall that $\Pi(\phi)$ is to be an enlargement of the nontempered L-packet $\Pi(\phi')$. Attached to ϕ' is a parabolic subgroup Q and a tempered L-packet ρ on a Levi factor of Q such that for all $\sigma \in \rho$, the parabolically induced representations $\mathrm{Ind}_Q^G(\sigma)$ each have a unique irreducible quotient, the so-called Langlands quotient. The elements of $\Pi(\phi')$ are the Langlands quotients of the representations $\mathrm{Ind}_Q^G(\sigma)$ as we vary σ in ρ. Let π^{gen} be the Langlands quotient of $\mathrm{Ind}_Q^G(\sigma^{\text{gen}})$ where σ^{gen} is the unique ψ-generic member of ρ.

Suppose that G is globally quasi-split. Fix an additive character ψ of $\mathbf{A}/F$. If ϕ is a global A-parameter, then

$$\pi^{\text{gen}} = \bigotimes \pi_v^{\text{gen}}$$

should serve as a base point of the global packet $\Pi(\phi)$. In particular, the character $\langle \bullet', \pi^{\text{gen}} \rangle$ of $\mathfrak{S}(\phi)$ should be trivial because it is locally trivial.

§5. The zeta-function conjecture

5.1. The representation r. Fix an element $h : \mathscr{R} \to G_{\mathbf{R}}$ of $\mathfrak{X}$. Over $\mathbf{C}$, we have $\mathscr{R}_{\mathbf{C}} = \mathbf{G}_m \times \mathbf{G}_m$, where the ordering of the factors is such that the first factor corresponds to the identity embedding $\mathbf{C} \to \mathbf{C}$. Let μ be the restriction of $h_{\mathbf{C}}$ to the first factor. Up to conjugacy, we may assume that the image of μ is contained in a maximal torus T of G defined over $\mathbf{Q}$, and hence that μ itself is defined over $\overline{\mathbf{Q}}$. Let $\Gamma_\mu \subset \Gamma_{\mathbf{Q}}$ be the subgroup of elements fixing the conjugacy class of μ. By definition, the *reflex field*

E is the subfield of $\overline{\mathbf{Q}}$ corresponding to Γ_μ. Observe that E is a subfield of any extension K of $\mathbf{Q}$ over which G splits. This follows because up to conjugacy, μ may be chosen to take values in a maximal torus of G which splits over K.

As before, $(\widehat{B}, \widehat{S})$ denotes a Borel pair that is fixed by the Galois action. Under the identification $X^*(\widehat{T}) = X_*(T)$, μ defines a character of $\widehat{T}$. Using the Weyl group orbit of canonical isomorphisms $\widehat{T} \to \widehat{S}$, we may associate to μ an $\Omega(\widehat{T})$-orbit of characters of $\widehat{S}$. Let

$$\widehat{\mu} : \widehat{T} \to \mathbf{C}^*$$

be the unique element in this orbit that is dominant relative to $\widehat{B}$. Let

$$r : \widehat{G} \to \mathrm{GL}(V)$$

be the irreducible representation whose highest weight relative to $(\widehat{B}, \widehat{S})$ is $\widehat{\mu}$. The action of $\Gamma_{\mathbf{Q}}$ preserves the set of dominant weights of $\widehat{T}$ and hence Γ_E fixes $\widehat{\mu}$. It follows that Γ_E fixes the isomorphism class of r, and this yields a projective representation of Γ_E on V. There is a unique lifting of this projective action to an action Γ_E on V fixing the one-dimensional $\widehat{\mu}$-weight space pointwise. This defines an extension of r to a representation of the L-group ${}^LG_E$, which we continue to denote by r.

EXAMPLES.

(a) $G = \mathrm{GL}(2)_{\mathbf{Q}}$, $h(a+bi) = \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$. Over $\mathbf{C}$, h can be diagonalized to $(z, w) \to \left(\begin{smallmatrix} z & 0 \\ 0 & w \end{smallmatrix}\right)$ and $\mu(t) = \left(\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Relative to the standard Borel pairs in G and $\widehat{G}$, $\widehat{\mu}(\left(\begin{smallmatrix} x & 0 \\ 0 & y \end{smallmatrix}\right)) = x$, the highest weight of the standard two-dimensional representation of $\widehat{G}$. The representation r is the extension of the standard representation to ${}^LG$ which is trivial on $W_{\mathbf{Q}}$.

(b) $G = \mathrm{GU}(\Phi)$ where Φ is a Hermitian form in n variables with respect to a quadratic imaginary extension $K/\mathbf{Q}$. Suppose that the signature of Φ at infinity is (p, q). We can assume that Φ is diagonal of the form

$$\begin{pmatrix} \alpha I_p & 0 \\ 0 & \beta I_q \end{pmatrix}$$

where $\alpha > 0$, $\beta < 0$. Identifying $G(\mathbf{C})$ with $\mathrm{GL}_n(\mathbf{C}) \times \mathrm{GL}_1(\mathbf{C})$, we may define

$$h(z, w) = \begin{pmatrix} zI_p & 0 \\ 0 & wI_q \end{pmatrix} \times zw.$$

Relative to the standard Borel pairs in G and $\widehat{G} = \mathrm{GL}_n(\mathbf{C}) \times \mathrm{GL}_1(\mathbf{C})$,

$$\widehat{\mu}\left(\begin{pmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{pmatrix} \times \mu\right) = t_1 t_2 \cdots t_p \mu.$$

This is the highest weight of $\widehat{G}$ on $\Lambda^p(r_{\mathrm{st}})$ where r_{st} the standard n-dimensional representation of $\mathrm{GL}_n(\mathbf{C})$ and $\mu \in \mathrm{GL}_1(\mathbf{C})$ acts by μ. The reflex field E is K if $p \neq q$ and $E = \mathbf{Q}$ if $p = q$. In the former case, r is the representation of ${}^L G_E$ on $\Lambda^p(r_{\mathrm{st}})$ on which W_E acts trivially and if $p = q$, the full L-group ${}^L G$ acts on $\Lambda^p(r_{\mathrm{st}})$.

(c) $G = \mathrm{GSp}(2n)_{\mathbf{Q}}$, defined relative to the usual symplectic form, and

$$h(a+bi) = \begin{pmatrix} aI_n & bI_n \\ -bI_n & aI_n \end{pmatrix}.$$

Over $\mathbf{C}$, h can be diagonalized to

$$(z, w) \to \begin{pmatrix} wI_n & 0 \\ 0 & zI_n \end{pmatrix}.$$

The representation r of $\widehat{G} = \mathrm{GSpin}(2n+1)$ is the spin representation of $\widehat{G}$ of dimension 2^n. Let T be the torus in G, consisting of diagonal matrices t with entries $t_1, \dots, t_n, \nu t_1^{-1}, \dots \nu t_n^{-1}$. Then $X_*(T)$ is the span of the cocharacters $\lambda_0, \lambda_1, \dots, \lambda_n$ where

$$\lambda_0(x) = \begin{pmatrix} I_n & 0 \\ 0 & xI_n \end{pmatrix}$$

and $\lambda_j(x)$ has $t_j = x$, $t_i = 1$ for $i \neq j$ and $\nu = 1$. An extremal weight of r is λ_0, viewed as a character of $\widehat{T}$. The set of weights of r is the orbit of λ_0 under the Weyl group and consists of the characters $\lambda_0 \prod_{i \in S} \lambda_i$ where S ranges over the subsets of $\{1, \dots, n\}$. In particular, if χ is an unramified character of $T(\mathbf{Q}_p)$ defined by

$$\chi(t) = z_0^{\mathrm{val}(\nu)} \prod z_j^{\mathrm{val}(t_j)}$$

then $r(g(\pi(\chi)))$ is the matrix with eigenvalues $z_0 \prod_{i \in S} z_i$.

5.2. Eigenspaces. Let $\phi : \mathscr{L}_{\mathbf{Q}} \times \mathrm{SL}_2(\mathbf{C}) \to {}^L G$ be a cohomological A-parameter, and let $\Pi = \Pi(\phi)$ be the associated A-packet. Recall that

$$L(s, \Pi, r) = L(s, r \circ \phi')$$

where ϕ' is the Langlands parameter associated to π. The conjecture asserts that $L(s, \rho_\lambda^i(\pi_f))$ is a factor of $L(s, \Pi, r)$. To pick out the appropriate factor, we first decompose $L(s, \Pi, r)$ into a product of pure L-functions. (by a pure L-function, we mean an L-function that can be twisted so that all of its local factors have the form $\det(1 - A_v q_v^{-s})^{-1}$ where A_v is a unitary matrix). Let $d = \dim S_K$. Choose an embedding $w : \mathbf{G}_m \to \mathrm{SL}_2(\mathbf{C})$, and decompose $V(\pi_f)$ according to the action of $\mathbf{G}_m$:

$$V = \bigoplus V^i, \quad \text{where} \quad V^i = \{v \in V : \phi(w(t))v = t^{i-d} v \text{ for } t \in \mathbf{G}_m\}.$$

Let r^i be the restriction of r to V^i. Then

$$L(s, \Pi, r) = \prod_i L(s, r^i \circ \phi).$$

The group $S(\phi)$ is abelian. Let $\{\chi\}$ be the set of characters that appear in the restriction of r to $S(\phi)$, and let $V(\chi)$ be the χ-eigenspace. Then

$$V = \bigoplus V(\chi) \quad \text{and} \quad V(\chi) = \bigoplus V^i(\chi)$$

where $V^i(\chi) = V^i \cap V(\chi)$. Let $\phi^i(\chi)$ denote the representation of $\mathscr{L}_E$ on $V^i(\chi)$. Up to equivalence, this representation is independent of the choice of w.

In §5.3 we associate to each $\pi_f \in \Pi(\phi)_f$ a function $\chi_\phi(\pi_f)$ on $S(\phi)$ which, conjecturally, is the character of a possibly reducible finite-dimensional representation $\sigma(\pi_f)$ of $S(\phi)$. Set

$$m(\chi, \pi_f) = \text{multiplicity of } \chi \text{ in } \sigma(\pi_f).$$

CONJECTURE 5.2. *For almost all places v of E, the following local factors are equal*

$$L_v(s, \rho^i_\lambda(\pi_f)) = \prod_\phi \prod_\chi L_v\left(s - \frac{i}{2}, \phi^i(\chi)\right)^{m(\chi, \pi_f)}$$

where the product is over the cohomological parameters ϕ such that $\Pi(\phi)_f$ contains π_f and χ ranges over the characters of $S(\phi)$ appearing in V.

In this statement of the conjecture we follow **[K3]**. A conjecture of this type, phrased in terms of endoscopic groups, was first formulated by Langlands in the mid 1970s for the case of L-packets whose Archimedean component is a discrete series L-packet. At that time, the Vogan-Zuckerman classification did not exist and it was not clear how to extend the conjecture to take into account the contributions of nontempered representations.

5.3. The character $\chi_\phi(\pi_f)$. Set $\chi = \chi_\phi(\pi_f)$. We assume the global pairing $\langle s, \pi\rangle$ is given as a product of local pairings $\prod_v \langle s, \pi_v\rangle$. Then χ is of the form

$$\chi(s) = \varepsilon_\varphi(s)\chi_\infty(s)\langle s, \pi_f\rangle$$

where $\langle s, \pi_f\rangle = \prod_{v<\infty}\langle s, \pi_v\rangle$ and χ_∞ is the restriction to $S(\phi)$ of a certain character of $S(\phi_\infty)$ described below. Recall that $\varepsilon_\phi(s)$ is assumed to be a character with values in $\{\pm 1\}$. The function $\langle s, \pi_v\rangle$ is assumed to be the character of a finite-dimensional representation of $\mathfrak{S}(\phi_v)$ which will be trivial for almost all v.

To define χ_∞, let (B, T) be a Borel pair in G such that T is an $\mathbf{R}$-elliptic maximal torus as in §4.1 and suppose that μ takes values in T. Let $\widehat{\mu}$ be the corresponding character of $\widehat{S}$ determined via (4.1.1). The elements μ and $\widehat{\mu}$ are determined up to conjugation by an element in the normalizer of T in $G(\mathbf{R})$, i.e., by an element in $\Omega_{\mathbf{R}}(T)$. Furthermore, the restriction of $\widehat{\mu}$ to $\widehat{T}^{\sigma_\infty}$ is independent of the choice of μ **[K3,** §5]. In other words, for

$w \in \Omega_{\mathbf{C}}(T)$ the character of $Z(\widehat{G})\widehat{S}^{\sigma_\infty}$ defined by $s \to \widehat{\mu}(w^{-1}sw)$ depends only on the image of w in $\Omega_{\mathbf{C}}(T)/\Omega_{\mathbf{R}}(T)$.

If $\pi_\infty \in \Pi_P$, then π_∞ is a derived functor module A_Q for some parabolic subgroup $Q = w^{-1}Pw$ where $w \in \Omega_{\mathbf{C}}(T)$. Let $\widehat{\mu}_{\pi_\infty}$ denote the restriction of the character $\widehat{\mu}(w^{-1}sw)$ to $S(\phi_P) = Z(\widehat{G})Z(\widehat{M})^{\sigma_\infty}$. This depends only on the image of w in $\Omega(T, M)\backslash\Omega_{\mathbf{C}}(T)/\Omega_{\mathbf{R}}(T)$. Kottwitz observes that the character $\widehat{\mu}_{\pi_\infty}(s)\langle s, \pi_\infty\rangle$ is independent of the choice of $\pi_\infty \in \Pi_P$ **[K3, Lemma 9.2]**. We set

$$\chi_\infty = \widehat{\mu}_{\pi_\infty}\langle \bullet, \pi_\infty\rangle.$$

If G is quasi-split, we may assume that the local pairings are normalized with respect to a choice of global additive character. Let π_∞^{gen} be the distinguished base point of Π_P as in §4.4. Then $\langle \bullet, \pi_\infty^{\text{gen}}\rangle$ is trivial and $\chi_\infty = \widehat{\mu}_{\pi_\infty^{\text{gen}}}$.

5.4. More explicit statement. The primary role of the Langlands group in the above formulation is to provide a means for factoring $L(s, \pi, r)$ since it is the individual factors that are L-functions of motives occurring in the varieties S_K. Fortunately, there is a down-to-earth way of factoring an L-function which does not make use of the Langlands group. Suppose that π is a weak transfer of an automorphic representation σ of a group H with respect to an L-map $\varphi : {}^LH \to {}^LG$. Furthermore, let $S \subset \widehat{G}$ be a subgroup of $\widehat{G}$ that centralizes $\varphi(\widehat{H})$ and such that the cocycles $\varphi(z)s\varphi(z)s^{-1}$ are locally trivial for all $s \in S$. Then $r \circ \varphi$ decomposes as a sum $\bigoplus r_\chi$ over the characters χ of S, and this yields a factorization

$$L(s, \pi, r) = \prod L(s, \sigma, r_\chi).$$

The groups H that arise in the theory are called **endoscopic groups**. Since it has been our intention to avoid entering into the details of this subject, we refer to **[K1, K3]** for an excellent discussion.

Example. Let $U(n)$ be the quasi-split unitary group in n variables with respect to a quadratic extension E/F. The endoscopic groups are the groups $U(n_1) \times U(n_2)$ where $n_1 + n_2 = n$. A discrete parameter ϕ is stable ($\mathfrak{S}(\phi) = \{1\}$) if and only if the restriction ϕ_E to $\mathscr{L}_E \times \mathrm{SL}_2(\mathbf{C})$ is irreducible. In this case, $L(s, \phi_E) = L(s, \pi_E)$ where π_E is a discrete automorphic representation of $\mathrm{GL}(n)_E$. It is simple to show that every discrete parameter $\phi : \mathscr{L}_F \times \mathrm{SL}_2(\mathbf{C}) \to {}^LG$ factors through an embedding

$$\varphi : {}^LH' \to {}^LU(n)$$

where $H' = U(n_1) \times \cdots \times U(n_r)$ for some partition $n = n_1 + \cdots + n_r$ **[R3]**. To define this embedding, it is necessary to use the Weil form of the L-group. We have $\phi = \varphi \circ \phi_{H'}$ where $\phi_{H'}$ is a parameter for H which factors as $\phi_1 \times \cdots \times \phi_r$ and the ϕ_j are inequivalent stable parameters for $U(n_j)$. In this case, $\mathfrak{S}(\phi) = (\mathbf{Z}/2)^{r-1}$.

In the case of the unitary group $U(3)$, A-packets have been defined and their classification corresponds to a pattern suggested by the parameter formalism [R1]. Thus, the A-packets Π fall into three types, according as the standard L-function $L(s, \Pi_E)$ (Π = base change of Π to E) is of the form:

(1) $\prod_{j=1}^{3} L(s, \chi_j)$ where the χ_j are distinct Hecke L-series of E ($H' = U(1)^3$, $\mathfrak{S} = (\mathbf{Z}/2)^2$).

(2) $L(s, \sigma)L(s, \chi)$ where σ is a discrete automorphic representation of $GL(2)_E$ and χ is a Hecke L-series of E ($H' = U(2) \times U(1)$, $\mathfrak{S} = \mathbf{Z}/2$).

(3) $L(s, \pi)$ where π is a discrete automorphic representation of $\mathrm{GL}(3)_E$ ($H' = U(3)$, $\mathfrak{S} = \{1\}$).

5.5. Justification for r. The appearance of the representation r with highest weight $\widehat{\mu}$ can be justified by Archimedean and p-adic considerations. In the Archimedean case, let π_f be an element in the finite part of a cohomological A-packet $\Pi(\phi)$, and let $\mathrm{Inf}_\phi(\pi_f)$ be the set of $\pi_\infty \in \Pi(\phi)_\infty$ such that $\pi_\infty \otimes \pi_f$ occurs discretely. The vector space $H(\pi_f) \otimes \mathbf{C}$ is isomorphic to a sum over ϕ of the cohomology spaces

$$H_\phi(\pi_f) = \bigoplus_{\pi_\infty \in \mathrm{Inf}_\phi(\pi_f)} m(\pi_\infty \otimes \pi_f) H^*(\mathfrak{G}, K, \pi_\infty).$$

The Killing form restricted to $\mathscr{P}^+ \times \mathscr{P}^-$ yields a $(1, 1)$ class in $H^*(\mathfrak{G}, K, \pi_{\mathrm{triv}})$ (cf. (4.3.1)). Cup product with this class yields a Lefschetz decomposition and hence an action of $\mathrm{SL}_2(\mathbf{C})$ on $H_\phi(\pi_f)$ such that $\left(\begin{smallmatrix} t & \\ & t^{-1} \end{smallmatrix}\right)$ acts by t^{d-i} in degree i. Define an action of $\mathbf{C}^*$ on $H_\phi(\pi_f)$ by letting z act by $z^{\frac{q-p}{2}} \overline{z}^{\frac{p-q}{2}}$ on the part of Hodge type (p, q). This yields a representation σ_ϕ of $\mathbf{C}^* \times \mathrm{SL}_2(\mathbf{C})$. The twist by $(z\overline{z})^{-d/2}$ ($d = \dim S_K$) of the $\mathbf{C}^*$-action defined by

$$\sigma_\phi\left(z, \begin{pmatrix} |z|^{1/2} & 0 \\ 0 & |z|^{-1/2} \end{pmatrix}\right)$$

yields the Hodge type on $H_\phi(\pi_f)$.

The reflex field E has a distinguished Archimedean place v corresponding to the fixed embedding into $\mathbf{C}$. The group L_{E_v} is either $W_{\mathbf{C}} = \mathbf{C}^*$ or $W_{\mathbf{R}}$ according as v is complex or real. By restricting $r \circ \phi$, we obtain a representation ψ_v of $L_{E_v} \times \mathrm{SL}_2(\mathbf{C})$ on the space

$$\bigoplus_{\chi} m(\chi, \pi_f) V(\chi)$$

where χ ranges over the characters of $S(\phi_v)$. It follows from the multiplicity formula and [A1, §9] that σ_ϕ is isomorphic to the restriction of ψ_v to $\mathbf{C}^* \times \mathrm{SL}_2(\mathbf{C})$. If v is complex, this implies that Conjecture 5.2 holds at the Archimedean place v. However, if v is real, one would have to calculate in addition the action of the Frobenius at v.

Let p be a prime such that G_p is unramified, K_p is hyperspecial. Then p is unramified in E. Let v be a place of E dividing p. Set $n_v = [E_v : \mathbf{Q}_p]$ and let $q = q_v$ be the cardinality of the residue field at v. To verify Conjecture 5.2 at v using the trace formula and Lefschetz fixed-point formula, one needs to express the number of points in $S_K(\mathbf{F}_{q^r})$ as a sum of orbital integrals for all $r \geq 1$. The twisted orbital integrals

$$\int_{G_{\delta\sigma}\backslash G(F)} \phi_n(g^{-1}\delta\sigma(g))\, dg$$

enter into the expression for the number of points. Here F is the unramified extension of $\mathbf{Q}_p$ of degree $n = [\mathbf{F}_{q^r} : \mathbf{F}_p]$ and σ is the absolute Frobenius. Let K_F be a hyperspecial maximal compact subgroup of $G(F)$, and let S be a maximal split F-torus such that the Cartan decomposition $K_F S(F) K_F$ holds. We may assume that μ takes values in S. Then ϕ_n is the characteristic function of the double coset $K_F \mu(p) K_F$. Roughly speaking, the p-component of the function to which one applies the trace formula is the image of ϕ_n under the base change homomorphism. This is a function f_n in the Hecke algebra of G_p. Kottwitz has shown that the Satake transform of f_n satisfies

$$\widehat{f}_n(\pi) = \mathrm{Trace}(\pi(f_n)) = p^{nd/2}\,\mathrm{Trace}(r(g(\pi_p)^n))$$

for any unramified representation π, where $d = \dim(S_K)$ **[K4]**. This is the coefficient of p^{-ns}/n in $-\log(L(s - \frac{d}{2}, \pi, r))$. We refer to **[K3]** for a full discussion.

6. The Hecke polynomial

We retain the notation of 5.4. Define

$$H_v(T) = \det(T - q_v^{d/2} r(g \times \Phi_p)^{n_v}).$$

This is a polynomial in T whose coefficients are $\mathrm{ad}(\widehat{G})$-invariant polynomials on the set $\widehat{G} \times \Phi_p$. As in §1.13, we view $H_v(T)$ as a polynomial with coefficients in the Hecke algebra $\mathscr{H}_{\mathbf{C}}$. Hence the coefficients of $H_v(T)$ are functions on $\Pi_{\mathrm{un}}(G)$. For any $\pi_p \in \Pi_{\mathrm{un}}(G)$, let $H_v(T, \pi_p)$ be the specialization of $H_v(T)$ at π_p.

CONJECTURE: THE CONGRUENCE RELATION. *Assume that G_p is unramified. Let $\pi_f \in \mathrm{Coh}_f$ and assume that π_p is unramified. Suppose that λ is prime to p. Then $H_v(\Phi_p, \pi_p)$ acts nilpotently on $H^*_\lambda(\pi_f)$. In particular, the eigenvalues of Φ_p acting on $H^*_\lambda(\pi_f)$ form a subset of the set of eigenvalues of $q_v^{d/2} r(g(\pi_p) \times \Phi_p)^{n_v}$.*

6.2. EXAMPLES.

(a) Let $G = \mathrm{GL}(2)_{\mathbf{Q}}$. This is the setting of the original work of Eichler and Shimura for modular curves. In this case, r is the standard two-dimensional representation of $\mathrm{GL}_2(\mathbf{C})$ and is trivial on $W_{\mathbf{Q}}$. Suppose that $K_p = \mathrm{GL}_2(\mathbf{Z}_p)$ and that π_f is unramified at p. If

$$g(\pi_p) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix},$$

then the Eichler-Shimura relation asserts that eigenvalues of Φ_p acting on $H^1_\ell(\pi_f)$ are contained in the set $\{p^{1/2}\alpha, p^{1/2}\beta\}$ for $\ell \neq p$.

(b) For symplectic groups, the congruence relation has been studied by Shimura in the case of GSp(4) **[S3]** and in general by Deligne **[D]** and Faltings-Chai **[FC]**.

Let us assume the congruence relation in the following case. Let E/F be a CM extension and let v_0 be the Archimedean place corresponding to the identity embedding of F in $\overline{\mathbf{Q}}$. Let G be the group of unitary similitudes of a Hermitian form $\Phi \in \mathrm{GL}_n(E)$ such that G_{v_0} has semisimple rank one and G_v is compact modulo the center for the remaining Archimedean places. Let S_K be an associated Shimura variety. The dimension of S_K is $d = n - 1$.

PROPOSITION 6.1. *Let ρ_j be the action of Γ_F on $H^j(S_K, \mathbf{Q}_\ell)$. Then ρ_j^{ss} is abelian for $j \leq [d/2]$.*

SKETCH. It suffices to check the claims for $H^j_\lambda(\pi_f)$. In this case, the reflex field is E and $H_w(T, \pi_p)$ has degree n. Suppose that $j \leq d$. Let $q^{j/2}\alpha$ be an eigenvalue of Frobenius occurring in $H^j_\lambda(\pi_f)$. By Deligne's theorem, $|\alpha| = 1$. Then $q^{j/2+k}\alpha$ is an eigenvalue of Frobenius occurring in $H^{j+2k}_\lambda(\pi_f)$ for $0 \leq k \leq d - j$ since the cup product with the polarization commutes with the action of the Hecke operators. If another root $q^{j/2}\beta$ occurs, then $H_w(T)$ will have $2(d - j + 1)$ distinct roots. The degree of $H_w(T)$ is n; hence,

$$2(d - j + 1) = 2(n - j) \leq n.$$

This shows that for almost all w, Φ_w has only one eigenvalue in $H^j_\lambda(\pi_f)$; hence, its semisimplification is scalar. Each ρ_j^{ss} is Hodge-Tate and hence is associated to an algebraic Hecke character χ. □

We recover the following result of Kumar Murty and D. Ramakrishnan **[KM]**.

COROLLARY 6.2. *The albanese variety of S_K is of CM type.*

PROOF. For $j = 1$, ρ_1 is semisimple and abelian, and hence the commutant of the $\rho_1(\Gamma_E)$ contains a commutative semisimple subalgebra of $\mathrm{End}(H^1(S_K, \mathbf{Q}_\ell))$ of maximal rank. Faltings's theorem implies the same is true of $\mathrm{End}_{\mathbf{Q}}(\mathrm{Alb})$. □

PROPOSITION 6.3. *Let $\pi_f \in \mathrm{Coh}_f$ and suppose that $\mathrm{Inf}(\pi_f)$ contains a nontempered representation π_∞. Then π_f is nontempered at almost all finite places.*

PROOF. If π_v is unramified, then π_v is tempered if and only if the eigenvalues of $r(g(\pi_v))$ all have the same absolute value. The representation π_∞ has cohomology in degree j for some $j < \dim(S_K)$. If α is an eigenvalue of Φ_v on $H^j_\ell(\pi_f)$, then $q_v\alpha$ occurs in $H^{j+2}_\ell(\pi_f)$. The congruence relation implies that both α and $q_v\alpha$ are eigenvalues of $r(g(\pi_v))$. □

For applications of the congruence relation to the Tate conjectures, see **[BS]**.

7. Examples

In this section we discuss Conjecture 5.2 in some special cases.

7.1. Modular curves. This is the case $G = \mathrm{GL}(2)_{\mathbf{Q}}$. The set Coh_∞ consists of the trivial representation, the character $\mathrm{sgn} \circ \det$, and the lowest-weight discrete series representation π^1 of $G(\mathbf{R})$. Suppose that $\mathrm{Inf}(\pi_f) = \pi^1$. Then $H^1_\lambda(\pi_f)$ is two dimensional by the multiplicity one theorem for GL(2) **[JL]**. Conjecture 5.2 asserts that

$$L_p(s, \rho^1(\pi_f)) = L(s - \tfrac{1}{2}, \pi_p) = (1 - \alpha_p p^{1/2-s})^{-1}(1 - \beta_p p^{1/2-s})^{-1},$$

for almost all p, where α_p, β_p are the eigenvalues of $g(\pi_p)$. It follows from the congruence relation, once one knows that $\det(\rho^1(\pi_f)(\Phi_p)) = p\alpha_p\beta_p$. This almost everywhere result has been improved in a series of important works to a precise result valid at all primes p: **[D3, L2, Car]**.

7.2. Shimura varieties associated to GL(2). Let F be a totally real number field of degree $d > 1$ and let D be a quaternion algebra central over F. Let $\underline{D}^*$ be the F-algebraic group defined by the multiplicative group of D and let $G = \mathrm{Res}_{F/\mathbf{Q}}(\underline{D}^*)$. Let S_∞ be the set of Archimedean places of F and let S be the subset of $v \in S_\infty$ such that D_v is split. Since F has a given embedding into $\overline{\mathbf{Q}}$, we may identify S_∞ with $\Gamma_F\backslash\Gamma_{\mathbf{Q}}$ and view S as a subset of $\Gamma_F\backslash\Gamma_{\mathbf{Q}}$. Then $G(\mathbf{R})$ is isomorphic to $(\mathrm{GL}_2(\mathbf{R}))^s \times (\mathbf{H}^*)^{d-s}$, where $s = |S|$ and $\mathbf{H}$ is the algebra of Hamilton quaternions. The Shimura variety is defined relative to the map h, which is given on real points by

$$a + bi \to \left(\underbrace{\begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \dots, \begin{pmatrix} a & b \\ -b & a \end{pmatrix}}_{v \in S}, \underbrace{1, \dots, 1}_{v \in S_\infty - S} \right).$$

The theory of the zeta function in this case has been worked out in [L] for nonsplit quaternion algebras and in **[BL, HLR]** for the split case.

The dual group of G is

$$\widehat{G} = \prod_{\Gamma_F\backslash\Gamma_{\mathbf{Q}}} \mathrm{GL}_2(\mathbf{C}),$$

and the L-group is the semidirect product ${}^LG = \widehat{G} \rtimes W_{\mathbf{Q}}$ where $W_{\mathbf{Q}}$ acts by translation of the factors through its projection onto $\Gamma_{\mathbf{Q}}$. Let $\widehat{T}$ be the product of the diagonal subgroups in $\widehat{G}$. Up to conjugacy by the Weyl group, the weight $\mu : \widehat{T} \to \mathbf{G}_m$ looks like

$$\mu\left(\begin{pmatrix} a_{v_1} & 0 \\ 0 & b_{v_1} \end{pmatrix}, \dots, \begin{pmatrix} a_{v_d} & 0 \\ 0 & b_{v_d} \end{pmatrix}\right) = \prod_{v \in S} a_v.$$

The stabilizer of μ in $\Gamma_{\mathbf{Q}}$ coincides with the stabilizer of the subset S. The reflex field E is the fixed field of this stabilizer.

Let r_0 be the standard representation of $\mathrm{GL}_2(\mathbf{C})$. Then r is the representation of ${}^LG_E$ on the s-fold tensor product $r_0 \otimes \cdots \otimes r_0$ defined as follows. For $g = (g_{v_1}, \dots, g_{v_d}) \in \widehat{G}$,

$$r(g) = r_o(g_{v_1}) \otimes \cdots \otimes r_o(g_{v_s})$$

where $S = \{v_1, \dots, v_s\}$. This is extended to ${}^LG_E$ by letting $W_{\mathbf{E}}$ act via permutation of the factors. We have $\dim(r) = 2^s$.

Let triv_v be the trivial character of G_v. For $v \in S$, let $\mathrm{sgn} \circ \det_v$ be the sign character and let π_v^1 be the lowest-weight discrete series representation of $G_v = \mathrm{GL}_2(F_v)$. A unitary representation $\pi_\infty = \bigotimes_{v \in S_\infty} \pi_v$ belongs to Coh_∞ if and only if $\pi_v \in \{\mathrm{triv}_v, \mathrm{sgn} \circ \det_v, \pi_v^1\}$ for $v \in S$ and $\pi_v = \mathrm{triv}_v$ for $v \in S_\infty - S$. The one-dimensional representations in Coh_f are the finite parts of finite-order automorphic characters of G. If π_f is infinite dimensional, then $\mathrm{Inf}(\pi_f)$ consists of a unique element of the form

$$\pi_\infty^1 = \underbrace{\pi_{v_1}^1 \otimes \cdots \otimes \pi_{v_s}^1}_{v \in S} \otimes \underbrace{\mathrm{triv}_{v_{s+1}} \otimes \cdots \otimes \mathrm{triv}_{v_d}}_{v \in S_\infty - S}$$

and can belong to $\mathrm{Inf}(\pi_f)$ for $\pi_f \in \mathrm{Coh}_f$. This follows from the well-known fact that an automorphic representation of G with a one-dimensional component at any noncompact place is one dimensional.

Let $\pi_f \in \mathrm{Coh}_f$ and assume that $\mathrm{Inf}(\pi_f) = \{\pi_\infty^1\}$. By the Künneth formula, $H^j(\mathfrak{G}, K_\infty, \pi_\infty^1)$ vanishes for $j \neq s$, and $\dim(H^s(\mathfrak{G}, K_\infty, \pi_\infty)) = 2^s$. Since cuspidal representations of G occur with multiplicity one, the representation of Γ_E on $H_\lambda^s(\pi_f)$ has dimension 2^s. We fix the place λ and denote this representation by $\rho^s(\pi_f)$.

Let p be a rational prime that is unramified in E. The L-group of G over $\mathbf{Q}_p$ is

$$G_p = \left(\prod_{v \in \mathrm{Hom}(F, \overline{\mathbf{Q}}_p)} GL_2(\mathbf{C})\right) \rtimes W_{\mathbf{Q}_p}.$$

Fix an embedding $\iota : \overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_p$ inducing a prime v of E. Then we obtain an embedding $W_{\mathbf{Q}_p} \hookrightarrow W_{\mathbf{Q}}$ and an identification of $\mathrm{Hom}(F, \overline{\mathbf{Q}}_p)$ with $\Gamma_F \backslash \Gamma_{\mathbf{Q}}$.

This yields an embedding ${}^LG_p \hookrightarrow {}^LG$. Suppose that π_p is unramified and that p does not divide λ. Let $g(\pi_p)\times\Phi_p \in {}^LG_p$ be a representative for the associated Langlands class. Let $e_v = [E_v : \mathbf{Q}_p]$. Then $(g(\pi_p)\times\Phi_p)^{e_v}$ belongs to ${}^LG_E$ and $r(g(\pi_p)\times\Phi_p)^{e_v})$ is defined. The local factor at v is

$$L_v(s, \rho^s(\pi_f)) = \det(1 - Nv^{-s} r(g(\pi_p)\times\Phi_p)^{e_v}).$$

Two-dimensional λ-adic representations. We recall that there exists a representation, first constructed by R. Taylor,

$$\rho_2(\pi_f) : \Gamma_F \longrightarrow \mathrm{GL}_2(V_\lambda)$$

with $\dim(V_\lambda) = 2$ such that $L(s, \rho_2(\pi_f)) = L(s, \pi)$ **[Ta1, BR1, BR2]**. The semisimplification of $\rho^s(\pi_f)$ can be described in terms of $\rho_2(\pi_f)$ using the following procedure.

Let G be any group and let $H, K \subset G$ be subgroups. Starting from a representation

$$\tau : H \longrightarrow \mathrm{GL}(W)$$

and a double coset $H\sigma K$ such that $d(\sigma) = |H\backslash H\sigma K| < \infty$, we define a representation $\tau_{H\sigma K}$ of K on the $d(\sigma)$-fold tensor product $W^{\otimes d(\sigma)}$. Choose representatives $\{\sigma_1, \dots, \sigma_{d(\sigma)}\}$ such that $H\sigma K = \bigcup H\sigma_j$. For $\gamma \in K$, there exists $\varepsilon_j \in H$ and an index $\gamma(j)$ such that $\sigma_j\gamma = \varepsilon_j \sigma_{\gamma(j)}$. Define

$$\tau_{H\sigma K}(\gamma)(w_1 \otimes \cdots \otimes w_{d(\sigma)}) = \tau(\varepsilon_1) w_{\gamma^{-1}(1)} \otimes \cdots \otimes \tau(\varepsilon_{d(\sigma)}) w_{\gamma^{-1}(d(\sigma))}.$$

It is easy to check that the equivalence class of $\tau_{H\sigma K}$ is independent of the choice of representatives $\{\sigma_j\}$.

Now write S as a disjoint union of double cosets:

$$S = \bigcup_{j=1}^{k} \Gamma_F \sigma_j \Gamma_E,$$

and let ρ_j be the representation of Γ_E defined by $\rho_2(\pi_f)$ and the double coset $\Gamma_F\sigma_j\Gamma_E$. Then $\rho^s_\lambda(\pi_f)^{ss}$ is isomorphic to the tensor product $\rho_1 \otimes \cdots \otimes \rho_k$

As a special case, suppose that $D = M_2(F)$ is the split quaternion algebra and that F is a normal extension of $\mathbf{Q}$. Then $s = d$ and $E = \mathbf{Q}$. The representation ρ^d occurs as a subrepresentation of

$$\mathrm{Ind}_F^{\mathbf{Q}}(\rho_2(\pi_f)^{\sigma_1} \otimes \cdots \otimes \rho_2(\pi_f)^{\sigma_d}),$$

where the σ_j are representatives for $\Gamma_F\backslash\Gamma_{\mathbf{Q}}$ and

$$\rho_2(\pi_f)^{\sigma}(\gamma) = \rho_2(\pi_f)(\sigma\gamma\sigma^{-1}).$$

This follows from the formula

$$\mathrm{Ind}_H^G(\mathrm{Res}_H(\rho)) \xrightarrow{\sim} \rho \otimes \mathrm{Ind}_H^G(1),$$

valid whenever ρ is a representation of a group G and H is a subgroup of finite index. In particular, the L-function is invariant under induction, so we see that $L(s, \rho^d(\pi_f))$ is a factor of

$$L(s, \rho_2(\pi_f)^{\sigma_1} \otimes \cdots \otimes \rho_2(\pi_f)^{\sigma_d}).$$

Specializing to the case $d = 2$, we have

$$L(s, \rho_2(\pi_f) \otimes \rho_2(\pi_f)^{\sigma}) = L(s, \rho^2(\pi_f))L(s, \rho^2(\pi_f) \otimes \omega_{F/\mathbf{Q}}).$$

7.3. Shimura varieties attached to $\mathrm{GSp}(4)_{\mathbf{Q}}$. Let $G = \mathrm{GSp}(4)_{\mathbf{Q}}$. Then $\widehat{G} = \mathrm{GSp}_4(\mathbf{C})$ with the trivial Galois action. Conjecture 5.2 implies that the L-functions of the irreducible motives that occur in the cohomology of the Shimura variety attached to $\mathrm{GSp}(4)_{\mathbf{Q}}$ are of the form $L(s, \pi)$ where π is a cuspidal representation of $\mathrm{GL}(m)$ for $m = 1, 2,$ or 4.

Parameters for G will be considered as homomorphisms into $\widehat{G}$. We first describe the cohomological parameters explicitly. We shall denote by ρ_m the standard representation of $\mathrm{SL}_2(\mathbf{C})$ of dimension m.

Archimedean parameters. Let T be an elliptic maximal torus in G and let $\phi_P : W_{\mathbf{Q}} \times \mathrm{SL}_2(\mathbf{C}) \to \mathrm{GSp}(4)$ for P a parabolic subgroup containing T.

(i) ϕ_B is the discrete series parameter whose restriction to $\mathbf{C}^*$ is

$$\phi_B(z) = \begin{pmatrix} (z/\overline{z})^{\frac{3}{2}} & 0 & 0 & 0 \\ 0 & (z/\overline{z})^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & (z/\overline{z})^{\frac{-1}{2}} & 0 \\ 0 & 0 & 0 & (z/\overline{z})^{\frac{-3}{2}} \end{pmatrix}.$$

The L-packet Π_B contains two representations $\{\pi^w, \pi^h\}$, where π^w contributes cohomology of Hodge type $(2, 1)$, $(1, 2)$ and π^h has cohomology of Hodge type $(3, 0)$, $(0, 3)$.

To describe the remaining cohomological parameters, set

$$P_1 = \left\{ \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix} \right\}, \qquad P_2 = \left\{ \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & 0 & 0 & * \end{pmatrix} \right\}.$$

The Levi factors of P_1 and P_2 are both isomorphic to $\mathrm{GL}_2(\mathbf{C}) \times \mathrm{GL}_1(\mathbf{C})$. Under the identification of $\mathrm{GSp}(4)$ with its dual group, the two classes of maximal parabolic subgroups are interchanged: $\widehat{P}_1 = P_2$ and $\widehat{P}_2 = P_1$. Let ρ_d denote the d-dimensional irreducible representation of $\mathrm{SL}_2(\mathbf{C})$.

(ii) If $P = P_1$, then ϕ_{P_1} maps $\mathrm{SL}_2(\mathbf{C})$ nontrivially to $\widehat{M_2}$. Let sgn be the unique character of $W_{\mathbb{R}}$ of order two. As a four-dimensional representation,

ϕ_P is isomorphic to $\phi_3 \oplus \rho_2$ or $\phi_3 \oplus \rho_2 \otimes \mathrm{sgn}$, where ϕ_3 is the irreducible two-dimensional representation of $W_{\mathbf{R}}$ whose restriction to $\mathbf{C}^*$ is

$$\phi_3(z) = \begin{pmatrix} (z/\overline{z})^{\frac{3}{2}} & 0 \\ 0 & (z/\overline{z})^{-\frac{3}{2}}. \end{pmatrix}$$

On $\mathbf{C}^*$, the associated Langlands parameter is

$$\phi'_P(z) = \begin{pmatrix} (z/\overline{z})^{\frac{3}{2}} & 0 & 0 & 0 \\ 0 & (z\overline{z})^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & (z\overline{z})^{-\frac{1}{2}} & 0 \\ 0 & 0 & 0 & (z/\overline{z})^{-\frac{3}{2}} \end{pmatrix}.$$

The two possible A-packets have the form $\Pi_{P_1} = \{\pi^{P_1}, \pi^h\}$, where π^{P_1} is a representation whose cohomology has Hodge types $(1, 1)$ and $(2, 2)$.

(iii) If $P = P_2$, then ϕ_{P_2} maps $\mathrm{SL}_2(\mathbf{C})$ nontrivially into the Levi factor of $\widehat{P_2} = P_1$ and, as a four-dimensional representation, ϕ_{P_2} is isomorphic to $\phi_2 \otimes \rho_2$ where ϕ_2 is the irreducible two-dimensional representation of $W_{\mathbf{R}}$ whose restriction to $\mathbf{C}^*$ is

$$\phi_2(z) = \begin{pmatrix} z/\overline{z} & 0 \\ 0 & \overline{z}/z \end{pmatrix}.$$

On $\mathbf{C}^*$, the associated Langlands parameter is

$$\phi'_{P_2}(z) = \begin{pmatrix} z^{\frac{3}{2}}\overline{z}^{-\frac{1}{2}} & 0 & 0 & 0 \\ 0 & z^{\frac{1}{2}}\overline{z}^{-\frac{3}{2}} & 0 & 0 \\ 0 & 0 & z^{-\frac{1}{2}}\overline{z}^{\frac{3}{2}} & 0 \\ 0 & 0 & 0 & z^{-\frac{3}{2}}\overline{z}^{\frac{1}{2}} \end{pmatrix}.$$

The L-packet Π_{P_2} contains a single representation π^{P_2} which has cohomology of Hodge types $(3, 1), (2, 0), (0, 2), (1, 3)$.

(iv) If $P = G$, then ϕ_G is trivial on $W_{\mathbf{R}}$ and its restriction to $\mathrm{SL}_2(\mathbf{C})$ is the four-dimensional irreducible representation ρ_4, which is symplectic.

We now return to the global case. The representation r is the evident representation of dimension four. Let $\phi : \mathscr{L}_{\mathbf{Q}} \times \mathrm{SL}_2(\mathbf{C}) \to \mathrm{GSp}_4(\mathbf{C})$ be a discrete A-parameter. Recall that ϕ is called stable if $\mathfrak{S}(\phi) = \{1\}$, i.e., $S(\phi) = Z(\widehat{G})$. This is equivalent to the irreducibility of $r \circ \phi$. If $r \circ \phi$ is reducible, then it decomposes as $\sigma_1 \oplus \sigma_2$ where $\dim(\sigma_j) = 2$. Since $S(\phi)$ is finite modulo $Z(\widehat{G})$, σ_1 is not equivalent to a twist of σ_2. If this were the case, ϕ would factor through a parabolic subgroup. Hence ϕ factors through the group

$$H = \{(g, h) \in \mathrm{GL}(2) \times \mathrm{GL}(2) : \det(g) = \det(h)\}$$

viewed as a subgroup of $\mathrm{GSp}(4)$ in the natural way. We call ϕ **endoscopic** if $r \circ \phi$ factors through H.

Let $\Pi = \Pi(\phi)$. We suppose that Π_∞ is cohomological and analyze the conjectural possibilities.

Case 1. ϕ stable.

(i) ϕ tempered. Then $\Pi_\infty = \Pi_B$. In this case, $L(s, \phi, r)$ would be the L-function of a cuspidal representation π of $\mathrm{GL}(4)$. Conjecturally, the set of π obtained in this way from stable tempered parameters for $\mathrm{GSp}(4)$ are precisely those such that $L(s, \Lambda^2(\pi))$ has a pole at $s = 1$.

(ii) $\phi = \phi_2 \otimes \rho_2$ where ϕ_2 is an irreducible two-dimensional representation of $\mathscr{L}_{\mathbf{Q}}$. Then $\Pi_\infty = \Pi_{P_2}$. We observe that ϕ_2 must be induced from a character of $\mathscr{L}_K$ for some quadratic extension K of $\mathbf{Q}$. This is because $\phi \otimes \rho_2$, which decomposes as $\phi_2 \oplus \phi_2 \otimes \mathrm{Sym}^2(\rho_2)$, carries an orthogonal similitude. One would therefore expect that ϕ_2 is associated to a holomorphic cuspidal representation of $\mathrm{GL}(2)_{\mathbf{Q}}$ of classical weight 3 of CM type, i.e., with a quadratic self-twist. This can in fact be verified (cf. Proposition 7.1).

(iii) $\phi = \psi \otimes \rho_4$, where ψ is a character of $\mathscr{L}_{\mathbf{Q}}$. Then $\Pi_\infty = \Pi_G$ and Π consists of a single one-dimensional automorphic character of $\mathrm{GSp}(4)$ with trivial infinity type.

Case 2. ϕ endoscopic. Then ϕ factors as ϕ as $\phi_1 \times \phi_2$, where the ϕ_j are two-dimensional representations with the same central character such that ϕ_1 is not equivalent to a twist of ϕ_2.

(i) ϕ tempered. Then $\Pi_\infty = \Pi_B$ and ϕ_1 and ϕ_2 correspond to holomorphic cuspidal representations π_1 and π_2 of $\mathrm{GL}(2)_{\mathbf{Q}}$ of classical weights 2 and 4 having the same central character.

(ii) ϕ nontempered. Then $\Pi_\infty = \Pi_{P_1}$. In this case, ϕ_1 corresponds to a holomorphic cuspidal representation of $\mathrm{GL}(2)_{\mathbf{Q}}$ of classical weight four and ϕ_2 is a two-dimensional parameter with nontrivial restriction to $\mathrm{SL}_2(\mathbf{C})$ and whose restriction to $\mathscr{L}_{\mathbf{Q}}$ is a character.

We now prove that if $\pi = \pi_\infty \otimes \pi_f$ with $\pi_\infty = \pi_{P_2}$, the description of the π_f-isotypic component $IH^*(\pi_f) = IH^*(\overline{S}_K, \mathbf{Q}_\ell)(\pi_f)$ has the structure predicted by Case (1)(ii).

PROPOSITION 7.1. *Let π be a discrete cohomological representation of G such that $\pi_\infty = \pi_{P_2}$.*

(a) *Then there exists a stable A-parameter*

$$\psi : W_{\mathbf{Q}} \times \mathrm{SL}_2(\mathbf{C}) \to \mathrm{GSp}_4(\mathbf{C})$$

such that $\pi_v \in \Pi(\psi_v)$ for almost all places v such that π_v and ψ_v are unramified. More precisely, there exists a quadratic imaginary extension $K/\mathbf{Q}$ and a unitary Hecke character χ of K of infinity type $z/\overline{z}$ such that $\psi = \mathrm{Ind}_K^{\mathbf{Q}}(\chi) \otimes \rho_2$. For almost all places p, $L_p(s, \pi, r) = L_p(s-\frac{1}{2}, \theta)L_p(s+\frac{1}{2}, \theta)$ where θ is the unitary holomorphic cuspidal representation of CM type on $\mathrm{GL}(2)_{\mathbf{Q}}$ of classical weight three associated to χ.

(b) *There exists a positive integer* m *such that for almost all* p,

$$L_p(s, IH^*(\pi_f)^{ss}) = L_p(s - \tfrac{3}{2}, \pi, r)^m.$$

Proof. We sketch an argument, the essentials of which were first found by R. Weissauer [W] in his study of the Tate conjecture for H^2 of the Siegel modular threefolds.

Since $\pi_{P\infty}$ has Hodge types $(2, 0)$, $(0, 2)$, there exists a holomorphic 2-form ω on S_K in $H^2(\pi_f)$. Passing to a suitable finite cover, we can find an embedding of a product of two modular curves $C_1 \times C_2$ into S_K such that the pull-back of ω and its complex conjugate is nonzero. Hence, restriction yields a nonzero Galois equivariant map from $H^2_\lambda(\pi_f)$ to $H^2_\lambda(C_1 \times C_2)$ whose image W_λ has an irreducible constituent σ of dimension at least two. To show this, note that if the Galois action on W_λ^{ss} is abelian, then it is given by Hecke characters of $\mathbf{Q}$, necessarily of the form $\varepsilon|\cdot|^{-1}$ where ε has finite order. By Hodge-Tate theory, the associated Hodge structure would be purely of type $(1, 1)$, which is not possible since ω has type $(2, 0)$. We now show that $\dim(\sigma) = 2$. By the congruence relation and an argument similar to that used in Proposition 6.1, a Frobenius element Φ_p has at most two distinct eigenvalues on $H^2(\pi_f)$. We have seen that some Φ_p has at least two distinct eigenvalues since σ is not abelian. Now σ is a constituent of $H^2_\lambda(C_1 \times C_2)$ and hence must occur in

$$H^1_\lambda(C_1) \otimes H^1_\lambda(C_2).$$

The irreducible constituents of this space are factors of tensor products $\rho_1 \otimes \rho_2$, where the ρ_j are two-dimensional λ-adic representations associated to modular forms of weight 2 of $\mathrm{GL}(2)_{\mathbf{Q}}$. If $\dim(\sigma) > 2$, then either $\sigma = \rho_1 \otimes \rho_2$ or $\sigma = \mathrm{Sym}^2(\rho_1) \otimes \psi$ for some character ψ. In either case, infinitely many Frobenii have more than two distinct eigenvalues. Hence $\dim(\sigma) = 2$, and one can conclude that ρ_1 and ρ_2 are dihedral. Indeed, if $\rho_1 \otimes \rho_2 = \sigma \oplus \sigma'$, then the left-hand side carries a nondegenerate orthogonal similitude. If this form restricts nondegenerately to σ, we are done. Otherwise, it must define a pairing between σ and σ', hence, $\sigma' \xrightarrow{\sim} \sigma \otimes \psi$ for some character ψ. Taking exterior squares of both sides, we find that

$$\Lambda^2(\rho_1 \otimes \rho_2) = \Lambda^2(\rho_1) \otimes \mathrm{Sym}^2(\rho_2) \oplus \Lambda^2(\rho_2) \otimes \mathrm{Sym}^2(\rho_1) = \Lambda^2(\sigma \oplus \sigma \otimes \psi).$$

The right-hand side contains at least three Galois-invariant lines. Hence one of ρ_1 and ρ_2 is dihedral, say $\rho_1 = \mathrm{Ind}_K^{\mathbf{Q}}(\alpha)$ where K is a quadratic extension of $\mathbf{Q}$. Then $\rho_1 \otimes \rho_2$ has a self-twist by a character φ. Hence $\sigma \otimes \varphi$ is isomorphic to σ or $\sigma \otimes \psi$. In either case, σ itself has a self-twist, and this shows that it is dihedral. By Hodge-Tate theory, σ is induced from a Hecke character χ' of an imaginary quadratic extension of $\mathbf{Q}$ of infinity type z^{-2}. The proposition holds with $\chi = \chi'|\cdot|^2$ and $\psi = \mathrm{Ind}_K^{\mathbf{Q}}(\chi) \otimes \rho_2$.

REMARK. R. Taylor has shown that under certain hypotheses it is possible to attach to cohomological π on GSp(4) a four-dimensional system of λ-adic representations of $\Gamma_{\mathbf{Q}}$ [Ta2]. He proves that on a set of positive density the characteristic polynomials of Frobenius are what they should be.

7.4. Unitary groups. Let F be an extension of degree d of a Hermitian form Φ in n-variables and let G be the group of unitary similitudes of Φ. We assume that E/F is a CM extension: F totally real and E totally imaginary, since this is the only case in which there are associated Shimura varieties. Let $G' = \mathrm{Res}_{F/\mathbf{Q}}(G)$.

Let S_∞ be the set of Archimedean places of F. For $v \in S_\infty$ of F, choose an embedding $\iota_v : E \to \mathbf{C}$ that induces v on F. Let (p_v, q_v) be the signature of $\iota_v(\Phi)$ where $p_v \leq q_v$. There exists an isomorphism of

$$G'(\mathbf{R}) = G_\infty \longrightarrow \prod_{v \in S_\infty} \mathrm{GU}(p_v, q_v).$$

Define $\mu : \mathbf{G}_m \to G_{\mathbf{C}}$ such that $\mu(t) = (\mu_v(t))$ and

$$\mu_v(t) = \begin{pmatrix} t1_p & \\ & 1_q \end{pmatrix}.$$

Let r_{st} be the n-dimensional representation of $\widehat{G} = \mathrm{GL}_n(\mathbf{C}) \times \mathrm{GL}_1(\mathbf{C})$ defined by $(g, \lambda) \to \lambda g$. Extend r_{st} to $\widehat{G} \times W_E$ with W_E acting trivially (this representation extends further to ${}^LG$ if and only if $n = 2$). We define the standard L-function of an automorphic representation π of G to be $L(s, \pi_E, r_{\mathrm{st}})$ where π_E is the base change of π to G_E. Let π_0 denote the restriction of π to the unitary group G_0 and let π_{0E} denote the base change of π_0 to $G_{0E} = \mathrm{GL}(n)_E$. Let χ_π be the central character of π. The center of G is isomorphic to E^*, so we view χ_π as a Hecke character of E. Then

$$L(s, \pi, r_{\mathrm{st}}) = L(s, \pi_{0E} \otimes \overline{\chi}_\pi)$$

where $\overline{\chi}(z) \equiv \chi(\overline{z})$ (the bar denotes conjugation of E/F) and the L-function on the right is a standard L-function for $\mathrm{GL}(n)$.

Suppose that $F = \mathbf{Q}$. If $p \neq q$, then $E = K$ and the L-functions which arise in the description of the zeta function are factors of $L(s, \pi_E, r^p)$ where r^p is the representation on the p^{th}-exterior power given by $r^p(g, \lambda) = \lambda\Lambda^p(g)$. If $p = q$, $E = \mathbf{Q}$ and r^p extends to a representation of ${}^LG$. In this case, the L-functions $L(s, \pi, r)$ occur. If $F \neq \mathbf{Q}$, the description of r is more involved, as in the Hilbert modular case.

We will restrict attention to the case that G_∞ has precisely one non-compact factor. If Π is an A-packet on G, set $L(s, \Pi) = L(s, \pi_{0E} \otimes \overline{\chi}_\pi)$ where π is any element of Π. Conjecturally, $L(s, \Pi)$ is the L-function of an automorphic representation of $\mathrm{GL}(n)_E$ which need not be cuspidal. If π

is a tempered L-packet,

$$L(s, \Pi) = \prod_{j}^{r} L(s, \pi_j).$$

Here π_j is a cuspidal representation of $\mathrm{GL}(n_j)_E$, and $n_1 + \cdots + n_r$ is a partition of n. In this case, the role of the zeta-function conjecture is to pick out a factor such that $L(s, \rho_\lambda^i(\pi_f)) = L(s, \pi_j)$.

7.5. Picard modular surfaces. Let G be a quasi-split unitary group in three variables such that G_∞ has precisely one noncompact factor and the associated Shimura varieties are of dimension two and are called Picard modular surfaces. The theory of the zeta functions at almost all places for these surfaces is worked out in [M]. We review the results in the case $F = \mathbf{Q}$. Let Π be a cohomological A-packet and let $\pi_f \in \Pi_f$. In this case, $S(\phi)$ can be defined, although ϕ of course is not known to exist. We have $S(\phi) = Z(\widehat{G})C(\phi)$ where $C(\phi)$ is the centralizer of the image of ϕ in $\mathrm{SL}_n(\mathbf{C}) \subset \widehat{G}$.

Tempered case. Suppose first that $\Pi_\infty = \Pi_B$ is a discrete series L-packet. Then $\Pi_\infty = \{\pi^{20}, \pi^{11}, \pi^{02}\}$ where the labeling is according to Hodge types. Up to equivalence, the restriction of the Archimedean parameter to $\mathbf{C}^*$ is

$$\phi_\infty : z \to \begin{pmatrix} z/\overline{z} & & \\ & 1 & \\ & & \overline{z}/z \end{pmatrix},$$

and $S(\phi_\infty) = Z(\widehat{G})C(\phi_\infty)$ where

$$C(\phi_\infty) = \left\{ s = \begin{pmatrix} \pm 1 & & \\ & \pm 1 & \\ & & \pm 1 \end{pmatrix} : \det(s) = 1 \right\}.$$

Let $\chi_\infty : C(\phi_\infty) \to \mathbf{C}^*$ be the character sending an element of $C(\phi_\infty)$ to its middle eigenvalue.

The global centralizer is a subgroup of $C(\phi_\infty)$ and hence is either trivial or isomorphic to $(\mathbf{Z}/2)^r$ for $r = 1$ or 2. In this case,

$$\chi(\pi_f) = \chi'_\infty \cdot \langle \bullet, \pi_f \rangle$$

where χ'_∞ is the restriction of χ_∞ to $C(\Pi)$.

If $C(\Pi)$ is trivial, then Π is a stable L-packet and $L(s, \rho_\lambda^2(\pi_f)) = L(s, \Pi, r_{\mathrm{st}})$, a cuspidal L-function on $\mathrm{GL}(3)$. Suppose that $C(\phi)$ has order 2. Then

$$L(s, \Pi) = L(s, \pi_1)L(s, \pi_2)$$

where π_j is a cuspidal representation of $\mathrm{GL}(j)_E$. Furthermore, π_1 and π_2 are the base changes of L-packets σ_j on the unitary groups $U(1)$ and $U(2)$, respectively. Let s_k denote the unique nontrivial element in $C(\phi_\infty)$ whose

k^{th} diagonal entry is $+1$. Then $C(\Pi) = \{1, s_k\}$ for some k. The derived group of $U(2)$ is $\mathrm{SL}(2)_{\mathbf{Q}}$ and the L-packet σ_2 corresponds in a natural way to a classical modular newform on $\mathrm{GL}(2)_{\mathbf{Q}}$ whose weight m depends on $C(\Pi)$ or, more precisely, on k. The weight m is 3 if $k = 2$ and $m = 2$ for $k = 1, 3$. Observe that χ'_∞ is trivial if $k = 2$ and is nontrivial otherwise. According to the recipe in §5,

$$L(s, \rho_\lambda^2(\pi_f)) = \begin{cases} L(s, \pi_1) & \text{if } \chi'_\infty(s_k)\langle s_k, \pi_f\rangle = 1 \\ L(s, \pi_2) & \text{if } \chi'_\infty(s_k)\langle s_k, \pi_f\rangle = -1. \end{cases}$$

On the other hand, if $C(\phi)$ has order 4, then

$$L(s, \Pi) = L(s, \pi_1)L(s, \pi_2)L(s, \pi_3)$$

where the π_1, π_2, π_3 are Hecke characters of E which may be numbered to have infinity types $z/\overline{z}, 1, \overline{z}/z$, respectively. In this case, the recipe of §5 can be stated as follows: $L(s, \rho_\lambda(\pi_f)) = L(s, \pi_k)$ where k is the unique index such that $\chi'(s_k)\langle s_k, \pi_f\rangle = 1$.

Nontempered case. If Π_∞ is nontempered, then $\Pi_\infty = \Pi_P$ where P is either the $(2, 1)$ or the $(1, 2)$ parabolic. The A-packet Π_P contains two elements π^1, π^d where π^1 is a non-tempered representation which contributes cohomology in degrees 1 and 3, and π^d is a discrete series representation. Globally, the A-packets of this type depend essentially on abelian data. In fact, Π is associated to a parameter ϕ of the following form. There exists a Hecke character φ of E, which we view as a character of W_E, and a Hecke character ρ' of the norm one group E^1 such that for $w \in W_E$,

$$\phi(w, 1) = \begin{pmatrix} \rho'_E(w)\varphi(w) & 0 & 0 \\ 0 & \rho'_E(w) & 0 \\ 0 & 0 & \rho'_E(w)\varphi(w) \end{pmatrix} \times w$$

where $\rho'_E(z) = \rho'(z/\overline{z})$,

$$\phi\left(1, \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ c & 0 & d \end{pmatrix},$$

and

$$\psi(w_\sigma) = \begin{pmatrix} & & 1 \\ & 1 & \\ -1 & & \end{pmatrix} \times w_\sigma.$$

The restriction of φ to F is the character of order 2 associated to E/F by class field theory. The group $S(\phi)$ is equal to $Z(\widehat{G})C(\phi)$ where $C(\phi)$ is the group of order two generated by

$$s = \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix}.$$

The character ε_ϕ of $C(\phi)$ appearing in the multiplicity formula is trivial if $\varepsilon(\frac{1}{2}, \phi) = +1$ and is nontrivial otherwise. For a given π_f, $\mathrm{Inf}(\pi_f) = \{\pi^n\}$ if $\varepsilon_\phi\langle \bullet, \pi_f\rangle$ is trivial and $\mathrm{Inf}(\pi_f) = \{\pi^d\}$ otherwise. In the former case, $L(s, \rho^1(\pi_f)) = L(s - \frac{1}{2}, \rho'_E\varphi)$.

REMARK. For Shimura varieties attached to unitary groups in more variables our knowledge is not nearly so complete. However, if G is a unitary group attached to a division algebra over E with an involution of the second kind, then the zeta functions of the associated Shimura variety have been computed in **[K2]**. In contrast to the general unitary groups, the problems due to endoscopy do not appear due to a Galois cohomological vanishing result. In particular, only the full L-functions $L(s, \pi, r)$ occur, and it is not necessary to factor them.

7.6. The abelian nature of H^1 can be shown to be the case more generally for unitary similitude groups G over a totally real field F such that G_∞ has precisely one noncompact factor isomorphic to $\mathrm{GU}(n-1, 1)$ with $n > 2$.

PROPOSITION 7.2. *Assume that G_∞ is isomorphic to $GU(n-1, 1) \times \mathrm{GU}(n)^{d-1}$ where $n \geq 3$. Let π be a discrete cohomological representation of G such that π_∞ contributes cohomology in degree* 1. *Then there exists an A-parameter*

$$\phi : W_F \times \mathrm{SL}_2(\mathbf{C}) \to {}^L G$$

such that $\pi_v \in \Pi(\phi_v)$ for almost all places v of F.

PROOF. As shown in Proposition 6.1, the representation of Γ_E on $H^1(S_K, \mathbf{Q}_\ell)$ is abelian. Let χ be the unitary Hecke character such that Γ_E acts by $\chi|\cdot|^{-1/2}$ on $H^1(\pi_f)$. By the Lefschetz decomposition, the character $\chi|\cdot|^{-j+1/2}$ occurs in $H^{2j-1}(\pi_f)$ for $j = 1, \dots, n-1$. For almost all places w of E, the matrix $r_{\mathrm{st}}(g_w(\pi_f))$ must have among its eigenvalues the numbers $\chi(\varpi_w)q_w^{j-\frac{n}{2}}$ for $1 \leq j \leq n-1$, where ϖ_w is a prime element at w. Let $\psi(\omega_w)$ denote the remaining eigenvalue of $r_{\mathrm{st}}(g_w(\pi_f))$. The base change π_{0E} of π_0 to $\mathrm{GL}(n)_E$ is not defined, but the w-component of π_{0E} is defined for all places w that divide a place v of F such that π_{0v} is unramified. For convenience, we write π_{0E} below, although its components are not everywhere defined. The determinant of $r_{\mathrm{st}}(g_w(\pi_f)) = g(\pi_{0E} \otimes \overline{\chi_\pi})$ is

$$\chi_{\pi_{0E}}(\varpi_w)\chi_\pi(\overline{\varpi_w})^n) = \chi_\pi(\varpi_w)\chi_\pi(\overline{\varpi_w})^{n-1}$$

since $\chi_{\pi_{0E}}(z) = \chi_\pi(z/\overline{z})$. Hence

$$(*) \qquad \chi(\varpi_v)^{n-1}\psi(\varpi_v) = \chi_\pi(\varpi_w)\chi_\pi(\overline{\varpi_w})^{n-1},$$

and it follows that ψ is a Hecke character. At each infinite place, the infinity type of χ is $(z/|z|)^{\pm 1}$ and since χ_π is of finite order, the infinity type of ψ is $(z/|z|)^{\pm(n-1)}$.

We have

$$L\left(s-\frac{n-1}{2}, \pi_{0E}\otimes\overline{\chi}_\pi\right) = L\left(s-\frac{n-1}{2}, \psi\right)\prod_{j=1}^{n-1} L\left(s-j+\frac{1}{2}, \chi\right)$$

and hence

$$L\left(s-\frac{n-1}{2}, \pi_{0E}\right) = L\left(s-\frac{n-1}{2}, \psi\overline{\chi}_\pi^{-1}\right)\prod_{j=1}^{n-1} L\left(s-j+\frac{1}{2}, \chi\overline{\chi}_\pi^{-1}\right).$$

We call a representation ρ of $\mathrm{GL}(n)_E$ ε-invariant if $\varepsilon(\rho)$ is equivalent to ρ, where $\varepsilon(\rho)(g) = \rho({}^t\overline{g}^{-1})$. The representation π_{0E} is ε-invariant (i.e., at the unramified places where it is defined) since π_{0E} is a base change from a unitary group. The L-function $L(s-\frac{n-1}{2}, \pi_{0E})$ has a corresponding invariance which implies that

$$(**) \qquad \overline{\chi}_\pi(z\overline{z}) = \chi(z\overline{z}) = \psi(z\overline{z}) = 1$$

for all $z \in I_E$. Define

$$\phi : W_E \times \mathrm{SL}_2(\mathbf{C}) \to \mathrm{GL}_n(\mathbf{C}) \times \mathrm{GL}_1(\mathbf{C})$$

as follows. For $z \in W_E$ and $g \in \mathrm{SL}_2(\mathbf{C})$, let

$$\phi(z, g) = \left(\overline{\chi}_\pi^{-1}\begin{pmatrix}\chi(z)\rho_{n-1}(g) & \\ & \psi(z)\end{pmatrix}, \overline{\chi}_\pi\right) \times z \in {}^L G.$$

where ρ_{n-1} is an irreducible $(n-1)$-dimensional representation of $\mathrm{SL}_2(\mathbf{C})$. The matrix Φ_{n-1} defined in §1.8(b) is orthogonal or symplectic according as n is odd or even. It follows that we may choose the realization so that $\Phi_{n-1}\,{}^t\rho_{n-1}(g)\Phi_{n-1}^{-1} = \rho_{n-1}(g)$.

It remains to extend ϕ to W_F. A necessary condition for the extension to exist is that the representations $\mathrm{ad}(w_\sigma)\phi(z, g)$ and $\phi(\overline{z}, g)$ of $I_E \times \mathrm{SL}_2(\mathbf{C})$ be equivalent, where $w_\sigma \in W_F$ projects to the nontrivial element $\sigma \in \mathrm{Gal}(E/F)$. This follows by $(*)$, $(**)$, and the choice of ρ_{n-1}. Let $Y \in \mathrm{GL}_n(\mathbf{C})$ be the element with blocks Φ_{n-1} and 1 and set

$$\psi(w_\sigma) = (Y\Phi_n^{-1}, 1) \rtimes w_\sigma.$$

Then $\mathrm{ad}(\psi(w_\sigma))(\psi(z, g)) = \psi(\overline{z}, g)$, and to conclude we must check that $\psi(w_\sigma)^2 = \psi(w_\sigma^2)$, i.e., the equality

$$\begin{pmatrix}-I_{n-1} & \\ 0 & (-1)^{n-1}\end{pmatrix} = \overline{\chi}_\pi^{-1}(w_\sigma^2)\begin{pmatrix}\chi(w_\sigma^2)I_{n-1} & \\ 0 & \psi(w_\sigma^2)\end{pmatrix}.$$

Since χ_π is of finite order and E is a CM field, the infinity type of χ_π is trivial. Relation $(*)$ implies that χ^{n-1} and ψ have the same infinity type. Let $w_\infty \in I_F$ be an element whose component at a single Archimedean place is -1 and whose remaining components are all $+1$. Then w_∞ generates $I_F/N_{E/F}(I_E)$ and

$$\overline{\chi}_\pi^{-1}\chi(w_\sigma^2) = \chi(w_\infty) = -1$$

since the infinity type of χ is $(z/|z|)^{\pm 1}$. Furthermore,

$$\overline{\chi}_\pi^{-1}\psi(w_\sigma^2) = \psi(w_\infty) = (-1)^{n-1}$$

by $(*)$ as required.

REFERENCES

[AF] *Automorphic forms, Shimura varieties, and L-functions.* vols. I and II (L. Clozel and J. Milne, eds.), Academic Press, New York, 1990.

[ABV] J. Adams, D. Barbasch, and D. Vogan, *The Langlands classification and irreducible characters of real reductive groups*, Birkhäuser, Boston, MA, 1992.

[AJ] J. Adams and J. Johnson, *Endoscopic groups and packets of non-tempered representations*, Compositio Math. **64** (1987), 271–309.

[A1] J. Arthur, *Unipotent automorphic representations: conjectures*, Orbites Unipotentes et Représentations. II, Astérisque 171/172 (1989) 13–71.

[A2] ——, *Unipotent automorphic representations: Global motivation*, Automorphic Forms, Shimura Varieties and L-Functions, vol. I (L. Clozel and J. Milne, eds.), Academic Press, New York, 1990, pp. 1–76.

[A3] ——, *On some problems suggested by the trace formula*, Lie Group Representations II, Lecture Notes in Math., vol. 1041, Springer-Verlag, New York, 1983, pp. 1–49.

[B] D. Blasius, *On multiplicities for $SL(n)$*, Israel J. Math. (to appear).

[BR1] D. Blasius and J. Rogawski, *Galois representations for Hilbert modular forms*, Bull. Amer. Math. Soc. (N. S.) **21** (1989), 65–69.

[BR2] ——, *Motives for Hilbert modular forms*, Invent. Math. (to appear).

[BS] D. Blasius and J. Schwermer, *On Tate's conjectures for H^2 of Shimura varieties*, in preparation.

[Bo] A. Borel, *Automorphic L-functions*, Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, Part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 27–61.

[BC] A. Borel and W. Casselman, *L^2-cohomology of locally symmetric manifolds of finite volume*, Duke Math. J. **50** (1983), 625–647.

[Bn] N. Bourbaki, *Groupes et algèbres de Lie*, Chapitres 7–8, Hermann, Paris, 1975.

[BL] J.-L. Brylinski and J.-P. Labesse, *Cohomologie d'intersection et fonctions L de certaines variétés de Shimura*, Ann. Sci. École Norm. Sup. **17** (1984), 361–412.

[BSL] M. Burger, P. Sarnak, and J.-S. Li, *Ramanujan duals and automorphic spectrum*, Bull. Amer. Math. Soc. (N. S.) **26** (1992), 253–257.

[C] *Automorphic forms, representations, and L-functions*, Proc. Sympos. Pure Math., vol. 33 (A. Borel and W. Casselman, eds.), Amer. Math. Soc., Providence, RI, 1979.

[Car] H. Carayol, *Sur les répresentations ℓ-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. École Norm. Sup. **19** (1986), 409–468.

[Ca] W. Casselman, *An assortment of results on representations of $GL_2(k)$*, Lecture Notes in Math., vol. 349, Springer-Verlag, New York, 1973.

[CF] C.-L. Chai and G. Faltings, *Degeneration of abelian varieties*, Springer-Verlag, New York, 1990.

[Cl] L. Clozel, *Représentations Galoisiennes associées aux représentations automorphes autoduales de $GL(n)$*, Inst. Hautes Études Sci. Publ. Math. **73** (1991), 13–145.

[D1] P. Deligne, *Letter to Serre,* 1968.

[D2] ——, *Travaux de Shimura*, Sém. Bourbaki 1970/71, Exposé 389, Lecture Notes in Math., vol. 244, Springer-Verlag, New York, 1971.

[D3] ——, *Formes modulaires et représentations ℓ-adique*, Sém. Bourbaki 1969, Exposé 55, Lecture Notes in Math., vol. 179, Springer-Verlag, New York, 1969, pp. 139–172.

[E] M. Eichler, *Quaternäre quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion*, Arch. Math. **5** (1954), 355–366.

[F] D. Flath, *Decomposition of representations into tensor products*, Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, Part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 179–183.

[G] S. Gelbart, *Automorphic forms on adèle groups*, Ann. of Math. Stud., vol. 83, Princeton Univ. Press, Princeton, NJ, 1975.

[JL] H. Jacquet and R. Langlands, *Automorphic forms on* $GL(2)$, Lecture Notes in Math., vol. 114, Springer-Verlag, New York, 1970.

[HLR] G. Harder, R. Langlands, and M. Rapoport, *Algebraische Zyklen auf Hilbert-Blumenthal Flächen*, J. Reine Angew. Math. **366** (1986), 53–120.

[HPS] R. Howe and I. I. Piatetski-Shapiro, *A counterexample to the generalized Ramanujan conjecture for (quasi-) split groups*, Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, Part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 315–322.

[I] Y. Ihara, *Hecke polynomials as congruence* ζ*-functions in elliptic modular case*, Ann. of Math. (2) **85** (1967), 267–295.

[K1] R. Kottwitz, *Stable trace formula: Cuspidal tempered terms*, Duke Math. J. **51** (1984), 611–650.

[K2] ——, *On the* λ*-adic representations associated to some simple Shimura varieties*, Invent. Math. **108** (1992), 653–665.

[K3] ——, *Shimura varieties and* λ*-adic representations*, Automorphic forms, Shimura varieties, and L-functions, vol. 1 (L. Clozel and J. Milne, eds.), Academic Press, New York, 1990.

[K4] ——, *Shimura varieties and twisted orbital integrals*, Math. Ann. **269** (1984), 287–300.

[K5] ——, *Points on some Shimura varieties over finite fields*, J. Amer. Math. Soc., vol. 5, No. 2 (1992), 373–444.

[KS] R. Kottwitz and D. Shelstad, *Twisted Endoscopy.* I: *Definitions, Norm mappings, and transfer factors*, preprint 1992.

[Ku] S. Kudla, *The local Langlands correspondence: The non-Archimedean case*, these Proceedings, vol. 2, pp. 365–391.

[KM] V. Kumar Murty and D. Ramakrishnan, *The Albanese of unitary Shimura varieties*, Zeta-functions of Picard modular surfaces (R. Langlands and D. Ramakrishnan, eds.), Univ. Montreal Press, Montreal, 1992.

[KR] N. Kurokawa, *Examples of eigenvalues of Hecke operators of Siegel cusp forms of degree two*, Invent. Math. **49** (1978), 365–399.

[LL] J.-P. Labesse and R. Langlands, *L-indistinguishability for* $SL(2)$, Canad. J. Math. **31** (1979), 726–785.

[LS] J.-P. Labesse and J. Schwermer, *On liftings and cusp cohomology of arithmetic groups*, Invent. Math. **83** (1986), 383–401.

[L1] R. P. Langlands, *On the zeta functions of some simple Shimura varieties*, Canad. J. Math. **31** (1979), 1121–1216.

[L2] ——, *Modular forms and* ℓ*-adic representations*, Modular functions of one variable. II, Lecture Notes in Math., vol. 349, Springer-Verlag, New York, 1973, pp. 362–499.

[L3] ——, *Les débuts d'une formule des traces stable*, Publ. Math. Univ. Paris VII, 1982.

[L4] ——, *Automorphic representations, Shimura varieties, and motives*, Automorphic Forms, Representations, and L-functions, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 205–246.

[L5] ——, *On the classification of irreducible representations of real algebraic groups*, Representation theory and harmonic analysis on semisimple groups, Math. Surveys Monographs, vol. 31, Amer. Math. Soc, Providence, RI, 1989, pp. 101–170.

[L6] ——, *Problems in the theory of automorphic forms*, Lectures in modern analysis and applications, Lecture Notes in Math., vol. 170, Springer-Verlag, New York, 1970, pp. 18–86.

[L7] ——, *Stable conjugacy; definitions and lemmas*, Canad. J. Math. **31** (1979), 700–725.

[LR] R. Langlands and M. Rapoport, *Shimuravarieten und gerben*, J. Reine Angew. Math. **378** (1987), 113–220.

[Li] J.-S. Li, *Non-vanishing theorems for the cohomology of certain arithmetic quotients*, J. Reine Angew. Math. **428** (1992), 177–217.

[M] *Zeta-functions of Picard Modular Surfaces* (R. Langlands and D. Ramakrishnan, eds.), Les Publications CRM, Univ. Montreal Press, 1992.

[Mi1] J. Milne, *Shimura varieties and motives,* these Proceedings, vol. 2, pp. 447–523.

[Mi2] ——, *The points on a Shimura variety modulo a prime of good reduction,* in *Zeta-functions of Picard Modular Surfaces* (R. Langlands and D. Ramakrishnan, eds.), Univ. Montreal Press, 1992.

[Mi3] ——, *The conjecture of Langlands and Rapoport for Siegel modular varieties,* Bull. Amer. Math. Soc. (N. S.) **24** (1991), 335–341.

[MW] C. Moeglin and J.-L. Waldspurger, *Le spectre residuel de* $GL(n)$, Ann. Sci. École Norm. Sup. (4) **22** (1989), 605–674.

[Ra] M. Rapoport, *On the bad reduction of Shimura varieties,* vol. 2 (L. Clozel and J. Milne, eds.), Academic Press, New York, 1990, pp. 253–321.

[R1] J. D. Rogawski, *Automorphic representations of unitary groups in three variables,* Ann. of Math. Stud., vol. 123, Princeton Univ. Press, Princeton, NJ, 1990.

[R2] ——, *The multiplicity formula for A-packets,* in *Zeta-functions of Picard Modular Surfaces* (R. Langlands and D. Ramakrishnan, eds.), Univ. Montreal Press, 1992.

[R3] ——, *Analytic expression for the number of points mod* p, in *Zeta-functions of Picard Modular Surfaces* (R. Langlands and D. Ramakrishnan, eds.), Univ. Montreal Press, 1992.

[Sc] J. Schwermer, *On arithmetic quotients of the Siegel upper half space of degree two,* Compositio Math. **58** (1986), 233–258.

[Sh1] D. Shelstad, *Notes on L-indistinguishability (based on a lecture by R. P. Langlands),* Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, Part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 193–203.

[Sh2] ——, *L-indistinguishability for real groups,* Math. Ann. **259** (1982), 385–430.

[S1] G. Shimura, *Correspondences modulaires et les fonctions* ζ *des courbes algébriques,* J. Math. Soc. Japan **10** (1958), 1–18.

[S2] ——, *On the zeta-functions of the algebraic curves uniformized by certain automorphic functions,* J. Math. Soc. Japan **13** (1961), 275–331.

[S3] ——, *On modular correspondences for* $Sp(N, \mathbf{Z})$ *and their congruence relations,* Proc. Nat. Acad. Sci. USA **46** (1963), 824–828.

[T] J. Tate, *Number theoretic background,* Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, part 2, Amer. Math. Soc., Providence, RI, 1979, pp. 3–26.

[Ta1] R. Taylor, *On Galois representations associated to Hilbert modular forms,* Invent. Math. **98** (1989), 265–280.

[Ta2] ——, *On the* ℓ*-adic cohomology of Siegel threefolds,* Invent. Math. (to appear).

[Ti] J. Tits, *Reductive groups over local fields,* Automorphic Forms, Representations, and L-Functions, Proc. Sympos. Pure Math., vol. 33, Part 1, Amer. Math. Soc., Providence, RI, 1979, pp. 29–69.

[VZ] D. Vogan and G. Zuckerman, *Unitary representations with non-zero cohomology,* Compositio Math. **53** (1984), 51–90.

[W] R. Weissauer, *Differentialformen zu untergruppen der Siegelschen Modulgruppe zweiten grades,* J. Reine Angew. Math. **391** (1988), 100–156.

[Z] S. Zucker, L^2*-Cohomology of Shimura varieties,* vol. 2 (L. Clozel and J. Milne, eds.), Academic Press, New York, 1990, pp. 377–391.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Hodge-de Rham Structures and Periods of Automorphic Forms

MICHAEL HARRIS

Introduction

The Hodge-de Rham (HdR) structure (the name is apparently due to Harder) of a (mixed) motive over a number field is what is left of the motive when one forgets everything except its (mixed) Hodge structure and periods.[1] If one believes the Hodge conjecture, then one believes that a pure motive is characterized up to isomorphism by its HdR structure (cf. Proposition 1.2). This fact is reflected in the conjectures of Deligne and Beilinson, which predict (among other things) relations between the L-functions of motives and their HdR structures. This would seem to justify studying HdR structures for their own sakes, but in fact very little of interest is known about them, perhaps because one does not know how to begin to characterize those HdR structures, abstractly defined, which actually come from motives. The question properly belongs to transcendence theory, but this remark sheds little light on the problem.

The cohomology of Shimura varieties, with coefficients in local systems, provides an abundant supply of examples of motives. One can say a lot about their HdR structures, thanks largely to Shimura's theory of canonical models and to the theory of automorphic forms. This is fortunate, since such motives are the only ones whose L-functions are at all well understood analytically (and even then only in certain cases). The present article is primarily a report on what is known about the HdR structures underlying these motives.

1991 *Mathematics Subject Classification.* Primary 11F67, 11G18; Secondary 11F70, 14D07, 14G35.

This work was partially supported by NSF Grant No. DMS-8901101 and by the NAS exchange program with the USSR.

This paper is in final form and no version of it will be submitted for publication elsewhere.

[1] G. Anderson has introduced the more general notion of arithmetic Hodge structures, whose Hodge types are allowed to be fractional.

A characteristic feature of the theory of Shimura varieties is that the same compatible systems of ℓ-adic representations occur in the cohomology of different Shimura varieties. The Tate conjecture then predicts isomorphisms between the corresponding motives, and a fortiori between the associated HdR structures. This leads to the problem of proving *period relations* for arithmetically normalized automorphic forms on different groups. An extreme case is the problem of proving that a Tate class in the cohomology (with twisted coefficients) of a Shimura variety is also a Hodge class. As mentioned in Tate's talk, this has now been carried out successfully in a number of cases (as of this writing it appears to be known for all codimension one Tate classes). Other examples are discussed in §§4.5 and 5.2. Since automorphic forms naturally occur as differential forms of pure Hodge type, it appears that the automorphic theory does not suffice to prove isomorphisms of HdR structures, but only to compare the periods of cohomology classes of pure Hodge type that are rational in *coherent cohomology*, viewed as the associated graded object for the Hodge filtration on de Rham cohomology. The data consisting of such periods forms what in §1.1 is called an *$H^{p,q}$-structure.* The $H^{p,q}$-structures seem to be the geometric structures most closely adapted to the automorphic theory, and the proof of isomorphisms between sub-$H^{p,q}$-structures of the cohomology of different Shimura varieties is, in a sense, the principal objective of the theory described in this article.

A particular family of examples appears several times in these notes. These are the (polarized) *regular* motives, which have the property that their Hodge components are of rank ≤ 1 over the coefficient field. Such motives are emphasized by Clozel in his work on the Galois representations attached to automorphic forms on $\mathrm{GL}(n)$ **[Cl]**. Clozel's work relies heavily on the study of Kottwitz of a class of Shimura varieties whose cohomology is especially simple. It is conjectured that the essential part of the cohomology of these varieties consists of exterior powers of regular motives. The ℓ-adic version of this conjecture has been established in most cases by Kottwitz and R. Taylor. As explained in §§4.5 and 5.7, the corresponding statement about HdR structures (or $H^{p,q}$-structures) would have interesting consequences for the Deligne conjecture on critical values of L-functions, in the context of regular motives over $\mathbb{Q}$.

The first section presents HdR structures abstractly, as objects belonging to linear algebra, and introduces the quadratic period invariants attached to regular HdR structures and their polarizations. The relation between HdR and $H^{p,q}$-structures is also discussed. The second section describes the mixed HdR structures on the cohomology of Shimura varieties, and explains how a conjecture of Faltings on the degeneration of a certain spectral sequence computing this cohomology follows immediately from M. Saito's theory of mixed Hodge modules; the latter result does not seem to have appeared in print up to now. The most familiar automorphic forms are vector-valued harmonic differential forms; thus de Rham and Dolbeault cohomology are the invariants of Shimura varieties most directly connected with the analytic theory of automorphic forms. This was first observed in the 50s and 60s; the third section explains the relation between automorphic forms and cohomology in the language of representation theory. In particular, it is shown that the analytic Hodge structure on L_2-cohomology of Shimura varieties, defined by Langlands and Arthur, coincides (via Zucker's conjecture; cf. Theorem 3.2.3) with Saito's functorial Hodge structure on intersection cohomology, when the coefficients are sufficiently regular. The example studied by Clozel and Kottwitz is worked out in some detail in the fourth section; here the Hodge types are especially easy to describe (4.2.1).

The material presented in the first four sections is reasonably systematic; difficult questions remain, but the correct framework seems to be in place. By contrast, the theory of automorphic L-functions, discussed in the last two sections, is still somewhat ad hoc. The fifth section describes some examples in which the special values of L-functions can be related to periods of automorphic forms. The message of this section, if any, is that this is a comparatively rare phenomenon, and that the well-known example of elliptic modular forms sheds little light on the general case. The final section describes a technique, originating in work of Rallis on the theta correspondence, for proving some of the period relations predicted by the Tate conjecture.

Most of the new material in §§2 and 3, especially Proposition 3.3.9, Theorem 3.5.3, and the presentation of Theorem 2.2.7 (due in substance to M. Saito), represents joint work with Zucker. The ideas in §6 have developed during the course of many discussions with Kudla and Rallis, and while I was a visitor at the Steklov Institute of Mathematics in Moscow in 1989–90. The discussion in §5.7 depends crucially on a recent analytic computation by Garrett. In preparing my talk at the conference and the present article I also benefited from conversations with Blasius, Brylinski, Clozel, Deligne, Franke, Gross, Jannsen, Kudla, Ramakrishnan, Rogawski, Morihiko Saito, and Zucker.

1. Hodge-de Rham structures

Let k be a field of characteristic zero and E a number field. A *mixed Hodge-de Rham structure* H (briefly, HdR structure) *over the field* k, *with coefficients in* E, consists of the following data:

(1.0.1) A $(k \otimes_{\mathbb{Q}} E)$-module H_{DR} of finite rank, endowed with an increasing filtration $\cdots \subset W_i H_{\mathrm{DR}} \subset W_{i+1} H_{\mathrm{DR}} \subset \cdots$ (*weight filtration*) and a decreasing filtration $\cdots \subset F^i H_{\mathrm{DR}} \subset F^{i-1} H_{\mathrm{DR}} \subset \cdots$ (*Hodge filtration*), both $(k \otimes_{\mathbb{Q}} E)$-linear;

(1.0.2) For each embedding $\sigma\colon k \hookrightarrow \mathbb{C}$, an E-module H_σ of finite rank, endowed with an increasing filtration $\cdots \subset W_i H_\sigma \subset W_{i+1} H_\sigma \subset \cdots$ (*weight filtration*) and, if $\sigma(k) \subset \mathbb{R}$, an E-linear involution F_σ of H_σ that preserves $W_\cdot H_\sigma$;

(1.0.3) For each $\sigma: k \hookrightarrow \mathbb{C}$, an E-linear isomorphism

$$I_\sigma: H_\sigma \otimes_{\mathbb{Q}} \mathbb{C} \xrightarrow{\sim} H_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C}$$

that respects the weight filtrations, and such that, if $\sigma(k) \subset \mathbb{R}$, then

$$I_\sigma \circ (F_\sigma \otimes \imath) \circ I_\sigma^{-1} = 1 \otimes \imath \in \mathrm{Aut}(H_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C}),$$

where $\imath$ = complex conjugation;

(1.0.4) For each $i \in \mathbb{Z}$ let $\mathrm{gr}_i^W H$ be the collection $(\mathrm{gr}_i^W H_{\mathrm{DR}}, \{\mathrm{gr}_i^W H_\sigma\})$, with induced $k \otimes E$- (resp. E-) linear structures, induced involutions F_σ, and induced Hodge filtration

$$F^q \mathrm{gr}_i^W H_{\mathrm{DR}} = (F^q H_{\mathrm{DR}} \cap W_i H_{\mathrm{DR}})/(F^q H_{\mathrm{DR}} \cap W_{i-1} H_{\mathrm{DR}}).$$

For each σ, the filtration induced by $F^\cdot \mathrm{gr}_i^W H_{\mathrm{DR}}$ on $\mathrm{gr}_i^W H_\sigma \otimes_{\mathbb{Q}} \mathbb{C}$, via I_σ, is a pure $\mathbb{Q}$-Hodge structure of weight i.

In (1.0.3) and (1.0.4), the $\mathbb{Q}$-structure on H_σ is obtained by forgetting the E-structure. Then (1.0.1–1.0.4) are the data (a), (c), and (d) of the definition of Jannsen's category of mixed realizations **[J,** 2.1], to which we have added the coefficient field and the involutions F_σ (which are actually redundant). The category of mixed HdR structures (without coefficients) over k, denoted $(\mathrm{MHdR})_k$, is abelian (cf. **[J,** 2.3]). The standard example of a mixed HdR structure is that attached to a mixed motive M for absolute Hodge cycles over k **[J,** §§3, 4; **D4,** §1]; the coefficient field E is assumed to act by endomorphisms of M. We denote this structure $H(M)$. If X is a smooth quasi-projective variety over k and $i \in \mathbb{Z}$, $H^i_{\mathrm{DR}}(X)$, together with its Hodge and weight filtrations, $H^i_\sigma(X) := H^i_B(X \times_{\sigma(k)} \mathbb{C}, \mathbb{Q})$, together with its weight filtration, and F_σ = complex conjugation when $\sigma(k) \subset \mathbb{R}$, form a mixed HdR structure (without coefficients) with weights $\geq i$. The HdR structure H is *pure* if $\mathrm{gr}_i^W H \neq 0$ for at most one i; then M is said to be of weight i. Let $(\mathrm{HdR})_k$ be the subcategory of $(\mathrm{MHdR})_k$ whose objects are direct sums of pure HdR structures. The Tate objects $\mathbb{Q}(n) \in (\mathrm{HdR})_k$ are defined for any k.

When $k = \mathbb{Q}$, σ is unique, and we write I_∞ and F_∞ for I_σ and F_σ.

1.1. Suppose H is pure of weight w. Let $H^{p,w-p}_{\mathrm{DR}} = F^p H_{\mathrm{DR}}/F^{p+1} H_{\mathrm{DR}}$, with its induced k-rational structure. Fix $\sigma: k \hookrightarrow \mathbb{C}$ and, for any q, define

$$F^q H = I_\sigma^{-1}(F^q H_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C}), \qquad \overline{F^q} H = I_\sigma \circ (1 \otimes \imath) \circ I_\sigma^{-1}(F^q H_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C}),$$

$$H^{p,w-p} := F^p H \cap \overline{F}^{w-p} H,$$

as subspaces of $H_\sigma \otimes_{\mathbb{Q}} \mathbb{C}$. Then (1.0.4) implies that the natural map

$$(1.1.1) \qquad H^{p,w-p} \to F^p H_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C} \to H^{p,w-p}_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C},$$

where the first map is I_σ and the second is the natural projection, is an isomorphism. In other words, we have the Hodge decomposition

$$(1.1.2) \quad H_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C} \xrightarrow[\sim]{I_\sigma^{-1}} H_\sigma \otimes_{\mathbb{Q}} \mathbb{C} \cong \bigoplus H^{p,w-p} \cong \bigoplus H^{p,w-p}_{\mathrm{DR}} \otimes_{\sigma(k)} \mathbb{C}.$$

The data consisting of the E-modules H_σ, the graded $(k \otimes_{\mathbb{Q}} E)$-module $\bigoplus H_{DR}^{p,w-p}$, and the isomorphisms (1.1.2) (for each σ) will be called an $H^{p,q}$*-structure over* k (with coefficients in E).

The significance of the HdR structure is indicated by the following proposition, which was observed independently by Blasius and myself (at least):

1.2. PROPOSITION. *Assume the Hodge conjecture. Let* k *be a subfield of* $\mathbb{C}$. *Then the functor* $\mathcal{M}_k \to (\mathrm{HdR})_k$ *which, to any (pure) motive* M *over* k *assigns its* HdR *structure* $H(M)$, *is fully faithful.*

PROOF. It is well known that the Hodge conjecture implies that the functor $\mathcal{M}_{\mathbb{C}} \to (\mathrm{HdR})_{\mathbb{C}}$ is fully faithful. Suppose $M, N \in \mathcal{M}_k$ and $f \in \mathrm{Hom}_{(\mathrm{HdR})_k}(H(M), H(N))$. Without loss of generality, we may assume $M = H^w(X)$, $N = H^w(Y)$, where X and Y are algebraic varieties over k of dimension m and n, respectively. Then f comes from a morphism $\Lambda \in \mathrm{Hom}_{\mathcal{M}_{\mathbb{C}}}(M_{\mathbb{C}}, N_{\mathbb{C}})$, where Λ is $(\mathbb{Q}(m)\otimes)$ a codimension m cycle C on $X \times Y$; moreover, Λ induces a k-rational homomorphism of de Rham realizations. By deforming C within its numerical equivalence class, we may assume $\Lambda \in \mathrm{Hom}_{\mathcal{M}_{\bar{k}}}(M_{\bar{k}}, N_{\bar{k}})$, where $\bar{k}$ is the algebraic closure of k in $\mathbb{C}$. Then Λ belongs to the submotive $\mathcal{A} \subset \check{M} \otimes N$ generated by the images of codimension m algebraic cycles on $X \times Y$ over $\bar{k}$, twisted by $\mathbb{Q}(m)$. Now $\mathcal{A}$ is a motive of pure Hodge type $(0, 0)$, so it suffices to show that the functor is fully faithful on the subcategory of motives of pure Hodge type $(0, 0)$ (Artin motives over k). But this is an immediate consequence of **[DMOS,** II, Proposition 6.17], which identifies the latter category with the category of representations of $\mathrm{Gal}(\bar{k}/k)$.

1.3. The category $(\mathrm{MHdR})_k$ is obviously a (neutral) tensor category; if we are given a privileged embedding $\sigma: k \hookrightarrow \mathbb{C}$, then $(\mathrm{MHdR})_k$ comes with two obvious fiber functors, namely H_σ and H_{DR}. When $k = \mathbb{Q}$, we write H_B instead of H_σ. Then the fiber functor H_B, restricted to the essential image of $\mathcal{M}_{\mathbb{Q}}$ in $(\mathrm{HdR})_{\mathbb{Q}}$, defines the motivic Galois group $G_{\mathcal{M}}$, and $\mathrm{Isom}^{\otimes}(H_B, H_{DR})$ defines the *period torsor* $\mathfrak{P}$, which is a principal homogeneous space under $G_{\mathcal{M}}$.

This all actually makes sense if we restrict our attention to the category **CM**/$\mathbb{Q}$ of motives of abelian varieties over $\mathbb{Q}$ of potential **CM** type, and use absolute Hodge cycles as morphisms. For this Tannakian category with fiber functor H_B, the Galois group is the Taniyama group (cf. Schappacher's talk), and the period torsor $\mathfrak{P}_{CM}$ attached to $\mathrm{Isom}^{\otimes}(H_B, H_{DR})$ has been characterized up to isomorphism by Blasius (unpublished; cf. **[Mi,** I, Theorem 7.4]). An explicit identification of $\mathfrak{P}_{CM}$ would require an explicit determination of periods of CM abelian varieties.

1.4. The HdR structure of a pure motive M over $\mathbb{Q}$ contains all information used in the definition of the Deligne periods $c^+(M)$, $c^-(M)$, which appears in Deligne's conjecture on critical values of L-functions **[D3]**. Accordingly, for any HdR structure H of weight w over $\mathbb{Q}$, with coefficients in E, we can define the Deligne periods $c^+(H)$, $c^-(H) \in (E \otimes \mathbb{C})^\times / E^\times$, by the formulas in **[D3]**, provided Deligne's hypothesis is satisfied:

(1.4.1) $\quad F_\infty$ acts as a scalar on $H^{w/2, w/2}$ when w is even.

Under certain supplementary hypotheses, the Deligne period can be expressed simply in terms of inner products. Applications involving automorphic forms are described in §5. The formulas in 1.4.6 and 1.4.10 are proved in **[H7]**; the computations are all elementary. In what follows, the coefficient field E will always be *totally real*.

1.4.2. Definition. Let $H = (\mathrm{HdR})_{\mathbb{Q}}$ be pure of weight w, with coefficients in E. We say H is *regular* if $\dim_E H^{p,q} \leq 1$ for all p, q. If $H = H(M)$ for some motive M over $\mathbb{Q}$, we say M is *regular*.

Note that (1.4.1) is automatic for regular HdR structures.

1.4.3. Definition. Let $H \in (\mathrm{HdR})_{\mathbb{Q}}$ be pure of weight w. A (false) *polarization* of H is a homomorphism $\langle \cdot , \cdot \rangle : H \otimes H \to \mathbb{Q}(-w)$ of HdR structures, which determines a nondegenerate bilinear form $\langle \cdot , \cdot \rangle_{\mathrm{DR}}$ on H_{DR}, which is symmetric if w is even, and skew-symmetric is w is odd. If H has coefficients in E, we require that $\langle ex , y \rangle = \langle x , ey \rangle$ for all x, $y \in H$, $e \in E$.

To avoid confusion we use only the de Rham polarization $\langle \cdot , \cdot \rangle_{\mathrm{DR}}$ for calculation; via I_∞ this defines a pairing on H_B with the property that $(2\pi i)^w \langle H_B , H_B \rangle_{\mathrm{DR}} \subset \mathbb{Q}$. To get a true polarization one imposes the condition that the pairing $(2\pi i)^w \langle x , Cy \rangle_{\mathrm{DR}}$ on $H_{B,\mathbb{R}}$ be symmetric positive-definite, where C is the Weil operator. This is certainly true for the HdR structures of motives, but for us it is enough to know that $\langle \cdot , \cdot \rangle$ is nondegenerate. The fact that $\langle \cdot , \cdot \rangle$ is a morphism in $(\mathrm{HdR})_{\mathbb{Q}}$ immediately implies the following *Hodge-Riemann bilinear relations*:

(1.4.3.1) $$\langle F^p H_{\mathrm{DR}} , F^q H_{\mathrm{DR}} \rangle = 0 \quad \text{if} \quad p + q < w ;$$

(1.4.3.2) $$\langle \cdot , \cdot \rangle \text{ defines a nondegenerate pairing } H^{p,q} \otimes H^{q,p} \to \mathbb{Q}(-w)_{\mathrm{DR}} = \mathbb{Q} \quad \text{whenever } p + q = w .$$

We assume henceforward that H is regular and (falsely) polarized, with coefficients in E. Let $n = \dim_E(H)$; the hypotheses imply that $(-1)^{nw} = 1$. Let p_i, $i = 1 , \dots , n$, be integers such that $H^{p_i , w - p_i} \neq \{0\}$, so that $\dim_{E^+} H^{p_i , w - p_i} = 1$. Write $q_i = w - p_i$. We assume $p_1 > p_2 > \cdots > p_n$, so that $p_{n-i+1} = q_i$ for all i. Let

$$F_{\infty , i} : H^{p_i , q_i} \otimes \mathbb{C} \xrightarrow{\sim} H^{q_i , p_i} \otimes \mathbb{C}$$

be the isomorphism induced by F_∞ via the Hodge decomposition. Then $F_{\infty , i}$ is an isomorphism of one-dimensional $(E \otimes \mathbb{C})$-modules. Let ω_i denote a nontrivial $\mathbb{Q}$-rational element of $H_{\mathrm{DR}}^{p_i , q_i}$, viewed as an element of $H_{\mathrm{DR}} \otimes$

$\mathbb{C}$ via the Hodge decomposition (1.1.2). Then ω_i is an $(E \otimes \mathbb{C})$-basis for $H^{p_i, q_i} \otimes \mathbb{C}$. By (1.4.3.2), we may assume

(1.4.3.3) $$\langle \omega_i, \omega_{n-i+1} \rangle_{\mathrm{DR}} = 1, \qquad i = 1, \dots, [n/2].$$

Define the *quadratic periods* $Q_i = Q_i(H)$ of H by the relation

(1.4.3.4) $$F_\infty \omega_i = Q_i \cdot \omega_{n-i+1}, \qquad i = 1, \dots, n.$$

These invariants are well defined as elements of $(E \otimes \mathbb{C})^\times / E^\times$. If $a, b \in (E \otimes \mathbb{C})^\times$, we write a $\sim_E b$ if $a/b \in E^\times$. Since F_∞ is an involution, it is clear that

(1.4.3.5) $$Q_i(H) \cdot Q_{n-i+1}(H) \sim_E 1.$$

Combining (1.4.3.4) and (1.4.3.3), we see that

(1.4.3.6) $$Q_i(H) \sim_E \langle \omega_i, F_\infty \omega_i \rangle_{\mathrm{DR}},$$

where the right-hand side is interpreted in the obvious way as an element of $(E \otimes \mathbb{C})^\times$.

We attach discriminant factors to the polarization as follows. If w is odd (so n is even), let $d_B(H) = d_{\mathrm{DR}}(H) = 1$. Suppose w is even. Then we let

(1.4.4) $$d_{\mathrm{DR}}(H) = (-1)^{n/2} \quad \text{if } n \text{ is even},$$
$$d_{\mathrm{DR}}(H) = (-1)^{(n-1/2)} \cdot \langle \omega_{\frac{1}{2}(n+1)}, \omega_{\frac{1}{2}(n+1)} \rangle_{\mathrm{DR}} \quad \text{if } n \text{ is odd}.$$

Since w is even, $\langle \cdot, \cdot \rangle_{\mathrm{DR}}$ is symmetric, hence its restriction to $I_\sigma(H_B)$ can be diagonalized. Moreover, $\langle \cdot, \cdot \rangle_{\mathrm{DR}}$ takes values in $(2\pi i)^{-w}\mathbb{Q}$ on $I_\sigma(H_B)$, so it can even be diagonalized over $\mathbb{Q}$. Let $e_1, \dots, e_n$ be a basis for $I_\sigma(H_B)$, orthogonal relative to $\langle \cdot, \cdot \rangle_{\mathrm{DR}}$, and let

(1.4.5) $$d_B(H) = (2\pi i)^{nw} \cdot \prod_{i=1}^{n} \langle e_i, e_i \rangle_{\mathrm{DR}}.$$

1.4.6. PROPOSITION **[H7]**. *If H is a (falsely) polarized regular* HdR *structure, of weight w and rank n over the coefficient field E, then the Deligne periods $c^+(H)$ and $c^-(H)$ satisfy the relation*

$$c^+(H) \cdot c^-(H) \sim_E (2\pi i)^{-nw/2} \cdot [d_{\mathrm{DR}}(H)/d_B(H)]^{1/2} \prod_{j=1}^{[n/2]} Q_j(H).$$

We can go further if we now twist H by the HdR structures attached to algebraic Hecke characters (cf. Schappacher's talk). Let $\mathscr{K}$ be an imaginary quadratic field, and let $\chi: \mathscr{K}_{\mathbb{A}}^\times / \mathscr{K}^+ \to \mathbb{C}^\times$ be an algebraic Hecke character. Recall that this means that $\chi_\infty = \chi|_{\mathbb{C}^\times}$ is an algebraic character of $\mathbb{C}^\times$; i.e. $\chi_\infty(z) = z^{-k} \cdot \bar{z}^{-\lambda}$ for some integers $k = k(\chi)$, $\lambda = \lambda(\chi)$. We always assume $\lambda = 0$, $k > 0$, since the general case can easily be derived from this.

Let $\mathbb{Q}(\chi)$ denote the field generated by the values of $\chi_f = \chi|_{\mathbb{A}_{\mathscr{K}}^{f,\times}}$. Then $\mathbb{Q}(\chi)$ is either $\mathbb{Q}$ or a CM field. The constructions of **[DMOS; Bl1,** §3**]** attach to χ a rank-one absolute Hodge motive $M(\chi)$ over $\mathcal{K}$, of weight k, with coefficients in $\mathbb{Q}(\chi)$. The Galois representations on $M(\chi)_l$ are determined by the identity $L_{\mathcal{K}}(M(\chi), s) = L_{\mathcal{K}}(\chi, s)$, where the latter is the usual Hecke L-function. The motive $M(\chi)$ is potentially of CM type, and is constructed by applying absolute Hodge correspondences to the cohomology of elliptic curves with complex multiplication by $\mathcal{K}$.

Let $RM(\chi)$ denote the restriction of scalars of $M(\chi)$ to $\mathbb{Q}$; then $RM(\chi)$ is an absolute Hodge motive of rank two with coefficients in $\mathbb{Q}(\chi)$, and $RM(\chi)_{\mathrm{DR}}$ is naturally a $(\mathcal{K}, \mathbb{Q}(\chi))$-bimodule. We have

$$L_{\mathbb{Q}}(RM(\chi), s) = L_{\mathcal{K}}(M(\chi), s) = L(\chi, s). \tag{1.4.7}$$

1.4.8. Remark. The L-function $L(\chi, s)$ may also be realized as the L-function of a holomorphic new form $\theta(\chi)$ on $\mathrm{GL}(2)_{\mathbb{Q}}$, of weight $k+1$: for $\theta(\chi)$ we take the classical binary theta function of Hecke attached to χ, appropriately normalized. Thus Scholl's theorem **[Sc]** and (1.4.7) show that $L_{\mathbb{Q}}(RM(\chi), s)$ is the L-function of a Grothendieck motive, realized in the cohomology of a Kuga-Sato variety. It can be shown that the quadratic period $Q(\chi)$, to be introduced below, can be computed in either geometric realization of $RM(\chi)$ (this idea is basically due to Shimura **[Sh4]**; cf. **[HK]** for a refinement).

The Hodge type of $RM(\chi)$ is $(k, 0)+(0, k)$. Choose $\omega(\chi) \in RM(\chi)^{k,0}$, $\omega'(\chi) \in RM(\chi)^{0,k}$ to be $\mathbb{Q}(\chi)$-bases for the $\mathbb{Q}$-structures on the respective spaces. The restriction of χ to $\mathbf{A}^\times/\mathbb{Q}^\times$ is of the form $\chi_0 \cdot \|\cdot\|_{\mathbf{A}}^{-k}$, where χ_0 is a character of finite order. Let $\varepsilon_{\mathcal{K}}\colon \mathbf{A}^\times/\mathbb{Q}^\times \to \{\pm 1\}$ be the quadratic character attached to $\mathcal{K}/\mathbb{Q}$. For any Dirichlet character η let $g(\eta)$ be the Gauss sum attached to η and some additive character. Define $Q(\chi) \in (\mathbb{Q}(\chi) \otimes \mathbb{C})^\times/\mathbb{Q}(\chi)^\times$ by the formula

$$F_\infty(\omega(\chi)) = g(\chi_0^{-1} \cdot \varepsilon_{\mathcal{K}})^{-1} Q(\chi)\omega'(\chi), \tag{1.4.9}$$

where this formula is to be interpreted in analogy with (1.4.3.4). Up to an element of $\mathbb{Q}(\chi)^\times$, the Gauss sum does not depend on the choice of additive character. The appearance of the Gauss sum in (1.4.9) is related to computations in **[H7]**, specifically the analogue of (1.4.3.6).

Let $E(\chi) = E \cdot \mathbb{Q}(\chi)$. For simplicity of exposition we assume $E \cap \mathbb{Q}(\chi) = \mathbb{Q}$; however, the results below, appropriately modified, do not depend on this hypothesis. In this case $H \otimes RM(\chi)$ is a HdR structure with coefficients in $E(\chi)$. The following proposition describes when $H \otimes RM(\chi)$ satisfies hypothesis (1.4.1), and computes $c^+(H \otimes RM(\chi))$ in terms of quadratic periods and other elementary invariants.

1.4.10. Proposition **[H7**, 1.7.1, 1.7.6, 1.8]). *Assume the hypotheses of Proposition* 1.4.6. *The structure* $H \otimes RM(\chi)$ *satisfies* (1.4.1) *if and only if* $k = k(\chi)$ *does not equal any of the values* $w - 2p_i$, $i = 1, \dots, n$; *this is true if and only if* $(H \otimes RM(\chi))^{p,p} = 0$ *for all* p. *Set* $p_0 = -\infty$, $p_{n+1} = \infty$, *and suppose* $w - 2p_r + 1 \leq k \leq w - 2p_{r+1} - 1$ (*we say* χ belongs to the rth critical interval for H), *for some* $r \in \{0, \ldots, n\}$. *Let* $r' = \min(r, n-r)$. *Then*

(a) *If* w *is odd and* n *is even, then*

$$c^+(H \otimes RM(\chi)) \sim_{E(\chi)} (2\pi i)^{(-w-k)n/2} \cdot g(\chi_0 \cdot \varepsilon_{\mathscr{K}})^r \cdot Q(\chi)^{r-n/2} \cdot \prod_{j=1}^{r'} Q_j(H).$$

(b) *If* w *and* n *are both even, then*

$$c^+(H \otimes RM(\chi)) \sim_{E(\chi)} (2\pi i)^{(-w-k)n/2} \cdot g(\chi_0 \cdot \varepsilon_{\mathscr{K}})^r \cdot [d_{\mathrm{DR}}(H)/d_B(H)]^{1/2} \cdot Q(\chi)^{r-n/2} \cdot \prod_{j=1}^{r'} Q_j.$$

(c) *If* w *is even,* $n = 2\nu - 1$, *then*

$$c^+(H \otimes RM(\chi)) \sim_{E(\chi)} a^{\pm}(\chi) \cdot (2\pi i)^{-nw/2-k\nu} \cdot g(\chi_0 \cdot \varepsilon_{\mathscr{K}})^r \cdot [d_{\mathrm{DR}}(H)/d_B(H)]^{1/2} \cdot Q(\chi)^{r-\nu} \cdot \prod_{j=1}^{r'} Q_j,$$

where $a^{\pm}(\chi)$ *is a* (*simple*) *period of* $\omega(\chi)$, *depending on the sign of* F_∞ *acting on* $H^{\nu,\nu}$.

1.4.11. REMARKS. (a) The HdR structure $H(M)$ also determines the rational structures on the highest exterior power of Deligne cohomology, used in the formulation of Beilinson's conjectures (cf. the talks of Soulé, Nekovar, and Scholl). It would be interesting to know whether the formula for the regulator (in Scholl's reformulation of the Beilinson conjectures, say) simplifies in an analogous way.

(b) B. Gross has pointed out that the argument used to prove 1.4.10 implies similar formulas for $c^+(H \otimes H')$, where H and H' are two polarized regular motives of arbitrary dimension, provided $H \otimes H'$ satisfies (1.4.1).

1.5. Let H be an HdR structure over $\mathbb{Q}$, with coefficients in E. Serre's recipe (cf. Deninger's talk) attaches to the Hodge structure underlying H a Γ-factor $L_\infty(H, s)$. Following Deligne, we call the integer m a *critical value* for H if neither $L_\infty(H, s)$ nor $L_\infty(\check{H}, 1-s)$ has a pole at $s = m$. Note that

(1.5.1) If $H^{p,p} = \{0\}$ for all p, then H has at least one critical value.

If H and χ are as in §1.4 then the set of critical values of $H \otimes RM(\chi)$ depends only on w and $k(\chi)$; we call this set $C(w, k(\chi))$.

The following conjecture, which is certainly not due to me, is a special case of a far-reaching conjecture, which appears to be generally believed by analytic number theorists. As far as I know it has not appeared in print.

1.5.2. NONVANISHING CONJECTURE. Let M be a regular motive over $\mathbb{Q}$ of weight w. Let $\mathcal{K}$ be an imaginary quadratic extension of $\mathbb{Q}$. For every r, every weight k in the rth critical interval for $H(M)$, and every integer $m \in C(w, k(\chi))$, there exists an algebraic Hecke character χ of $\mathcal{K}$, of weight k, such that the restriction of χ to $\mathbf{A}^\times/\mathbb{Q}^\times$ equals $\|\cdot\|_{\mathbf{A}}^{-k}$, and such that $L(M \otimes RM(\chi), m) \neq 0$.

We may then derive the following corollary to Proposition 1.4.10:

1.5.3. COROLLARY. *Let M be a (polarized) regular motive over $\mathbb{Q}$, with coefficients in the totally real field E. The invariants $Q_j(M) = Q_j(H(M)) \in (E \otimes \mathbb{C})^\times/E^\times$, $j = 1, \dots, [n/2]$, can be expressed, up to elementary factors, in terms of the periods of $M(\chi)$ and $c^+(M \otimes RM(\chi))$, as χ runs through the distinct critical intervals for M. In particular, if Deligne's conjecture and the nonvanishing conjecture (1.5.2) hold for the motives $M \otimes RM(\chi)$ and M, respectively, then the $Q_j(M)$ can be expressed, up to multiplication by $E^\times$ and elementary factors, in terms of critical values of $L(M \otimes N(\chi), s)$ and of L-functions of algebraic Hecke characters of $\mathcal{K}$.*

PROOF. By (1.5.1) and Proposition 1.4.10, $M \otimes RM(\chi)$ has critical values whenever k is in one of the critical intervals for $H(M)$. It thus suffices to remark that the $Q(\chi)$ and $a^\pm$ can be expressed as in terms of special values of L-functions of algebraic Hecke characters of $\mathcal{K}$, by the theorem of Blasius **[Bl1]** (and, in this case, of Goldstein-Schappacher **[GSc]**).

The Tate conjecture implies that the HdR structure $H(M)$ is determined by $L(M, s)$. Deligne's conjecture on critical values can be seen in this light: at least part of the period isomorphism I_∞ can be reconstructed as the irrational part of a critical value of $L(M, s)$. It is reasonable to ask "how many" of the periods of M can be recovered in terms of Deligne's conjecture. Corollary 1.5.3 asserts that, provided M is regular and polarized over a totally real coefficient field E, at least the quadratic periods of M can be recovered, assuming the nonvanishing conjecture. This accounts roughly for the square root of the number of entries in the period matrix.

1.6. Recall that the quadratic periods were defined in terms of cohomology classes ω_i of pure Hodge type: the $H^{p,q}$-structure defined in 1.1. As Deligne already observed in **[D3]**, his conjecture on critical values really depends only on the $H^{p,q}$-structure, and not on the whole HdR structure. This is fortunate, since, as noted in the introduction, automorphic forms naturally arise as differential forms of pure Hodge type; the DR rational Hodge filtration, unlike the Hodge decomposition, is not obviously accessible by the theory of harmonic forms. We are left with the awkward possibility that two motives with nonisomorphic HdR structures might have isomorphic $H^{p,q}$-structures. The following proposition shows that this is unlikely:

1.6.1. PROPOSITION. *Assume Deligne's "hope" that every Hodge cycle is absolutely Hodge. Let M and N be pure motives over $\mathbb{Q}$ with isomorphic*

$H^{p,q}$-structures. Then the Hodge-de Rham structures $H(M)$ and $H(N)$ are isomorphic over $\mathbb{Q}$.

1.6.2. REMARK. In the original version of this proposition, Deligne's "hope" was replaced by the Hodge conjecture; Jannsen pointed out the sufficiency of the present formulation.

PROOF. Let $\psi\colon H_B(M) \xrightarrow{\sim} H_B(N)$ be an isomorphism inducing an isomorphism of $H^{p,q}$-structures. Thus (via I_∞):

(1.6.3) ψ takes the filtration $F^{\cdot}H_{\mathrm{DR}}(M)$ to $F^{\cdot}H_{\mathrm{DR}}(N)$, and induces a $\mathbb{Q}$-rational isomorphism on the associated graded spaces.

We consider $\psi \in H(\check{M}) \otimes H(N) = H(\check{M} \otimes N)$ as a pair $(\psi_B, \psi_{\mathrm{DR}})$, with $\psi_B \in H_B(\check{M} \otimes N)$, $\psi_{\mathrm{DR}} \in H_{\mathrm{DR}}(\check{M} \otimes N) \otimes \mathbb{C}$, and $I_\infty(\psi_B) = \psi_{\mathrm{DR}}$; the goal is to prove that $\psi_{\mathrm{DR}} \in H_{\mathrm{DR}}(\check{M} \otimes N)$. But the hypothesis (1.6.2) implies that $\psi_{\mathrm{DR}} \in F^0 H_{\mathrm{DR}}(\check{M} \otimes N) \otimes \mathbb{C}$, hence ψ is a Hodge cycle. By our assumption, ψ is absolutely Hodge, hence $\psi_{\mathrm{DR}} \in H_{\mathrm{DR}}(\check{M} \otimes N) \otimes \overline{\mathbb{Q}}$. It remains to show that ψ is fixed by $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ (acting on the factor ψ_{DR}). Let $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and consider $\xi(\tau) = \psi - \psi^\tau$. By hypothesis, $\xi(\tau)$ induces the zero map from $\mathrm{gr}_F^{\cdot}(M)$ to $\mathrm{gr}_F^{\cdot}(N)$. Thus $\xi(\tau) \in F^1 H_{\mathrm{DR}}(\check{M} \otimes N) \otimes \overline{\mathbb{Q}}$. But $\xi(\tau)$ is an absolute Hodge cycle, hence $\xi(\tau) = 0$. This completes the proof.

Let $G_{\mathcal{M},\mathrm{DR}}$ be the Galois group attached to the Tannakian category $\mathcal{M}_{\mathbb{Q}}$, with the fiber functor H_{DR}. The Hodge filtration on $\mathcal{M}_{\mathbb{Q}}$ corresponds to a parabolic subgroup $P_F \subset G_{\mathcal{M},\mathrm{DR}}$, rational over $\mathbb{Q}$. Let U_F be the unipotent radical of P_F. If H is a pure HdR structure of weight w over $\mathbb{Q}$, the *Hodge numbering* on $\mathrm{gr}_F^{\cdot}(H)$ is the homomorphism $\mu_H\colon \mathbb{G}_m \to \mathrm{GL}(\mathrm{gr}_F^{\cdot}(H_{\mathrm{DR}}))$ such that $\mu(t)$ acts as t^{-p} on $H_{\mathrm{DR}}^{p,w-p}$. Correspondingly, there is a homomorphism $\mu\colon \mathbb{G}_m \to P_F/U_F$ which induces μ_H on any $\mathrm{gr}_F^{\cdot}(H)$, and μ is defined over $\mathbb{Q}$. The Hodge decomposition defines a canonical lifting $\mu_\infty\colon \mathbb{G}_{m,\mathbb{C}} \to P_{F,\mathbb{C}}$. On the other hand, by Levi's theorem there exist $\mathbb{Q}$-rational liftings $\tilde{\mu}\colon \mathbb{G}_m \to P_F$; the set of such liftings is homogeneous under $U_F(\mathbb{Q})$. Let $d\tilde{\mu}$ and $d\mu_\infty$ be the corresponding maps on Lie algebras, and let

$$X = d\mu_\infty(t \cdot d/dt) - d\tilde{\mu}(t \cdot d/dt) \in \mathrm{Lie}(U_F(\mathbb{Q}))\backslash \mathrm{Lie}(U_F(\mathbb{C})).$$

Now the Lie algebra of the Mumford-Tate group of any motive M is itself a motive, and Proposition 1.6.1, applied to these Lie algebras, says that the element X is somehow determined by the motivic Galois group $G_{\mathcal{M}}$ (attached to H_B) and μ_∞; that the transition from the $H^{p,q}$- structure to the HdR structure on $\mathrm{Lie}(G_{\mathcal{M},\mathrm{DR}})$ introduces no new periods. Is this related to the characterization of the essential image of $\mathcal{M}_{\mathbb{Q}}$ in $(\mathrm{HdR})_{\mathbb{Q}}$? The only result I know of in this direction treats the case of elliptic curves, and is due to Katz [**Ka**,§4]. He shows that, for an elliptic curve in the Weierstrass model, the complex conjugate $\overline{\omega}$ of the holomorphic differential $\omega = dx/y$ can be written $\overline{\omega} = q \cdot \eta + u \cdot \omega$, where $\eta = x\,dx/y$ is the canonical differential of the second kind and u/q can be expressed in terms of a classical modular form of weight 2 which is nearly holomorphic in Shimura's sense. Since η is rational with respect to the de Rham structure, u/q precisely measures the discrepancy between the $H^{p,q}$-structure and the HdR structures in this case.

2. Shimura varieties and mixed Hodge-de Rham structures

2.1. Automorphic vector bundles. Let G be a reductive group over $\mathbb{Q}$, and let X be a $G(\mathbb{R})$-conjugacy class of homomorphisms $h\colon \underline{S} = R_{\mathbb{C}/\mathbb{R}}\mathbb{G}_{m,\mathbb{C}} \to G_{\mathbb{R}}$, such that the pair (G, X) satisfies the axioms for a Shimura variety $\mathrm{Sh}(G, X)$:

$$_K\mathrm{Sh}(G, X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbf{A}^f)/K$$
$$\mathrm{Sh}(G, X)(\mathbb{C}) = \varprojlim_K {}_K\mathrm{Sh}(G, X)(\mathbb{C}),$$

as in Milne's talk. Here K runs through open compact subgroups of $G(\mathbf{A}^f)$, which we call *level subgroups*. Let $E(G, X)$ be the reflex field of (G, X), so that $\mathrm{Sh}(G, X)$ has a canonical model over $E(G, X)$. We do not assume that $w = h|_{\mathbb{G}_{m,\mathbb{R}}}$ is defined over $\mathbb{Q}$. Milne has explained that, if (ρ, V) is a $\mathbb{Q}$-rational representation of G, such that $\rho \circ w$ is defined over $\mathbb{Q}$, then the local system

$$\widetilde{V}^{\nabla} := \varprojlim_K G(\mathbb{Q})\backslash X \times V(\mathbb{Q}) \times G(\mathbf{A}^f)/K$$

is a variation of Hodge structures (VHS) over $\mathrm{Sh}(G, X)$, of weight w_ρ, where $\rho \circ w(t) = t^{-w_\rho}$. Recall that, if $[h, g]$ is the image in $\mathrm{Sh}(G, X)(\mathbb{C})$ of $(h, g) \in X \times G(\mathbf{A}^f)$, then the Hodge structure on the fiber $(\widetilde{V}^{\nabla} \otimes \mathbb{C})_{[h,g]}$ is given by $((\widetilde{V}^{\nabla} \otimes \mathbb{C})_{[h,g]})^{p,q} = V^{p,q}$, where for $z \in \underline{S}(\mathbb{R}) \cong \mathbb{C}^\times$, $h(z)$ acts as $z^{-p} \cdot \overline{z}^{-q}$ on $V^{p,q} \subset V \otimes \mathbb{C}$. We always assume $\rho \circ w$ defined over $\mathbb{Q}$. Let $d = \dim(X) = \dim(\mathrm{Sh}(G, X))$.

Let $\widetilde{V}$ be the vector bundle associated to $\widetilde{V}^{\nabla}$. This is an example of an *automorphic vector bundle*. More general automorphic vector bundles are defined by the following procedure (for more details, see **[H1, Mi]**). Let $h \in X$, and let $K_h \subset G$ be its centralizer. Let $\mathfrak{k}_h = \mathrm{Lie}(K_h)$, $\mathfrak{g} = \mathrm{Lie}(G)$, and write

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{k}_{h,\mathbb{C}} \oplus \mathfrak{p}^+ \oplus \mathfrak{p}^- = \mathfrak{g}^{0,0} \oplus \mathfrak{g}^{1,-1} \oplus \mathfrak{g}^{-1,1}, \tag{2.1.1}$$

where $\mathfrak{g}^{p,q}$ is the (p, q)-Hodge component for the adjoint representation of G on $\mathfrak{g}$ and the point h; $\mathfrak{p}^+$ and $\mathfrak{p}^-$ correspond, respectively, to the holomorphic and antiholomorphic tangent spaces at $h \in X$. Let $\mathfrak{P}_h = \mathfrak{k}_{h,\mathbb{C}} \oplus \mathfrak{p}^-$; this is a parabolic subalgebra of $\mathfrak{g}_{\mathbb{C}}$ with Levi component $\mathfrak{k}_{h,\mathbb{C}}$ and (abelian) unipotent radical $\mathfrak{p}^-$. Let $\mathscr{P}_h \subset G$ be the connected subgroup with Lie algebra $\mathfrak{P}_h$, and let $\check{X} = \check{X}(G, X)$ be the flag variety of parabolic subgroups of G conjugate to $\mathscr{P}_h$; it has a natural canonical model over

$E(G, X)$. The map $\beta\colon h \mapsto \mathscr{P}_h$ defines a $G(\mathbb{R})$-equivariant embedding (the Borel embedding) of X in $\check{X}(\mathbb{C})$. If $\mathscr{E}$ is a G-equivariant algebraic vector bundle on $\check{X}$, let

$$[\mathscr{E}] = \varprojlim_{K} G(\mathbb{Q})\backslash\beta^*(\mathscr{E}) \times G(\mathbf{A}^f)/K. \tag{2.1.2}$$

Thus $\tilde{V} = [V \times \check{X}]$, where G acts diagonally on $V \times \check{X}$. For any $\mathscr{E}$, $[\mathscr{E}]$ is a $G(\mathbf{A}^f)$-equivariant algebraic vector bundle over $\mathrm{Sh}(G, X)$ for the obvious right action of $G(\mathbf{A}^f)$. The automorphic vector bundles are the $[\mathscr{E}]$ obtained in this way; those of the form $\tilde{V}$ are called *flat*; they are endowed with natural flat connections $\nabla_V\colon \tilde{V} \to \tilde{V} \otimes \Omega^1_{\mathrm{Sh}(G,X)}$.

2.1.3. REMARK. If we fix a point $h \in X$, then the tensor category of G-equivariant vector bundles on $\check{X}$ is naturally equivalent to the category of representations of $\mathscr{P}_h$. In particular, any finite-dimensional representation (τ, W_τ) of K_h gives rise to an automorphic vector bundle $\mathscr{W}_\tau$ on $\mathrm{Sh}(G, X)$.

2.2. We write ${}_K\mathrm{Sh} = {}_K\mathrm{Sh}(G, X)$ ($K \subset G(\mathbf{A}^f)$ open), $\mathrm{Sh} = \mathrm{Sh}(G, X)$, until further notice. The Shimura varieties ${}_K\mathrm{Sh}$ are, in general, quasi-projective but not complete. A good class of compactifications, the *toroidal compactifications*, is constructed in **[AMRT]** (and adelically in **[H2, P1]**); these have turned out to be very useful in applications to Hodge theory. Their precise definition is omitted here, but we will have more to say about their structure in §3.5 below. For the time being, it is enough to say that we can choose toroidal compactifications $j_\Sigma\colon {}_K\mathrm{Sh} \hookrightarrow {}_K\mathrm{Sh}_\Sigma$, depending on auxiliary data Σ, which are smooth, projective, and with the property that $Z_\Sigma = {}_K\mathrm{Sh}_\Sigma - {}_K\mathrm{Sh}$ is a divisor with normal crossings. We only consider such ${}_K\mathrm{Sh}_\Sigma$. The choice of Σ introduces an inevitable element of arbitrariness, but the inverse limit $\mathrm{Sh}^\sim = \varprojlim_{K,\Sigma} {}_K\mathrm{Sh}_\Sigma$ taken over all K and Σ is well defined, if unwieldy from the geometric point of view. The coherent cohomology of $\mathrm{Sh}^\sim$ turns out to be a natural invariant of Sh.

If $[\mathscr{E}]$ is an automorphic vector bundle on Sh, we write $[\mathscr{E}]$ again for its descent to ${}_K\mathrm{Sh}$ (any K). In **[H2]** a *canonical extension* $[\mathscr{E}]^{\mathrm{can}} = [\mathscr{E}]_\Sigma$ is defined, extending a construction of Mumford **[Mu]** and Faltings **[FC]**; $[\mathscr{E}]^{\mathrm{can}}$ is a vector bundle over ${}_K\mathrm{Sh}_\Sigma$, and the construction of $[\mathscr{E}]^{\mathrm{can}}$ is compatible with the inverse limit over K and Σ. We can thus define $[\mathscr{E}]^{\mathrm{can}}$ over $\mathrm{Sh}^\sim$.

The main property of the canonical extension is that the functor $\mathscr{E} \mapsto [\mathscr{E}]^{\mathrm{can}}$ (from G-equivariant vector bundles on $\check{X}$ to vector bundles on ${}_K\mathrm{Sh}_\Sigma$) commutes with tensor operations. If $\tilde{V}$ is a flat automorphic vector bundle, then $\tilde{V}^{\mathrm{can}}$ is the Deligne extension of $\tilde{V}$ to a vector bundle for which ∇_V has regular singularities **[D1]**. Viewing the bundles Ω^i_{Sh} (K implicit), $i = 0, \ldots, d$, as the automorphic vector bundles $[\Omega^i_{\check{X}}]$, we have $(\Omega^i_{\mathrm{Sh}})^{\mathrm{can}} \cong \Omega^i_{\mathrm{Sh}}(\log(Z_\Sigma))$, the differential forms with logarithmic poles along Z_Σ **[D1]**. Let $[\mathscr{E}]^{\mathrm{sub}} = [\mathscr{E}]^{\mathrm{sub}}_\Sigma = [\mathscr{E}]^{\mathrm{can}} \otimes \mathscr{I}_{Z_\Sigma}$, where $\mathscr{I}_{Z_\Sigma}$ is the sheaf of ideals defining

Z_Σ; thus

$$(\Omega^d_{\mathrm{Sh}})^{\mathrm{sub}} \cong \Omega^d_{\mathrm{Sh}_\Sigma}. \tag{2.2.1}$$

2.2.2. PROPOSITION [**H4** §2]. *Let* ${}_K\mathrm{Sh}_\Sigma$, ${}_K\mathrm{Sh}_{\Sigma'}$ *be two toroidal compactifications of* ${}_K\mathrm{Sh}_\Sigma$. *Then there are canonical isomorphisms of sheaf cohomology*

$$H^\bullet({}_K\mathrm{Sh}_\Sigma, [\mathscr{E}]_{\Sigma'}) \xrightarrow{\sim} H^\bullet({}_K\mathrm{Sh}_{\Sigma'}, [\mathscr{E}]_{\Sigma'});$$
$$H^\bullet({}_K\mathrm{Sh}_\Sigma, [\mathscr{E}]^{\mathrm{sub}}_\Sigma) \xrightarrow{\sim} H^\bullet({}_K\mathrm{Sh}_{\Sigma'}, [\mathscr{E}]^{\mathrm{sub}}_{\Sigma'}).$$

Write $H^\bullet_K([\mathscr{E}]^{\mathrm{can}}) = H^\bullet({}_K\mathrm{Sh}_\Sigma, [\mathscr{E}]_\Sigma)$ and $H^\bullet_K([\mathscr{E}]^{\mathrm{sub}}) = H^\bullet({}_K\mathrm{Sh}_\Sigma, [\mathscr{E}]^{\mathrm{sub}}_\Sigma)$ (any Σ); then

$$\widetilde{H}^\bullet([\mathscr{E}]^{\mathrm{can}}) = \varinjlim_K H^\bullet_K([\mathscr{E}]^{\mathrm{can}}), \qquad \widetilde{H}^\bullet([\mathscr{E}]^{\mathrm{sub}}) = \varinjlim_K H^\bullet_K([\mathscr{E}]^{\mathrm{sub}})$$

are naturally graded $G(\mathbf{A}^f)$-modules which are *admissible* in the sense of representation theory (cf. the article of Kudla), and the natural map $\widetilde{H}^\bullet([\mathscr{E}]^{\mathrm{sub}}) \to \widetilde{H}^\bullet([\mathscr{E}]^{\mathrm{can}})$ is $G(\mathbf{A}^f)$-equivariant.

We let $H^\bullet_!([\mathscr{E}]) = \mathrm{Im}(\widetilde{H}^*([\mathscr{E}]^{\mathrm{sub}}) \to \widetilde{H}^\bullet([\mathscr{E}]^{\mathrm{can}}))$. For now, these cohomology groups all have coefficients in $\mathbb{C}$; descent to number fields is explained in 2.3.

Any G-equivariant vector bundle $\mathscr{E}$ on $\check{X}$ has a natural *Hodge filtration* $F^\bullet\mathscr{E}$, which is a decreasing filtration by G-equivariant algebraic subbundles. If $h \in X$, then the fiber $\mathscr{E}_{h,\mathbb{C}}$ decomposes as the sum $\bigoplus(\mathscr{E}_h)^{p,q}$, where $(\mathscr{E}_h)^{p,q} \subset \mathscr{E}_{h,\mathbb{C}}$ is the $z^{-p}\cdot\bar{z}^{-q}$-eigenspace for $h(z)$, $z \in \underline{S}(\mathbb{R}) \cong \mathbb{C}^\times$, and then

$$F^p\mathscr{E}_h = \bigoplus_{p' \geq p} (\mathscr{E}_h)^{p,q} \tag{2.2.3}$$

is the fiber at h of $F^p\mathscr{E}$ (cf. [**H1**, §3]). Let $\mathscr{E}^\infty$ be the C^∞-vector bundle associated to $\mathscr{E}$; then the decomposition $\mathscr{E}_{h,\mathbb{C}} \cong \bigoplus(\mathscr{E}_h)^{p,q}$ extends to a global splitting of C^∞-vector bundles, usually nonholomorphic [loc. cit]:

$$\mathscr{E}^\infty \cong \bigoplus \mathscr{E}^{p,q}. \tag{2.2.3.1}$$

The definition (2.2.3) extends immediately to complexes $(\mathscr{E}^\bullet, d^\bullet)$ of G-equivariant vector bundles $\mathscr{E}^i$ and G-equivariant algebraic differential operators d^i [loc. cit.]. In particular, if (ρ, V) is a representation of G, the *de Rham complex* $(\Omega^\bullet(V), d^\bullet)$, where $\Omega^i(V) = V \otimes \Omega^i\check{X}$ and d^i is exterior differentiation, has a Hodge filtration $F^\bullet\Omega^\bullet(V)$. Applying the functor $[\cdot]$, we obtain a filtered complex $([\Omega^\bullet(V)], [d^\bullet])$ of automorphic vector bundles on Sh, which is easily seen to coincide with the algebraic de Rham complex $(\Omega^\bullet(\widetilde{V}), (\nabla_V)^\bullet)$ of $(\widetilde{V}, \nabla_V)$. Of course, ∇_V satisfies Griffiths transversality (cf. [**Z1**, p. 263]), and the Hodge filtration is the one discussed in [**Z1**], up to a Tate twist; we refer to [loc. cit.] for details on the present construction.

In [F], Faltings constructed a subcomplex $\mathscr{K}^{\cdot}(\tilde{V}) \subset \Omega^{\cdot}(\tilde{V})$, the (dual) BGG complex of $\tilde{V}$, such that the inclusion is a quasi-isomorphism. To explain his construction, we assume V to be an *absolutely irreducible* representation of G; in particular, we drop the hypothesis that V be defined over $\mathbb{Q}$ (but $\rho \circ w$ is still assumed $\mathbb{Q}$-rational). Fix a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{k}_{h,\mathbb{C}}$; then $\mathfrak{h}$ is automatically a Cartan subalgebra of $\mathfrak{g}_{\mathbb{C}}$, and we let $W_c = W(\mathfrak{k}_{h,\mathbb{C}}, \mathfrak{h})$, $W = W(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h})$ be the respective Weyl groups, and let l be the length function on W. Fix a set R_c^+ of positive roots for $(\mathfrak{k}_{h,\mathbb{C}}, \mathfrak{h})$, and let R_n^+ be the set of roots for $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h})$ contained in $\mathfrak{p}^+$. Then $R^+ = R_c^+ \cup R_n^+$ is a set of positive roots of $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h})$; let $\rho = \frac{1}{2}\sum_{\alpha \in R^+} \alpha$. Let $W^1 = \{w \in W \mid w(R^+) \supset R_c^+\}$, $W^1(p) = \{w \in W^1 \mid l(w) = p\}$; then W^1 is a set of coset representatives of shortest length for $W_c \backslash W$, and $W^1 = \bigcup_{0 \le p \le d} W^1(p)$.

Now let $\Lambda \in \mathfrak{h}^*$ be the highest weight of V relative to R^+, and for $w \in W^1$ let $\Lambda(w) = w(\Lambda + \rho) - \rho$. Denote by $\tau(\Lambda, w)$ the representation of K_h with highest weight $\Lambda(w)$, and let $\mathscr{W}_{\tau(\Lambda, w)}$ be the corresponding automorphic vector bundle on Sh, as in Remark 2.1.3. Finally, let $\mathscr{K}^p(\tilde{V}) = \bigoplus_{w \in W^1(p)} \mathscr{W}_{\tau(\Lambda, w)}$. Then Faltings' dual BGG complex has the form $0 \to \mathscr{K}^0(\tilde{V}) \to \mathscr{K}^1(\tilde{V}) \to \cdots \to \mathscr{K}^d(\tilde{V}) \to 0$ and is a subcomplex of $(\Omega^{\cdot}(\tilde{V}), (\nabla_V)^{\cdot})$.

The quasi-isomorphism $\mathscr{K}^{\cdot}(\tilde{V}) \to \Omega^{\cdot}(\tilde{V})$ is evidently compatible with the Hodge filtrations (2.2.3). Faltings has shown that, over any ${}_K\mathrm{Sh}_\Sigma$, this extends to a filtered quasi-isomorphism of canonical extensions [F, §7]

$$\mathscr{K}^{\cdot}(\tilde{V})^{\mathrm{can}} \to \Omega^{\cdot}(\tilde{V})^{\mathrm{can}} = \Omega^{\cdot}(\log(Z_\Sigma)) \otimes \tilde{V}^{\mathrm{can}}.$$

(In general, the existence of a Hodge filtration on the logarithmic de Rham complex $\Omega^{\cdot}(\log(Z_\Sigma)) \otimes \tilde{V}^{\mathrm{can}}$ is a deep theorem of Schmid [Sch], but for Shimura varieties it is a simple consequence of the functoriality of the canonical extension; cf. [HZ, I, 4.4–4.5] for a conceptual explanation of this fact.) Each automorphic vector bundle $\mathscr{W}_{\tau(\Lambda, w)}$ is attached to an irreducible representation of K_h, hence is of pure Hodge type for the Hodge decomposition (2.2.3.1); we write $\mathscr{W}_{\tau(\Lambda, w)} = (\mathscr{W}_{\tau(\Lambda, w)})^{p(\Lambda, w), r(\Lambda, w)}$.

The canonical quasi-isomorphism $\tilde{V}^{\mathrm{can}} \to \Omega^{\cdot}(\log(Z_\Sigma)) \otimes \tilde{V}^{\mathrm{can}}$ factors through a quasi-isomorphism $\tilde{V}^{\mathrm{can}} \to \mathscr{K}^{\cdot}(\tilde{V})^{\mathrm{can}}$. On the other hand, Deligne's theory [D1] shows that $\tilde{V}^{\mathrm{can}}$ and $Rj_{\Sigma,*}(\tilde{V})$ are quasi-isomorphic; i.e., for all K and Σ

$$H^{\cdot}({}_K\mathrm{Sh}_\Sigma, \tilde{V}^{\mathrm{can}}) \cong H^{\cdot}({}_K\mathrm{Sh}_\Sigma, Rj_{\Sigma,*}(\tilde{V})) \cong H^{\cdot}({}_K\mathrm{Sh}, \tilde{V}^{\nabla}) \tag{2.2.4}$$

and these isomorphisms commute with maps ${}_{K'}\mathrm{Sh} \to {}_K\mathrm{Sh}$. Thus there are spectral sequences of cohomology computing $H^{\cdot}({}_K\mathrm{Sh}, \tilde{V}^{\nabla})$ which, in the limit over K, yield

(2.2.5)
$$E_1^{p,q} = \tilde{H}^q(\mathscr{K}^p(\tilde{V})^{\mathrm{can}}) \Rightarrow H^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla) \left(= \varinjlim_K H^{p+q}({}_K\mathrm{Sh}, \tilde{V}^\nabla)\right),$$

(2.2.6)
$$E_1^{p,q} = \tilde{H}^{p+q}(\mathrm{gr}_F^p \mathscr{K}^\cdot(\tilde{V})^{\mathrm{can}}) \Rightarrow H^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla).$$

Here (2.2.5), resp. (2.2.6), is associated to the stupid filtration, resp. Hodge filtration, on $\mathscr{K}^\cdot(\tilde{V})^{\mathrm{can}}$. A priori the left-hand side of (2.2.6) should be hypercohomology, but the differentials in $\mathrm{gr}_F \mathscr{K}^\cdot(\tilde{V})^{\mathrm{can}}$ are all trivial.

M. Saito's theory of mixed Hodge modules implies that the right-hand side of (2.2.5/6) possesses a functorial mixed Hodge structure **[Sa3]**, at least when V is defined over $\mathbb{Q}$; this is at least implicit in the constructions of **[LR]**. Moreover, as Zucker and Saito explained to me, Saito shows that the Hodge filtration, in this special case, is induced by the Hodge filtration on $\Omega^\cdot(\log(Z_\Sigma)) \otimes \tilde{V}^{\mathrm{can}}$ **[Sa2,** Remark 3.3]. Since $\mathscr{K}^\cdot(\tilde{V})^{\mathrm{can}}$ is filtered quasi-isomorphic to $\Omega^\cdot(\log(Z_\Sigma)) \otimes \tilde{V}^{\mathrm{can}}$, it follows formally (**[D2,** 8.1.9(v)]) that

2.2.7. THEOREM. (i) *The spectral sequences* (2.2.5) *and* (2.2.6) *degenerate at* E_1. *If* (ρ, V) *is defined over* $\mathbb{Q}$, *then the filtration defined by* (2.2.6) *is the Hodge filtration for a mixed Hodge structure on* $H^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla)$, *of weights* $\geq p+q+w_\rho$. *If* (ρ, V) *is defined over the number field* $E(V)$, *then* $H^\cdot(\mathrm{Sh}, \tilde{V}^\nabla)$ *only obtains an* $E(V)$*-mixed Hodge structure* (in particular, if $E(V)$ is not totally real, Hodge symmetry need not be satisfied, and $\mathrm{gr}^W_\cdot(H^\cdot(\mathrm{Sh}, \tilde{V}^\nabla)\otimes\mathbb{C})$ only has a complex Hodge structure, cf. **[Z1,** (4.6.)]).

(ii) *The natural spectral sequences* (cf. **[FC,** 5.5(ii); **H3,** 4.5])

(2.2.7.1)
$$E_1^{p,q} = \tilde{H}^q(\mathscr{K}^p(\tilde{V})^{\mathrm{sub}}) \Rightarrow H_c^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla) \left(\varinjlim_K H^{p+q}({}_K\mathrm{Sh}, \tilde{V}^\nabla)\right),$$

(2.2.7.2)
$$E_1^{p,q} = \tilde{H}^{p+q}(\mathrm{gr}_{F^\cdot}^p \mathscr{K}^\cdot(\tilde{V})^{\mathrm{sub}}) \Rightarrow H_c^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla)$$

degenerate at E_1 *and the filtration defined by* (2.2.7.2) *is the Hodge filtration for an* $E(V)$*-mixed Hodge structure on* $H_c^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla)$, *of weights* $\leq p+q+w_\rho$.

(iii) *Define the interior cohomology* $H_!^\cdot(\mathrm{Sh}, \tilde{V}^\nabla)$, *to be the image of the natural map* $H_c^\cdot(\mathrm{Sh}, \tilde{V}^\nabla) \to H^\cdot(\mathrm{Sh}, \tilde{V}^\nabla)$. *Then the natural spectral sequence*

(2.2.7.3)
$$E_1^{p,q} = H_!^{p+q}(\mathrm{gr}_{F^\cdot}^p \mathscr{K}^\cdot(\tilde{V})) \Rightarrow H_!^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla)$$

degenerates at E_1 *and the corresponding filtration is the Hodge filtration of a pure* $E(V)$*-Hodge structure on* $H_!^{p+q}(\mathrm{Sh}, \tilde{V}^\nabla)$ *of weight* $p+q+w_\rho$. *Over* $\mathbb{C}$, *this filtration splits canonically, giving rise to a natural Hodge decomposition*:

(2.2.7.4)
$$H_!^i(\mathrm{Sh}, \tilde{V}^\nabla)\otimes\mathbb{C} \cong \bigoplus H_!^{p,i+w_\rho-p}(\mathrm{Sh}, \tilde{V}),$$

where (cf. [**Z1,** 5.9. 5.29; **FC**, 5.5])

$$(2.2.7.5) \qquad H_!^{p,i+w_p-p}(\mathrm{Sh}, \tilde{V}) = \bigoplus_{w\in W^1, p(\Lambda,w)=p} H_!^{i-l(w)}(\mathscr{W}_{\tau(\Lambda,w)}).$$

This theorem was conjectured by Faltings and proved by him when G is anisotropic or Sh is the Siegel modular variety [**F, FC**]. In general it is a simple consequence of Saito's theory. More precisely, Saito's theory implies directly that (2.2.6) degenerates at E_1; but then, as pointed out in [**FC**, p. 234], the degeneration of (2.2.5) follows by dimension count, since $\mathrm{gr}_{F^\cdot}\mathscr{K}^\cdot(\tilde{V})^{\mathrm{can}}$ has trivial differentials. The results in (ii) follow from (i) by duality (cf. Remark 2.2.8 (iv), below), and the first part of (iii) follows immediately from (i) and (ii), since the natural morphism $\mathscr{K}^\cdot(\tilde{V})^{\mathrm{sub}} \to \mathscr{K}^\cdot(\tilde{V})^{\mathrm{can}}$ of complexes respects the Hodge filtration. Finally, the existence of the Hodge decomposition (2.2.7.5) follows from the harmonic theory discussed in [**H3**], just as in the compact quotient case [**Z1,** §5]. We remark that the statements regarding the Hodge filtration in [**H3,** 4.3] are incorrect.

2.2.8. REMARKS. (i) When $\tilde{V}$ is the trivial coefficient system, Theorem 2.2.7 reduces to Deligne's mixed Hodge theory for the open variety Sh.

(ii) When G is anisotropic, a substantially equivalent theorem was proved by Zucker in [**Z1**].

(iii) Saito's theory requires that $\tilde{V}$ be a *polarizable* variation of Hodge structures over Sh. If V is a real representation, then the existence of a polarization is shown in [**Z1,** p. 264 ff.]; thus $\tilde{V} \oplus \tilde{V}'$ is polarizable for any V, and the spectral sequence for $\tilde{V} \oplus \tilde{V}'$ is the direct sum of those for its two summands. In general, if $V = \bigoplus V_i$ is the sum of absolutely irreducible representations, let $\mathscr{K}^\cdot(\tilde{V}) = \bigoplus \mathscr{K}^\cdot(\tilde{V}_i)$; Theorem 2.2.7 remains true for such V.

(iv) Poincaré duality between $H^\cdot(\mathrm{Sh}, \tilde{V}^\nabla)$ and $H_c^\cdot(\mathrm{Sh}, \tilde{V}^{\nabla,*})$ is reflected by Serre duality between the E_1-terms of the spectral sequences above. In particular, if V is self-dual as a representation of G, so are the interior cohomology $H_!^\cdot(\mathrm{Sh}, \tilde{V}^\nabla)$ and $H_!^\cdot(\mathrm{gr}_{F^\cdot}\mathscr{K}^\cdot(\tilde{V}))$.

2.3. It is proved in [**H1**] that the functor $\mathscr{E} \mapsto [\mathscr{E}]$ from G-equivariant vector bundles on $\check{X}$ to $G(\mathbf{A}^f)$-equivariant vector bundles on Sh preserves rationality over the reflex field $E(G, X)$, and takes differential operators to differential operators. The functors $[\mathscr{E}] \mapsto [\mathscr{E}]^{\mathrm{can}}$ and $[\mathscr{E}] \mapsto [\mathscr{E}]^{\mathrm{sub}}$ preserve $E(G, X)$-rationality as well [**H2,** §4]. Finally, Milne [**Mi**] has proved the existence of *canonical models* for the automorphic vector bundles $[\mathscr{E}]$; more generally, for any $\sigma \in \mathrm{Aut}(\mathbb{C})$, he has shown that $[\mathscr{E}]^\sigma$ is canonically isomorphic to a specific automorphic vector bundle $[\mathscr{E}^{(\sigma)}]$ on the Shimura variety Sh^σ (the identification of Sh^σ as a Shimura variety is a conjecture of Langlands, proved by Milne and Borovoĭ; cf. [**Mi**]).

These methods extend without difficulty to prove that $[\mathscr{E}]^{\mathrm{can},\sigma} \cong [\mathscr{E}^{(\sigma)}]^{\mathrm{can}}$ (a sketch of this fact can be found in [**BHR**]). In any case, the published theorems have the following consequence:

2.3.1. Corollary. *Let* $RV = R_{E(V)/\mathbb{Q}}V$ *(Weil's restriction of scalars) as a representation of* G. *The dual BGG complex* $\mathscr{K}^{\bullet}(\widetilde{RV})^{\mathrm{can}}$ *has a canonical model over the field* $E(G, X)$. *Let* $H^{\bullet}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV})$ *be the* $E(G, X)$*-vector space* $\widetilde{\mathbb{H}}^{\bullet}(\mathscr{K}^{\bullet}(\widetilde{RV})^{\mathrm{can}}) := \varinjlim_K \mathbb{H}^{\bullet}({}_K\mathrm{Sh}_{\Sigma}, \mathscr{K}^{\bullet}(\widetilde{RV})_{\Sigma})$ *(any* Σ *will do for a given* K, *or take the limit over all* Σ), *with respect to this canonical model. For any embedding* $\sigma\colon E(G, X) \hookrightarrow \mathbb{C}$, *Let* $H^{\bullet}_{\sigma}(\mathrm{Sh}, \widetilde{RV})$ *be the* $\mathbb{Q}$*-vector space given by topological cohomology of* $\sigma(\mathrm{Sh})$ *with coefficients in the local system* $\widetilde{RV}$ *of* $\mathbb{Q}$*-vector spaces. View* $H^{\bullet}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV})$ *and* $H^{\bullet}_{\sigma}(\mathrm{Sh}, \widetilde{RV})$ *as* $E(V)$*-modules through the tautological action of* $E(V)$ *on* RV. *Then the pair* $(H^{\bullet}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV}), \{H^{\bullet}_{\sigma}(\mathrm{Sh}, \widetilde{RV})\})$, *together with the* $G(\mathbf{A}^f)$*-equivariant isomorphisms*

$$I_{\sigma}\colon H^{\bullet}_{\sigma}(\mathrm{Sh}, \widetilde{RV}) \otimes_{\mathbb{Q}} \mathbb{C} \xrightarrow{\sim} H^{\bullet}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV}) \otimes_{\sigma(E(G,X))} \mathbb{C},$$

is a mixed Hodge-de Rham structure $H^{\bullet}(\mathrm{Sh}, \widetilde{RV})$ *over* $E(G, X)$, *with coefficients in* $E(V)$. *We similarly define a mixed Hodge-de Rham structure* $H^{\bullet}_c(\mathrm{Sh}, \widetilde{RV})$ *and pure Hodge-de Rham structures* $H^i_!(\mathrm{Sh}, \widetilde{RV})$, $0 \leq i \leq 2d$, *all over* $E(G, X)$ *with coefficients in* $E(V)$, *such that* $H^{\bullet}_{c,\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV}) = \widetilde{\mathbb{H}}^{\bullet}(\mathscr{K}^{\bullet}(\widetilde{RV})^{\mathrm{sub}})$, $H^i_{!,\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV}) = \mathrm{Im}(\widetilde{\mathbb{H}}^{\bullet}(\mathscr{K}^{\bullet}(\widetilde{RV})^{\mathrm{sub}}) \to \widetilde{\mathbb{H}}^{\bullet}(\mathscr{K}^{\bullet}(\widetilde{RV})^{\mathrm{can}}))$, *etc.*

2.3.2. Remark. Everything is a simple consequence of Theorem 2.2.7 and the results recalled above from **[H1, Mi]**, with the exception of the assertion that the weight filtration of $H^{\bullet}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV}) \otimes_{\sigma(E(G,X))} \mathbb{C}$ is defined over $E(G, X)$. But this follows from the $E(G, X)$-rationality of the toroidal compactifications and of the $\mathscr{D}_{\mathrm{Sh}}$-module associated to $\widetilde{RV}$, and the functoriality of Saito's constructions.

Let π_f be an irreducible representation of $G(\mathbf{A}^f)$, realized on a vector space $\mathscr{V}_f$ over the field $E(\pi_f)$, which we assume to contain $E(V)$. Define the Hodge-de Rham structure $H(\pi_f; V)$ over $E(G, X)$, with coefficients in $E(\pi_f)$, by the prescription $H_?(\pi_f; V) = \mathrm{Hom}_{E(V)[G(\mathbf{A}^f)]}(\mathscr{V}_f, H^*_?(\mathrm{Sh}, \widetilde{RV}))$, $? = \mathrm{DR}$ or σ. We similarly define $H_c(\pi_f; V)$ and $H_!(\pi_f; V)$. These are all graded by cohomological degree, and may well be mixed (except for $H_!$). For the most common choices of π_f, however, these Hodge-de Rham structures will be pure and concentrated in cohomological degree d; cf. 3.2, below.

2.3.3. When $\mathrm{Sh}(G, X)$ is a Shimura variety of abelian type, as in Milne's talk in this volume, the mixed Hodge-de Rham structures $H^{\bullet}(\mathrm{Sh}, \widetilde{RV})$, $H^{\bullet}_c(\mathrm{Sh}, \widetilde{RV})$, $H^{\bullet}_!(\mathrm{Sh}, \widetilde{RV})$, $H(\pi_f; V)$, $H_c(\pi_f; V)$, and $H_!(\pi_f; V)$ can all be shown by Jannsen's method **[J**, Corollary 1.4**]** to be the Hodge-de Rham realizations of mixed motives for absolute Hodge cycles; in particular, we define the λ-adic realizations $H_{\lambda}(\pi_f; V)$, etc., whenever λ is a finite place of the coefficient field $E(\pi_f)$. At least when (G, X) has a symplectic embedding, this is completely clear: the local systems $\widetilde{RV}$ are direct summands, defined by absolute Hodge cycles, in the relative cohomology of abelian schemes $\mathscr{A}$ over $\mathrm{Sh}(G, X)$, so $H^\cdot(\mathrm{Sh}, \widetilde{RV})$ is itself realized as a piece of the cohomology of $\mathscr{A}$. The general case follows as in Milne's talk.

The *semisimplification* of $H_\lambda(\pi_f; V)$ as a Galois-module is denoted $H_\lambda(\pi_f; V)_{\mathrm{ss}}$; likewise for $H_c(\pi_f; V)$, etc.

2.3.4. REMARK. Let H be any of the mixed Hodge-de Rham structures defined above over the reflex field $E(G, X)$. Using Milne's results on conjugation of automorphic vector bundles, one can compute $R_{E(G,X)/\mathbb{Q}}H$ as a mixed Hodge-de Rham structure over $\mathbb{Q}$, with coefficients in $E(V)$, purely in terms of cohomology of Shimura varieties.

2.4. The *minimal*, or *Baily-Borel-Satake* compactification ${}_K\mathrm{Sh} \hookrightarrow {}_K\mathrm{Sh}^*$ is a normal projective variety which is canonically associated to ${}_K\mathrm{Sh}$; thus $\mathrm{Sh}^* = \varprojlim_K {}_K\mathrm{Sh}^*$ is well defined. In general, Sh^* is highly singular, and the middle intersection cohomology groups $IH^\cdot(\mathrm{Sh}^*, \widetilde{V})$ (cohomology of the extension of $\widetilde{V}$ to a pure perverse sheaf on Sh^*) have been the object of much attention; cf. Theorem 3.2.3. Suppose $\mathrm{Sh}(G, X)$ is of abelian type. We would like to show that $IH^\cdot(\mathrm{Sh}^*, \widetilde{RV})$ can again be regarded as an absolute Hodge motive. Fix a level subgroup K. With $\pi: \mathscr{A} \to {}_K\mathrm{Sh}$ as above, choose smooth projective toroidal compactifications $j_{\Sigma,\mathscr{A}}: \mathscr{A} \hookrightarrow \mathscr{A}_\Sigma$, $j_\Sigma: {}_K\mathrm{Sh} \hookrightarrow {}_K\mathrm{Sh}_\Sigma$, defined over $E(G, X)$, such that π extends to $\pi_\Sigma: \mathscr{A}_\Sigma \to {}_K\mathrm{Sh}_\Sigma$ (the existence of such $\mathscr{A}_\Sigma$ is proved by Pink **[P1]**). Let $i_\Sigma: {}_K\mathrm{Sh}_\Sigma \to {}_K\mathrm{Sh}^*$ be the canonical map. By the decomposition theorem of Beilinson-Bernstein-Deligne-Gabber for pure perverse sheaves, $IH^\cdot(\mathrm{Sh}^*, \widetilde{RV})$ can be realized as a direct summand of $H^\cdot(\mathscr{A}_\Sigma)$ **[BBD, Sa1, Sa3]**. This is true both in ℓ-adic cohomology and in Betti cohomology; and if $\mathscr{A}_\Sigma$ is given with a polarization, Saito shows that $IH^\cdot(\mathrm{Sh}^*, \widetilde{RV})$ is canonically a polarized sub-Hodge structure of $H^\cdot(\mathscr{A}_\Sigma)$. (Here we have to use that $\widetilde{RV}$ has already been realized as a polarized sub-VHS of $R^\cdot\pi_*\mathbb{Q}_{\mathscr{A}}$.)

Furthermore, the ℓ-adic and Betti embeddings are compatible with respect to the comparison isomorphisms. In order to identify $IH^\cdot(\mathrm{SH}^\cdot, \widetilde{RV})$ with an absolute Hodge submotive of $H^\cdot(\mathscr{A}_\Sigma)$, we must show that the direct summand $IH^\cdot(\mathrm{Sh}^*, \widetilde{RV}) \subset H^\cdot(\mathscr{A}_\Sigma)$ is $E(G, X)$-rational for the de Rham rational structure on $H^\cdot(\mathscr{A}_\Sigma)$. Brylinski has explained how to do this (private communication) using the Riemann-Hilbert correspondence. Briefly, the point is that the de Rham rational structure can be computed using complexes of $\mathscr{D}$-modules, which can be defined over $E(G, X)$. Brylinski points out that this method shows in general that intersection cohomology with coefficients in a VHS of geometric origin can be realized as an absolute Hodge motive over an appropriate field of definition, and that the full strength of the decomposition theorem is not needed; it suffices to use hard Lefschetz to realize (e.g.) $IH^{\cdot}(\mathrm{Sh}^*, \widetilde{RV})$ as a direct summand of $H^{\cdot}(\mathscr{A}_\Sigma)$.

When the coefficient system V is trivial, the question of realizing $IH^{\cdot}(\mathrm{Sh}^*, \widetilde{RV})$ as a Grothendieck submotive of $H^{\cdot}({}_K\mathrm{Sh}_\Sigma)$ has been raised by Ramakrishnan in **[Ra]**.

3. Automorphic forms as differential forms

For any reductive group G over $\mathbb{Q}$, let $\mathfrak{g} = \mathrm{Lie}(G)_{\mathbb{C}}$, $\mathfrak{z}_G$ denote the center of the enveloping algebra $U(\mathfrak{g})$, and denote by Z_G the center of G. Choose maximal compact subgroups $K_\infty \subset G(\mathbb{R})$, $K_f \subset G(\mathbf{A}^f)$. Let $\mathscr{A}(G)$ denote the space of automorphic forms on $G(\mathbb{Q})\backslash G(\mathbf{A})$, $\mathscr{A}_0(G)$ the space of cusp forms, and $\mathscr{A}_{(2)}(G)$ the space of square integrable automorphic forms. Thus every element of $\mathscr{A}(G)$ is assumed to be slowly increasing, in Harish-Chandra's sense (cf. Ramakrishnan's lecture), and to be $\mathfrak{z}_G$- (resp. K_∞-, resp. K_f-) finite; i.e., to generate a finite-dimensional representation under right translation by $\mathfrak{z}_G$ (resp. K_∞, resp. K_f). We also assume our automorphic forms to be $Z_G(\mathbf{A})$-finite. We have the inclusions $\mathscr{A}(G) \supset \mathscr{A}_{(2)}(G) \supset \mathscr{A}_0(G)$. Let $\mathscr{A}_{\mathrm{si}}(G)$ be the space of slowly increasing $K_\infty \times K_f \times Z_G(\mathbf{A})$-finite functions on $G(\mathbb{Q})\backslash G(\mathbf{A})$.

Let $X(G) = \varprojlim_K G(\mathbb{Q})\backslash G(\mathbf{A})/Z_G(\mathbf{R})\cdot K_\infty \cdot K$, where K runs through open compact subgroups of $G(\mathbf{A}^f)$, be the adèlic Riemmannian locally symmetric space associated to G. If (ρ, V) is a finite-dimensional representation of G, let $\widetilde{V}^\nabla$ be the corresponding local coefficient system over $X(G)$, defined as in 2.1. If G is of hermitian type, then $X(G)$ is the Riemannian manifold underlying the Shimura variety $\mathrm{Sh}(G, X)(\mathbb{C})$. In this case we assume $K_\infty = K_h$ and define $\mathfrak{P}_h$ as in 2.1.

J. Franke has recently proved the following theorem, which had been conjectured by Borel:

3.1. THEOREM **[Fr]**. (i) *For any G, there is a canonical isomorphism*

$$H^{\cdot}(X(G), \widetilde{V}^\nabla(\mathbb{C})) \left(= \varinjlim_K H^{\cdot}(G(\mathbb{Q})\backslash G(\mathbf{A})/Z_G(\mathbb{R}) \cdot K_\infty \cdot K, \widetilde{V}^\nabla(\mathbb{C})) \right)$$
$$\xrightarrow{\sim} H^{\cdot}(\mathfrak{g}, K_\infty; V \otimes \mathscr{A}(G)),$$

where the right-hand side is relative Lie algebra cohomology.

(ii) *If G is of hermitian type and $\mathscr{W}_\tau$ is the automorphic vector bundle associated to the representation (τ, W_τ) of K_h (cf. 2.1), there is a canonical isomorphism*

$$\widetilde{H}^{\cdot}(\mathscr{W}_\tau^{\mathrm{can}}(\mathbb{C})) \xrightarrow{\sim} H^{\cdot}(\mathfrak{P}_h, K_h; W_\tau \otimes \mathscr{A}(G)).$$

Here $\widetilde{V}^\nabla(\mathbb{C})$ (resp. $\mathscr{W}_\tau^{\mathrm{can}}(\mathbb{C})$) are taken to be defined by complex-valued representations of G (resp. K_h).

The standard references for Lie algebra cohomology include **[BW, V]**. For G anisotropic ($X(G)$ compact), the results basically go back to work of

Matsushima-Murakami **[MM]**: the right-hand side of (i) (resp. (ii)) is naturally identified with the global sections of the de Rham complex on $X(G)$ with coefficients in $\widetilde{V}$ (resp. the Dolbeault complex on Sh with coefficients in $\mathscr{W}_\tau$). In the noncompact case, the first steps were taken by Borel **[Bo]**, who proved (i) with $\mathscr{A}(G)$ replaced by $\mathscr{A}_{\mathrm{si}}(G)$. The corresponding result for the coherent cohomology spaces in (ii) was obtained in **[H4, H3]**. Franke's contribution, which shows that the cohomology is computed by $\mathfrak{z}_G$-finite forms, is considerably more difficult.

Comparing Corollary 2.3.1 with Theorem 3.1, it is natural to ask to what extent the data associated with the HdR structure $H^{\cdot}(\mathrm{Sh}, \widetilde{RV})$ can be read off in terms of automorphic forms. We consider these data in turn:

3.1.1. The *Betti rational structures* $H^{\cdot}_{\sigma}(\mathrm{Sh}, \widetilde{RV})$ are theoretically completely accessible in terms of automorphic forms, but in practice it is difficult to construct enough rational cycles. The situation is quite well understood for GL(2) **[Ei, Sh1, Sh5, W2]**. For other groups our knowledge is much less satisfactory. We return to this point in §5.

3.1.2. Theorem 3.1 makes no direct reference to the *de Rham rational structure* $H^{\cdot}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV})$. Instead, the isomorphisms (ii) identify the direct summands of $\mathrm{gr}^{\cdot}_F H^{\cdot}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV})$ with spaces of automorphic forms. In other words, automorphic forms provide access, at least in principle, to the $H^{p,q}$-structures over $E(G, X)$ (1.1) attached to $H^{\cdot}_{\mathrm{DR}}(\mathrm{Sh}, \widetilde{RV})$. These are direct sums of spaces of coherent cohomology, and the problem in general is to read the rational structure on the left-hand side of 3.1(ii) in terms of something on the right-hand side.

The automorphic forms corresponding to the spaces $\widetilde{H}^0(\mathscr{W}_\tau^{\mathrm{can}})$ are *holomorphic*. In this case the problem has been completely solved, using ideas of Shimura **[H1]**. The rationality of a vector-valued holomorphic automorphic form $f\colon G(\mathbb{Q})\backslash G(\mathbf{A}) \to W_\tau(\mathbb{C})$, representing a section $s(f) \in \Gamma({}_K\mathrm{Sh}_\Sigma, \mathscr{W}_\tau^{\mathrm{can}})$, can be completely determined by evaluating f on the adelic points of subgroups $H \subset G$ corresponding to *zero-dimensional* Shimura subvarieties $\mathrm{Sh}(H, h) \subset \mathrm{Sh}(G, X)$ ($h \in X$ a point fixed by $H(\mathbb{R})$). In this way the problem is reduced to the problem for $\dim(\mathrm{Sh}) = 0$. But zero-dimensional Shimura varieties tend to be moduli spaces for CM motives (cf. Milne's talk). Thus for given $\mathrm{Sh}(H, h)$ and τ, there exist matrices $p(h, \tau)$ of periods of CM motives such that

(3.1.2.1)
$$s(f) \text{ is rational over } \overline{\mathbb{Q}} \text{ if and only if } p(h, \tau)^{-1} \cdot f(h, \gamma) \in W_\tau(\overline{\mathbb{Q}}) \text{ for all } \gamma \in G(\mathbf{A}^f).$$

By varying h, one can also descend to the reflex field. For more details, cf. **[H6,** §§1, 2]. If $\mathrm{Sh}(G, X)$ is noncompact, one can also use special values of Fourier-Jacobi series, cf. **[H1,** §6; **P1]**, as well as earlier work of Shimura, Garrett, and Brylinski cited in **[H1]**. This criterion generalizes the familiar *q-expansion principle* for elliptic modular forms **[DR]**; cf. §5.1 below.

Rationality criteria for forms in $\widetilde{H}^q(\mathscr{W}_\tau^{\mathrm{can}})$ when $q > 0$ are discussed in **[H4,** §5]. The criterion is based on restriction to Shimura subvarieties of $\mathrm{Sh}(G, X)$ of dimension $\geq q$, and is only effective under rather restrictive hypotheses.

3.1.3. When $E(G, X)$ is totally real, the *Archimedean Frobenius* involutions F_σ on $H^{\cdot}_\sigma(\mathrm{Sh}, \widetilde{RV})$, viewed as automorphisms of spaces of Lie algebra cohomology, should be determined by a formula defined by Langlands (p. 239, last line of **[L]**; however, there appears to be a misprint in the formula as stated). As far as I know this has not been checked except in certain cases.

3.1.4. Finally, the realization of the *weight filtration* in terms of automorphic forms remains mysterious. Special cases are described in **[Ha3]**. Some information is provided by the results discussed in §3.5 below.

3.2. When G is of hermitian type we write $\mathrm{Sh} = X(G)$, as before. Let

$$H^{\cdot}_{\mathrm{cusp}}(X(G), \widetilde{V}^{\nabla}) = H^{\cdot}(\mathfrak{g}, K_\infty; V \otimes \mathscr{A}_0(G)),$$
$$H^{\cdot}_{(2)}(X(G), \widetilde{V}^{\nabla}) = H^{\cdot}(\mathfrak{g}, K_\infty; V \otimes \mathscr{A}_{(2)}(G)),$$
$$H^{\cdot}_{\mathrm{cusp}}(\mathscr{W}_\tau) = H^{\cdot}(\mathfrak{P}_h, K_h; W_\tau \otimes \mathscr{A}_0(G)),$$
$$H^{\cdot}_{(2)}(\mathscr{W}_\tau) = H^{\cdot}(\mathfrak{P}_h, K_h; W_\tau \otimes \mathscr{A}_{(2)}(G)).$$

There are natural *injections*

$$(3.2.1)\quad H^{\cdot}_{\mathrm{cusp}}(X(G), \widetilde{V}^{\nabla}) \subset H^{\cdot}_!(X(G), \widetilde{V}^{\nabla}(\mathbb{C})) \subset H^{\cdot}_{(2)}(X(G), \widetilde{V}^{\nabla});$$

$$(3.2.2)\quad H^{\cdot}_{\mathrm{cusp}}(\mathscr{W}_\tau) \subset H^{\cdot}_!(\mathscr{W}_\tau^{\mathrm{can}}(\mathbb{C})) \subset H^{\cdot}_{(2)}(\mathscr{W}_\tau)$$

(cf. **[Schw,** 4.2], **[H3,** Theorem 2.7]). The *Zucker conjecture* **[Z2]**, proved independently by Looijenga and Saper-Stern, using two different methods, and reproved by Looijenga-Rapport using a third method, is the theorem:

3.2.3. THEOREM **[Lo, SS, LR]**. *There is a canonical isomorphism*

$$H^{\cdot}_{(2)}((\mathrm{Sh}, \widetilde{V}^{\nabla}) \xrightarrow{\sim} IH^{\cdot}(\mathrm{Sh}^*, \widetilde{V}) \qquad (\textit{notation } 2.4).$$

3.2.4. We say V, resp. W_τ, is *regular*, if the highest weight of V is regular, resp. if $\tau + \rho$ is regular, relative to the root system R^+ (ρ and R^+ as in 2.2). As Franke observes **[Fr,** Theorem 7.4.2, II], if V is regular then the inclusions (3.2.1) are *isomorphisms*; if W_τ is regular then the maps in (3.2.2) are similarly seen to be isomorphisms. Write

(3.2.4.1)
$$\mathscr{A}_0(G) = \bigoplus_\pi m_0(\pi)\pi, \qquad \mathscr{A}_{(2)}(G) = \bigoplus_\pi m_{(2)}(\pi), \pi, \qquad \pi \cong \pi_\infty \otimes \pi_f,$$

where π runs through the family of unitary representations $\{\pi_\infty \otimes \pi_f\}$ of $(\mathfrak{g}, K_\infty) \times G(\mathbf{A}^f)$ (for conventions regarding representations of adèlic groups cf. the articles of Knapp-Kudla and Ramakrishnan), and $m_0(\pi)$ and $m_{(2)}(\pi)$ are nonnegative integers, called the (cuspidal, resp. L_2) *multiplicity* of π. Then

$$(3.2.4.2)\quad H^{\cdot}_{\mathrm{cusp}}(X(G), \tilde{V}^{\nabla}) \cong \bigoplus_{\pi} m_0(\pi) H^{\cdot}(\mathfrak{g}, K_\infty; V \otimes \pi_\infty) \otimes \pi_f;$$

$$(3.2.4.3)\qquad H^{\cdot}_{\mathrm{cusp}}(\mathscr{W}_\tau) \cong \bigoplus_{\pi} m_0(\pi) H^{\cdot}(\mathfrak{P}_h, K_h; W_\tau \otimes \pi_\infty) \otimes \pi_f.$$

This reduces the computation of the left-hand side to the determination of the multiplicities $m_0(\pi)$ and the computation of $H^{\cdot}(\mathfrak{g}, K_\infty; V \otimes \pi_\infty)$ for all unitary representations (or $(\mathfrak{g}, K_\infty)$-modules) π_∞. The analogous formulas hold for $H^{\cdot}_{(2)}$. The problem of multiplicities is discussed in **[BR]**; much is known and much remains to be done. But the spaces $H^{\cdot}(\mathfrak{g}, K_\infty; V \otimes \pi_\infty)$ are completely understood, thanks to a long series of contributions culminating in the article **[VZ]** of Vogan and Zuckerman.

Assume G to be of hermitian type, and write $K_h = K_\infty$. If V is regular, then $H^{\cdot}(\mathfrak{g}, K_h; V \otimes \pi_\infty) = 0$ for π_∞ nontempered (cf. the proof of **[Fr,** loc. cit.]), and the complete description of $H^{\cdot}(\mathfrak{g}, K_h; V \otimes \pi_\infty)$ is very simple. Recall the definition of $W^1 \subset W(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h})$ from 2.2. Let $\Lambda \in \mathfrak{h}^*$ be the highest weight of V as before, and let $\lambda = \Lambda + \rho$. Then

(3.2.4.4) For any V, i, and any unitary representation π_∞, there is a canonical isomorphism (Hodge decomposition)

$$H^i(\mathfrak{g}, K_h; V \otimes \pi_\infty) \cong \bigoplus_{w \in W^1} H^{i-l(w)}(\mathfrak{P}_h, K_h; W_{\tau(\Lambda, w)} \otimes \pi_\infty).$$

(3.2.4.5) For each $w \in W^1$, there is a unique *discrete series* representation $\pi_{w(\lambda)}$ of $G(\mathbb{R})$ (or $(\mathfrak{g}, K_\infty)$); it is tempered and even has square-integrable matrix coefficients (modulo $Z_G(\mathbb{R})$), and every square-integrable representation of $G(\mathbb{R})$ has this form for some (not necessarily regular) V. The character $w(\lambda) \in \mathfrak{h}^*$ is the *Harish-Chandra parameter* of $\pi_{w(\lambda)}$.

(3.2.4.6) If V is regular, then $H^1(\mathfrak{g}, K_h; V \otimes \pi_\infty) = 0$ unless $i = d$ and $\pi_\infty \cong (\pi_{w(\lambda)})^*$ for some $w \in W^1$. Moreover,

$$H^d(\mathfrak{g}, K_h; V \otimes (\pi_{w(\lambda)})^*) \cong H^{d-l(w)}(\mathfrak{P}_h, K_h; W_{\tau(\Lambda, w)} \otimes (\pi_{w(\lambda)})^*)$$

is of dimension 1 for all $w \in W^1$. The set $\{(\pi_{w(\lambda)})^*\}$ is called the *discrete series L-packet* associated to V (or to λ).

References for these facts can be found in **[H3,** §§3–4].

3.3. The Hodge decomposition (3.2.4.4) induces similar decompositions on automorphic cohomology. For ? = cusp or (2), define

$$(3.3.1)\qquad H^{p,q}_{?}(\mathrm{Sh}, \tilde{V}^{\nabla}) = \bigoplus_{w \in W^1, p(\Lambda, w) = p} H^{i-l(w)}_{?}(\mathscr{W}_{\tau(\Lambda, w)}).$$

Then

$$H^i_?(\mathrm{Sh}, \tilde{V}^\nabla) \cong \bigoplus_{p+q=i+w_p} H^{p,q}_?(\mathrm{Sh}, \tilde{V}^\nabla), \tag{3.3.2}$$

via the isomorphisms (3.2.4.2), (3.2.4.3); call this the *analytic* Hodge decomposition. If V is regular, then the left-hand side of (3.3.2) coincides with interior cohomology for $? =$ cusp or (2), as remarked above. It follows from **[H3,** Theorem 4.5] that, in this case, the Hodge decomposition (3.3.1) coincides with the Hodge decomposition (2.2.7.4) on interior cohomology, which we call the *geometric* Hodge decomposition. More generally, the inclusion $H^\cdot_!(\mathrm{Sh}, \tilde{V}^\nabla(\mathbb{C})) \subset H^i_{(2)}(\mathrm{Sh}, \tilde{V}^\nabla)$ of (3.2.1) induces an analytic Hodge decomposition on the former, which again coincides with the geometric one. The following well-known conjecture remains open.

3.3.3. CONJECTURE. The analytic Hodge decomposition on $H^i_{(2)}(\mathrm{Sh}, \tilde{V}^\nabla)$ induces Saito's geometric Hodge decomposition **[Sa1, Sa3]** on $IH^\cdot(\mathrm{Sh}^*, \tilde{V})$ via the isomorphism of Theorem 3.2.3.

The analytic Hodge decomposition can be obtained in another way. In (2.1.1) write $\mathfrak{p}_\mathbb{C} = \mathfrak{p}^+ \oplus \mathfrak{p}^-$; thus $\mathfrak{g}_\mathbb{C} = \mathfrak{k}_{h,\mathbb{C}} \oplus \mathfrak{p}_\mathbb{C}$ is the complexification of the Cartan decomposition. For any $(\mathfrak{g}, K_h)$-module $\mathscr{V}$, the standard complex computing $H^\cdot(\mathfrak{g}, K_h; V \otimes \mathscr{V})$ is

$$C^\cdot(\mathfrak{g}, K_h; V \otimes \mathscr{V}) = [(\Lambda^\cdot \mathfrak{p}^*_\mathbb{C}) \otimes V \otimes \mathscr{V})]^{K_h},$$

the K_h-invariants in the tensor products, with differential given by **[BW,** II, 1.4(2)]. Write $V = \bigoplus V^{r,s}$, the Hodge decomposition corresponding to the point $h \in X$, as in §2.2 and Milne's talk. Let

$$C^{a,b;r,s}(\mathfrak{g}, K_h; V \otimes \mathscr{V}) = (\Lambda^a \mathfrak{p}^{+,*} \oplus \Lambda^b \mathfrak{p}^{-,*} \otimes V^{r,s} \otimes \mathscr{V})^{K_h};$$

then for all i,

$$C^i(\mathfrak{g}, K_h; V \otimes \mathscr{V}) = \bigoplus_{a+b=i} \bigoplus_{r,s} C^{a,b;r,s}(\mathfrak{g}, K_h; V \otimes \mathscr{V}). \tag{3.3.4}$$

The right-hand side is bigraded: let

$$C^{p,q}(\mathfrak{g}, K_h; V \otimes \mathscr{V}) = \bigoplus_{a+r=p} \bigoplus_{b+s=q} C^{a,b;r,s}(\mathfrak{g}, K_h; V \otimes \mathscr{V}). \tag{3.3.5}$$

Now let $\mathscr{V} = \mathscr{A}_{(2)}(G)$. If $\mathscr{W}$ is any irreducible $(\mathfrak{g}, K_h)$-module, the center $\mathfrak{z}_G$ of $U(\mathfrak{g})$ acts on $\mathscr{W}$ by a character $\chi_\mathscr{W}: \mathfrak{z}_G \to \mathbb{C}$, the infinitesimal character of $\mathscr{W}$. The harmonic theory **[BW,** II] has the following consequences. Let

$$\mathscr{A}_?(G)_V = \bigoplus_{\chi_{\pi_\infty} = \chi_V} m_?(\pi)\pi \subset \mathscr{A}_?(G), \qquad ? = (2),\ 0.$$

Then there are canonical isomorphisms (cf. **[BW,** II, 3.1]):

$$H^\cdot(\mathfrak{g}, K_h; V \otimes \mathscr{A}_?(G)) \cong C^\cdot(\mathfrak{g}, K_h; V \otimes \mathscr{A}_?(G)_V), \qquad ? = (2),\ 0. \tag{3.3.6}$$

Then $C^{p,q}(\mathfrak{g}, K_h; V \otimes \mathscr{A}_{(2)}(G)_V)$, defined as in (3.3.5), corresponds to a direct summand

$$H^{p,q}(\mathfrak{g}, K_h; V \otimes \mathscr{A}_{(2)}(G)) \subset H^{\cdot}(\mathfrak{g}, K_h; V \otimes \mathscr{A}_{(2)}(G)) \cong H^{\cdot}_{(2)}(\mathrm{Sh}, \widetilde{V}^{\nabla}),$$

and the grading on $H^{\cdot}_{(2)}(\mathrm{Sh}, \widetilde{V}^{\nabla})$ by degree induces a natural grading on $H^{p,q}(\mathfrak{g}, K_h; V \otimes \mathscr{A}_{(2)}(G))$. Let $H^{p,q}_{(2)}(\mathrm{Sh}, \widetilde{V}) \subset H^{\cdot}_{(2)}(\mathrm{Sh}, \widetilde{V}^{\nabla})$ be the corresponding direct summand, and let $H^{p,q}_{!}(\mathrm{Sh}, \widetilde{V}) = H^{p,q}_{(2)}(\mathrm{Sh}, \widetilde{V}) \cap H^{\cdot}_{!}(\mathrm{Sh}, \widetilde{V}^{\nabla})$. I claim that the decomposition

$$H^{i}_{!}(\mathrm{Sh}, \widetilde{V}^{\nabla}) = \bigoplus_{p+q=i+w_p} H^{p,q}_{!}(\mathrm{Sh}, \widetilde{V}) \tag{3.3.7}$$

equals the analytic Hodge decomposition on interior cohomology. To verify that this is consistent with (3.3.2), we note that, for any $\mathscr{V}$,

$$\begin{aligned}(\Lambda^a \mathfrak{p}^{+,*} \otimes \Lambda^b \mathfrak{p}^{-,*} \otimes V \otimes \mathscr{V})^{K_h} &= (\Lambda^b \mathfrak{p}^{-,*} \otimes C^a(\mathfrak{p}^+, V) \otimes \mathscr{V})^{K_h} \\ &= C^b(\mathfrak{P}_h, K_h; C^a(\mathfrak{p}^+, V) \otimes \mathscr{V}),\end{aligned}$$

where $C^{\cdot}(\mathfrak{p}^+, V)$ is the standard complex computing $H^{\cdot}(\mathfrak{p}^+, V)$. We thus obtain a spectral sequence

$$E_2^{a,b} = H^b(\mathfrak{P}_h, K_h; H^a(\mathfrak{p}^+, V) \otimes \mathscr{V}) \Rightarrow H^{a+b}(\mathfrak{g}, K_h; V \otimes \mathscr{V})). \tag{3.3.8}$$

As Zucker observes **[Z1,** (5.32)], Kostant's formula (cf. **[BW,** III, 3.1]) shows that the E_2-term is exactly

$$\bigoplus_{w \in W^1, l(w)=a} H^b(\mathfrak{P}_h, K_h; W_{\tau(\Lambda, w)} \otimes \mathscr{V}),$$

from which our claim follows immediately.

When Sh is projective, Zucker's arguments in **[Z1,** §5] show that (3.3.7) coincides with the geometric Hodge decomposition (2.2.7.4). Using our concrete description (Theorem 2.2.7) of Saito's Hodge filtration on $H^{\cdot}_{!}(\mathrm{Sh}, \widetilde{V}^{\nabla})$, the arguments in [loc. cit] extend immediately, and we again see that the geometric and analytic Hodge decompositions coincide. In summary:

3.3.9. PROPOSITION. *The decompositions* (3.3.7), (3.3.2), *and* (2.2.7.4) *of* $H^{i}_{!}(\mathrm{Sh}, \widetilde{V}^{\nabla})$ *coincide.*

Recall that if V is not defined over a totally real field, these decompositions need not satisfy Hodge symmetry.

3.3.10. Now the Hodge numbering (3.3.5) on $C^{\cdot}(\mathfrak{g}, K_h; V \otimes \mathscr{A}_{(2)}(G)_V)$ is the one defined by Langlands in **[L,** p. 239]. The theory of the L-group defines yet another Hodge numbering on $C^{\cdot}(\mathfrak{g}, K_h; V \otimes \mathscr{A}_{(2)}(G)_V)$, this one coming from the Langlands classification. Let $r = r_{(G,X)}$ be the representation of the L-group L_G associated to $\mathrm{Sh}(G, X)$, as in **[BR]**; recall that

$\dim(r) = |W^1|$. Let $W_{\mathbb{C}/\mathbb{R}}$ be the Weil group of $\mathbb{C}/\mathbb{R}$ (cf. Steenbrink's lecture), and let $\alpha\colon W_{\mathbb{C}/\mathbb{R}} \to \mathbb{R}^\times$ be the norm from $W_{\mathbb{C}/\mathbb{R}} \to \mathbb{R}^\times$, composed with the absolute value. The discrete series L-packet $\{(\pi_{w(\lambda)})^*\}$ of (3.2.4.6) corresponds, via the Langlands classification, to a conjugacy class of homomorphisms $\varphi_\lambda\colon W_{\mathbb{C}/\mathbb{R}} \to {}^LG$ (cf. **[loc. cit.]**). Then $\alpha^{-d/2} \cdot r \circ \varphi_\lambda|_{\mathbb{C}^\times}$ is an algebraic character of $\mathbb{C}^\times$, and defines a (complex) Hodge structure equivalent to that on $\bigoplus_{w\in W^1} H^d(\mathfrak{g}, K_h; V \otimes (\pi_{w(\lambda)})^*)$, given by (3.3.5) [**L**, p. 240, Lemma].

More generally, to every conjugacy class of homomorphisms $\psi\colon W_{\mathbb{C}/\mathbb{R}} \times \mathrm{SL}(2, \mathbb{C}) \to L_G$, whose restriction to $W_{\mathbb{C}/\mathbb{R}}$ has bounded image, Arthur has conjectured the existence of a packet Π_ψ of irreducible $(\mathfrak{g}, K_h)$-modules [**A**], satisfying certain character identities. The conjecture has been verified by Adams and Johnson for representations with cohomology [**AJ**]. Identifying $r \circ \psi(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}) \in \mathrm{SL}(2, \mathbb{C})$ with the Lefschetz operator L on cohomology, Arthur has shown that the Hodge structure (3.3.7) corresponds to one naturally associated to $r \circ \psi$ [**A**, Proposition 9.1].

The description of Arthur's Hodge structure is omitted; we simply mention that Arthur's normalization of the weights differs slightly from ours. The significance of this identity of Hodge structures should be emphasized, however. The geometric Hodge structure (2.2.7.4) is the one used by Deninger in his lecture to define the Γ-factors for the corresponding motivic L-functions. The Hodge structures attached by Langlands and Arthur to π_∞ are the ones to define the Archimedean Euler factors associated to π_∞ and the representation r of ${}^LG$. In light of the above discussion, Conjecture 3.3.3 asserts that these Hodge structures coincide. We have seen that this is true at least for π_∞ contributing to interior cohomology, and interior cohomology coincides with L_2-cohomology when the coefficient system is regular (cf. §3.2.4 above). This is the Archimedean counterpart of the (mostly conjectural) formulas for the zeta functions of Shimura varieties, discussed in [**BR**]. Arthur says as much in [A], but (as far as I know) the relation between the analytic and geometric Hodge structures has not previously been verified, except when Sh is compact. Of course, the identification of Saito's Hodge structures with motivic Hodge structures remains to be established, but at least for Shimura varieties of abelian type Saito's structures are motivic (for absolute Hodge motives; cf. §§2.3.3 and 2.4).

3.4. ***L*-groups and Mumford-Tate groups.** The pure Hodge structure on the d-dimensional cohomology of the discrete series L-packet

$$H^d((G, X); V) := \bigoplus_{w\in W^1} H^d(\mathfrak{g}, K_h; V \otimes (\pi_{w(\lambda)})^*), \tag{3.4.1}$$

discussed above, is of special importance for the cohomology of Shimura varieties. As K runs through the set of open compact subgroups of $G(\mathbf{A}^f)$, let $m_{0,K}(\pi_{w(\lambda)})$ denote the multiplicity of $\pi_{w(\lambda)}$ in the space $\mathscr{A}_0(G)^K$ of K-fixed vectors in $\mathscr{A}_0(G)$. Then the limit $\operatorname{vol}(K)\cdot m_{0,K}(\pi_{w(\lambda)})$, as $\operatorname{vol}(K)\to 0$, tends to a number depending only on the L-packet, and not on the choice of $\pi_{w(\lambda)}$ **[Sv]**. This suggests that, for any representation π_f of $G(\mathbf{A}^f)$, the HdR structure $H(\pi_f, V)$, discussed in 2.3, will "in general" have ($E(V)$-) Hodge structure isomorphic to $H^d((G, X); V)$. As **[BR]** indicates, the actual state of affairs is much more complicated, but the stable trace formula is expected to reduce the entire intersection cohomology $IH^{\cdot}(\mathrm{Sh}^*, \widetilde{V})$ to Hodge structures which are "general" in the above sense, possibly on smaller groups. The cohomology of the varieties discussed in §4 is known to be stable in this sense.

Recall that the morphism $\underline{S} = R_{\mathbf{C}/\mathbf{R}}\mathbb{G}_{m,\mathbf{C}} \to \mathrm{GL}(H^d((G, X); V))$, defining the Hodge structure on (3.4.1), factors (up to a twist by $\alpha^{-d/2}$) through the representation $r_{(G,X)}$ of ${}^LG$ on the latter space. This suggests that ${}^LG$ should be viewed as the generic Mumford-Tate group of motives appearing in $IH^d(\mathrm{Sh}^*, \widetilde{V})$ (or in $H^d_!(\mathrm{Sh}, \widetilde{V})$).

3.5. Mixed motives and boundary cohomology. The material in this section represents work in progress with S. Zucker. As explained at the end of the previous section, the Hodge structures of pure motives with (conjecturally) appear in the cohomology of Shimura varieties are related to the Langlands-Arthur parameters of automorphic representations by Conjecture 3.3.3. No conjectural description has yet been proposed for the mixed Hodge structures of 2.2.7 (i), (ii); see **[Ha3]** for some results in that direction.

It is more likely that a purely group-theoretic expression can be found for the Hodge structure of the *semisimplified* motive $\mathrm{gr}^W_{\cdot} H^{\cdot}(\mathrm{Sh}, \widetilde{V}^{\nabla})$. Since $H^{\cdot}_!(\mathrm{Sh}, \widetilde{V}^{\nabla})$ is pure, it suffices to consider $H^{\cdot}(\mathrm{Sh}, \widetilde{V}^{\nabla})/H^{\cdot}_!(\mathrm{Sh}, \widetilde{V}^{\nabla})$. This is still rather subtle, and it is easier to study the boundary cohomology

$$(3.5.1)\qquad H^{\cdot}(\partial\,\mathrm{Sh}, \widetilde{V}^{\nabla}) := \varinjlim_{K} H^{p+q}({}_K\mathrm{Sh}_{\Sigma}, \mathrm{Cone}\{j_{\Sigma,!}\widetilde{V} \to \mathbf{R}j_{\Sigma,*}\widetilde{V}\}).$$

In the preceding formula, ${}_K\mathrm{Sh}_{\Sigma}$ can be replaced by any topological compactification, for example the Borel-Serre compactification ${}_K\overline{\mathrm{Sh}}$ as a C^{∞} manifold with corners **[BS]**, or the Baily-Borel-Satake compactification ${}_K\mathrm{Sh}^*$. Unlike the torodial compactifications, both of these are canonical, hence pass to the limit over K to define compactifications $\overline{\mathrm{Sh}}$ and Sh^*. Moreover, Sh is homotopy equivalent to $\overline{\mathrm{Sh}}$; let $\overline{V}$ be an extension of $\widetilde{V}^{\nabla}$ to a local system on $\overline{\mathrm{Sh}}$ compatible with the retraction. Thus, letting $\partial\overline{\mathrm{Sh}} := \overline{\mathrm{Sh}} - \mathrm{Sh}$ and $\bar{i}\colon \partial\overline{\mathrm{Sh}} \hookrightarrow \overline{\mathrm{Sh}}$ be the closed immersion, $H^{\cdot}(\partial\,\mathrm{Sh}, \widetilde{V}^{\nabla})$ can be realized concretely as $H^{\cdot}(\partial\overline{\mathrm{Sh}}, \bar{i}^*(\overline{V}))$. Moreover, $\partial\,\overline{\mathrm{Sh}}$ has a concrete description in terms of nilmanifold fibrations over locally symmetric spaces of smaller dimension. Concretely, let $\mathscr{P}$ (resp. $\mathscr{P}^{\max}$) be the set of conjugacy classes of rational parabolic (resp. maximal parabolic) subgroups of G; by abuse of notation we identify parabolics with their conjugacy classes.

Then $\partial\,\overline{\mathrm{Sh}}$ has a closed cover $\partial\,\overline{\mathrm{Sh}} = \bigcup_{P\in\mathscr{P}^{\max}} \partial^P\overline{\mathrm{Sh}}$. If $Q \in \mathscr{P}$ is conjugate to $\bigcap_{1\leq j\leq r} P_j$, for some set $\{P_j,\ 1 \leq j \leq r\}$ of rational maximal parabolics, let $\partial^Q = \bigcap_{1\leq j\leq r} \partial^{P_j}\overline{\mathrm{Sh}}$. Then (up to some technical questions involving connected components), ∂^Q is canonically a nilmanifold fibration over the Borel-Serre compactification of the locally symmetric space $X(L^Q)$ attached to a Levi component L^Q of Q. To each $P \in \mathscr{P}^{\max}$ is attached a corresponding *rational boundary component* $F(P)$ of X by the theory of Baily-Borel **[BB]**.

This point of view has been systematically exploited by Harder and Schwermer, among others, who have expressed $H^{\cdot}(\partial\,\mathrm{Sh}, \widetilde{V}^{\nabla})$ in terms of the $H^{\cdot}(X(L^Q), \widetilde{V}^{Q,\nabla})$, for various V^Q, and have used Eisenstein series to lift classes from $H^{\cdot}(\partial\,\mathrm{Sh}, \widetilde{V}^{\nabla})$. Specifically, let $T \subset G$ be a maximally split Cartan subalgebra, which we assume to be contained in Q, $W_{G,T} = W(\mathfrak{g}_{\mathbb{C}}, \mathrm{Lie}(T)_{\mathbb{C}})$, $W_{Q,T} = W(\mathrm{Lie}(L^Q)_{\mathbb{C}}, \mathrm{Lie}(T)_{\mathbb{C}})$, and $W^{1,Q} \subset W_{G,T}$ the set of coset representatives in $W_{Q,T}\backslash W_{G,T}$ of shortest length. For simplicity, assume V to be absolutely irreducible and rational over $\mathbb{Q}$; the general case can be treated similarly. Let Λ be the highest weight of V, relative to T and a system Φ_T^+ of positive roots with respect to which Q is standard; let ρ_T be $\frac{1}{2}$ the sum of the elements of Φ_T^+. For $w \in W^{1,Q}$, let $w(\Lambda) = w(\Lambda + \rho_T) - \rho_T$, and let $V_{w(\Lambda)}$ be the representation of L^Q with highest weight $w(\Lambda)$. Then the closed cover of $\partial\,\overline{\mathrm{Sh}}$ gives rise to a spectral sequence:

(3.5.2)
$$E_1^{i,j} = \bigoplus_{\mathrm{rk}(Q)=i}\ \bigoplus_{w\in W^{1,Q}} I_Q^G\{H^{j-l(w)}(X(L^Q), W_{w(\Lambda)})\} \Rightarrow H^{i+j}(\partial\,\overline{\mathrm{Sh}}, \widetilde{V}^{\nabla}),$$

where $I_Q^G\{\cdot\}$ is induction of representations from $Q(\mathbf{A}^f)$ to $G(\mathbf{A}^f)$, slightly modified to account for disconnectedness of the real groups; cf. **[Ha2,** §1].

A proof of the following theorem will appear in forthcoming joint work with Zucker.

3.5.3. THEOREM. (a) *The spectral sequence* (3.5.2) *is naturally a spectral sequence of mixed Hodge-de Rham structures over* $E(G, X)$.

(b) *More precisely, each* L^Q *can naturally be written* $G_{h,Q}\cdot G_{l,Q}$, *where* $G_{h,Q}$ *is a* $\mathbb{Q}$*-subgroup of hermitian type, associated to a Shimura variety* $\mathrm{Sh}(G_{h,Q}, X(Q))$, *and the algebraic connected component of* $G_{h,Q}\cap G_{l,Q}$ *is* $Z_G\cdot A_Q$, *where* A_Q *is a split component of* Q. *Here if* Q *is conjugate to* $\bigcap_{1\leq j\leq r} P_j$, *as above, let* $F(Q)$ *be the smallest among the rational boundary components* $F(P_j)$; *then* $F(Q)$ *is a connected component of* $X(Q)$, *and this characterizes* $G_{h,Q}$. *Write* $W_{w(\Lambda)} = W_{w(\Lambda,h)}\otimes W_{w(\Lambda,l)}$, *corresponding to* $L^Q \cong G_{h,Q}\cdot G_{l,Q}$. *Then each term* $H^{\cdot}(X(L^Q), W_{w(\Lambda)})$ *in* (3.5.2) *may* *naturally be written*

$$H^{\cdot}(\mathrm{Sh}(G_{h,Q}, X(Q)), W_{w(\Lambda,h)}) \otimes H^{\cdot}(X(G_{l,Q}), W_{w(\Lambda,l)})$$

(*again ignoring annoying problems coming from connected components*); *the mixed HdR structures on the left hand side of* (3.5.2) *is induced from the mixed HdR structures on* $H^{\cdot}(\mathrm{Sh}(G_{h,Q}, X(Q)), W_{w(\Lambda,h)})$ (*Theorem* 2.2.7), *and the trivial HdR structures on* $H^{\cdot}(X(G_{l,Q}), W_{w(\Lambda,l)})$.

Related results, in the context of the Baily-Borel compactification, have been obtained by Looijenga and Rapoport **[LR]** and (for ℓ-adic cohomology) by Pink **[P2]**. A number of authors have observed the presence of Tate Hodge structures in the boundary; cf. the comments at the end of **[H3**, §6], as well as **[OS]**. In our formulation, these are built into the term $H^{\cdot}(\mathrm{Sh}(G_{h,Q}, X(Q)), W_{w(\Lambda,h)})$.

4. An example

4.0. Let $\mathscr{K}$ be an imaginary quadratic field, and let $\iota \in \mathrm{Gal}(\mathscr{K}/\mathbb{Q})$ be the nontrivial automorphism. Let D be a central simple algebra with center $\mathscr{K}$, of dimension n^2 over $\mathscr{K}$, and let $*: D \to D$ be an involution of the second kind; thus the restriction of $*$ to $\mathscr{K}$ is the Galois automorphism σ. Let $\mathrm{GU}(D, *) = \{g \in D \mid g \cdot g^* \in \mathbb{G}_{m,\mathbb{Q}}\}$, where the last equation is the obvious shorthand. There is a homomorphism $\nu_D: \mathrm{GU}(D, *) \to \mathbb{G}_m$ of algebraic groups over $\mathbb{Q}$, with kernel $U(D, *)$. Similarly, there is a natural diagonal embedding $\mathbb{G}_{m,\mathbb{Q}} \to \mathrm{GU}(D, *)$; let $G_D = \mathrm{GU}(D, *)/\mathbb{G}_{m,\mathbb{Q}}$.

There is a (noncanonical) isomorphism $D(\mathbb{R}) \xrightarrow{\sim} M(n, \mathbb{C})$, in terms of which the involution can be taken to be $g \mapsto \Phi_{r,s}{}^t g^{-1}(\Phi_{r,s})^{-1}$, where $\Phi_{r,s}$ is the $n \times n$ matrix $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$ (I_α is the $\alpha \times \alpha$ identity matrix), for some pair with $r + s = n$. Thus $\mathrm{GU}(D, *)(\mathbb{R})$ is isomorphic to the similitude group $\mathrm{GU}(r, s)$ of a hermitian form of signature (r, s). Let $\widetilde{K}_{r,s} = U(r) \times U(s) \cdot \mathbb{R}^\times \subset \mathrm{GU}(r, s)$, $K_{r,s}$ its image in $G_D(\mathbb{R})$. The quotient $X_{r,s} = \mathrm{GU}(r, s)/\widetilde{K}_{r,s}$ is naturally an open domain in the complex Grassmannian of r-planes in n-space; the pair $(G_D, X_{r,s})$ defines a Shimura variety $\mathrm{Sh}(D, *) := \mathrm{Sh}(G_D, X_{r,s})$ of dimension rs. The reflex field $E(G_D, X_{r,s})$ is $\mathscr{K}$ unless $r = s$, in which case $E(G_D, X_{r,s}) = \mathbb{Q}$.

The Shimura varieties $\mathrm{Sh}(D, *)$ figure on Shimura's list of moduli spaces of abelian varieties of PEL type **[SH2]**; thus the HdR structures constructed in §2 are all (at least) motives for absolute Hodge cycles, and have associated ℓ-adic realizations. When D is a division algebra, $\mathrm{Sh}(D, *)$ is a projective variety. In this case, the corresponding families of λ-adic representations have been studied in great detail by Kottwitz **[K]**. The difficult problems connected with instability in the trace formula (cf. **[BR]**) evaporate for the varieties $\mathrm{Sh}(D, *)$, making it possible, in particular, to compare the λ-adic representations for different signatures (r, s). We begin by describing the Hodge structures.

4.2. We may identify $GU(D, *)_{\overline{\mathbb{Q}}} \cong (GL(n) \times \mathbb{G}_m)_{\overline{\mathbb{Q}}}$, and the representations (ρ, V) of G_D over $\overline{\mathbb{Q}}$ may be identified with representations of $GL(n)$ on which $-I_n \in GL(n)$ acts trivially. Thus the highest weight Λ of V, relative to the compact Cartan subalgebra $\mathfrak{h}$, is given by an n-tuple $(a_1 \geq \cdots \geq a_n)$ with $\sum_{i=1}^n a_1 \equiv 0 \pmod 2$. The corresponding variation of Hodge structure $\widetilde{V}$ over $\mathrm{Sh}(D, *)$ has weight $w_\rho = 0$, since the $\mathbb{Q}$-split component of the center of G_D is trivial. For simplicity, we first assume (ρ, V) rational over $\mathbb{Q}$; this is true if and only if $a_i + a_{n+1-i} = 0$ for all i.

Then the Hodge numbering on the middle-dimensional cohomology $H_!^{rs}(\mathrm{Sh}(D, *), \widetilde{V})$ is easy to compute. First suppose $(r, s) = (n-1, 1)$. Then

$$H_!^{n-1}(\mathrm{Sh}(D, *), \widetilde{V}) = \bigoplus_{i=1}^{n} H_!^{n-1}(\mathrm{Sh}(D, *), \widetilde{V})^{p_i, n-1-p_i}, \quad p_i = a_i + n - 1, \ i = 1, \ldots, n. \tag{4.2.1}$$

This is the same as the Hodge numbering on $H^{n-1}((G_D, X)_{n-1,1}); V)$, in the notation of (3.4.1). Note that $H^{n-1}((G_D, X_{n-1,1}); V)$ is a *regular* Hodge structure, in the sense of Definition 1.4.2. For general (r, s), let

$$\mathscr{H}(r, s; V) = \left\{ \sum_{j=1}^{s} (a_{i_j} + n - i_j) - \tfrac{1}{2}s(s-1) \mid 1 \leq i_1 < i_2 < \cdots < i_s \leq n \right\}.$$

These are the p_i's for the Hodge numbering on the weight rs Hodge structure

$$H^{rs}((G_D, X_{r,s}); V) \cong \Lambda^s[H^{n-1}((G_D, X_{n-1,1}); V)](\tfrac{1}{2}s(s-1)), \tag{4.2.2}$$

where the last term indicates a Tate twist, and the isomorphism is easily verified (and does not depend on the hypothesis that V be $\mathbb{Q}$-rational). Then

(4.2.3)
$$H_!^{rs}(\mathrm{Sh}(D, *), \widetilde{V}) = \bigoplus H_!^{rs}(\mathrm{Sh}(D, *), \widetilde{V})^{p_\alpha, rs-p_\alpha}, \qquad p_\alpha \in \mathscr{H}(r, s; V).$$

4.3. If G and G' are reductive groups over $\mathbb{Q}$ that are inner twists of one another, then $G(\mathbb{Q}_p) \cong G'(\mathbb{Q}_p)$ for almost all p. Let π_f, π'_f be representations of $G_D(\mathbf{A}^f)$ and $G_{D'}(\mathbf{A}^f)$, respectively. We say π_f and π'_f are *nearly equivalent*, and write $\pi_f \sim \pi'_f$, if the local components π_p and π'_p of π_f and π'_f are isomorphic for almost all p. In this way one can relate automorphic representations on different groups.

In particular, let $(D, *)$ and $(D', *')$ be two central simple algebras of dimension n^2 over $\mathscr{K}$, with involutions of the second kind. Then G_D and $G_{D'}$ are inner twists of each other. Thus we can define near equivalence of representations π_f, π'_f of $G_D(\mathbf{A}^f)$ and $G_{D'}(\mathbf{A}^f)$. Work of Kottwitz and Clozel, recently strengtened by R. Taylor, implies:

4.3.1. THEOREM (cf. **[K, C]** and forthcoming work of Taylor). (a) *Suppose D is a division algebra, and let π_f be a representation of $G_D(\mathbf{A}^f)$ such that $\pi_{w(\lambda)} \otimes \pi_f \subset \mathscr{A}(G_D)$ $(= \mathscr{A}_0(G_D))$, for some $\pi_{w(\lambda)}$ in the discrete series L-packet attached to V. Then $\pi_{w'(\lambda)} \otimes \pi_f \subset \mathscr{A}(G_D)$ for every w'. More precisely, the multiplicity $m_0(\pi_f) = m_0(\pi_{w(\lambda)} \otimes \pi_f)$ (notation* 3.2) *depends only on π_f and V, and not on the choice of $\pi_{w(\lambda)}$.*

(b) *Suppose $(r, s) = (n-1, 1)$. Let $H_\lambda(\pi_f, V)_{ss}$ be the semisimplified λ-adic realization of the absolute Hodge motive $H(\pi_f; V)$ (cf.* 2.3.3). *Then $H_\lambda(\pi_f; V)$ is pure of weight $n-1$ and (possibly after extension of the coefficient field $E(\pi_f)$) $H_\lambda(\pi_f; V)_{ss}$ is the direct sum of $m_0(\pi_f)$ copies of a representation $M_\lambda(\pi_f, V)$ of* $\mathrm{Gal}(\overline{\mathscr{K}}/\mathscr{K})$ *unramified outside a finite set of places (which can be determined explicitly). The characteristic polynomials of Frobenius acting on $M_\lambda(\pi_f, V)$ satisfy Langlands' reciprocity formula* **[L, BR]**.

(c) *Retaining the hypotheses of* (b), *let $(D', *')$ be another division algebra with involution of the second kind, of signature (r, s); suppose π_f and π'_f are nearly equivalent representations of $G_D(\mathbf{A}^f)$ and $G_{D'}(\mathbf{A}^f)$, respectively. Let $H(\pi_f; V) \subset H^\cdot(\mathrm{Sh}(D, *), \tilde{V})$, $H(\pi'_f; V) \subset H^\cdot(\mathrm{Sh}(D', *'), \tilde{V})$ be the corresponding absolute Hodge motives. Then there exists an integer m' such that*

$$H_\lambda(\pi'_f; V)_{ss} \cong [\Lambda^s(M_\lambda(\pi_f, V))(\tfrac{1}{2}s(s-1))]^{m'}, \tag{4.3.2}$$

(*exterior product over the coefficient field*); $(\frac{1}{2}s(s-1))$ *denotes a Tate twist.*

In fact, one may allow $\mathscr{K}$ to be an arbitrary CM field; this complicates the statement of (4.3.2) but introduces no essential new difficulties.

4.4. Clozel applies this result to study self-dual cuspidal automorphic representations of $\mathrm{GL}(n)_{\mathbb{Q}}$ with cohomology. Let Π be such a representation; thus $\Pi = \Pi_\infty \otimes \Pi_f$ is isomorphic to its own contragredient, and

$$H^\cdot(\mathfrak{gl}(n)_{\mathbb{C}}, O(n); V \otimes \Pi_\infty) \neq 0 \quad \text{for some } (\rho, V) \text{ as above}.$$

Here $\mathfrak{gl}(n) = \mathrm{Lie}(\mathrm{GL}(n, \mathbb{R}))$ and $O(n)$ is the usual maximal compact subgroup of $\mathrm{GL}(n, \mathbb{R})$. Assume the local components Π_p and $\Pi_{p'}$ to be square integrable at (at least) two finite primes p and p' of $\mathbb{Q}$, one of which splits in $\mathscr{K}$. For each pair (r, s) with $r + s = n$, Clozel constructs **[Cl]**:

(4.4.1) A division algebra D over $\mathscr{K}$ with involution of the second kind, such that D is split except at primes dividing p and p', and such that $\mathrm{GU}(D, *)_{\mathbb{R}} \cong \mathrm{GU}(r, s)$ (if n is odd the condition at p' is superfluous);

(4.4.2) An automorphic representation $\pi = \pi^D = \pi^D_\infty \otimes \pi^D_f \subset \mathscr{A}(G^D)$, such that (i) π_∞^D is in the discrete series L-packet attached to V, so that $H^{rs}(\mathrm{Lie}(G^D(\mathbb{R})), K_{r,s}; V \otimes \pi_\infty^D) \neq 0$; (ii) all π^D are nearly equivalent; and (4.4.3) If $(r, s) = (n-1, 1)$, the family of λ-adic representations $\{M_\lambda(\pi_f, V)\}$ is pure of weight $n-1$, and satisfies the relation

$$L(M_\lambda(\pi_f, V), s) = L(\Pi_{\mathscr{K}}, s - \tfrac{1}{2}(n-1)). \tag{4.4.4}$$

Here $\Pi_{\mathscr{K}}$ is the base change of Π to an automorphic representation of $\mathrm{GL}(n)_{\mathscr{K}}$, and $L(\Pi_{\mathscr{K}}, s)$ is the standard (Godement-Jacquet) L-function for $\mathrm{GL}(n)$, with functional equation $s \mapsto 1-s$.

Clozel's construction has the following motivic interpretation. The Langlands conjectures predict that every irreducible regular motive M over $\mathbb{Q}$, pure of weight $n-1$, of rank n over a totally real coefficient field, is associated to a self-dual cuspidal automorphic representation of $\mathrm{GL}(n)_{\mathbb{Q}}$ with cohomology in some V; this can be seen as a generalization in higher dimensions of the Shimura-Taniyama-Weil conjecture regarding elliptic curves of $\mathbb{Q}$. The highest weight of V is determined by the Hodge numbers of M, by formula (4.2.1). Thus, ignoring the conditions at p and p' (which would be rendered unnecessary by further development of the stable trace formula), the relation (4.4.4) conjecturally realizes all M of the given type in the cohomology of Shimura varieties $\mathrm{Sh}(D, *)$ of signature $(n-1, 1)$. More precisely, Clozel (conjecturally) identifies the λ-adic realizations of their base changes $M_{\mathscr{K}}$ to $\mathscr{K}$, for almost all imaginary quadratic fields $\mathscr{K}$; Blasius has shown that the system of n-dimensional λ-adic representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathscr{K})$ can be descended to a system of n-dimensional λ-adic representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

If one admits the Tate conjecture, then M is determined up to isomorphism by $H_\lambda(M)$. In particular, the HdR structure $H(M_{\mathscr{K}})$ over $\mathscr{K}$ (and also $H(R_{\mathscr{K}/\mathbb{Q}}M_{\mathscr{K}})$, cf. Remark 2.3.4) can be realized in terms of the HdR structures $H^{n-1}(\mathrm{Sh}(D, *), \tilde{V})$ described in §2. The associated $H^{p,q}$-structures over $\mathscr{K}$ can be realized in terms of automorphic forms of discrete series type, as explained in §3.1.2. In particular, since M is assumed to be regular, the quadratic periods $Q_j(H(M))$ of §1.4 can in theory be computed in terms of automorphic forms. (Since we only have immediate access to $M_{\mathscr{K}}$, one can determine the $Q_j(H(M))$ only up to $\mathscr{K}^\times$-multiples.)

4.5. The Tate conjecture predicts that the isomorphism (4.3.2) of λ-adic realizations comes from an isomorphism of motives, and in particular of HdR structures. To simplify the exposition, we make the hypothesis that the multiplicities $m_0(\pi_f)$ and m' of Theorem 4.3.1 equal 1; this is expected to be true (cf. [**BR**]), but it matters little for what we are about to say. We write $\pi_f = \pi_f^{(n-1,1)}$, $\pi'_f = \pi_f^{',(r,s)}$ to emphasize the signatures. We expect isomorphisms of motives over $\mathscr{K}$, with coefficients in $E(\pi_f) = E(\pi'_f)$:

$$H(\pi_f^{',(r,s)}; V) \cong [\Lambda^s H(\pi_f^{(n-1,1)}; V)](\tfrac{1}{2}s(s-1)). \tag{4.5.1}$$

The representation theory actually implies more. Let $W^1(r, s)$ be the subset W^1 defined in 2.2 for the real group $\mathrm{GU}(r, s)$. Each member $\pi'_{w'(\lambda)}$, $w' \in W^1(r, s)$ (resp. $\pi_{w(\lambda)}$, $w \in W^1(n-1, 1)$) of the discrete series L-packet associated to V for $G_{D'}$ (resp. G_D) contributes a one-dimensional $E(\pi_f)$-subspace $L'_{w'(\lambda)} \subset \mathrm{gr}^{\cdot}_F H_{\mathrm{DR}}(\pi_f'^{,(r,s)}, V)$ (resp. $L_{w(\lambda)} \subset \mathrm{gr}^{\cdot}_F H_{\mathrm{DR}}(\pi_f^{(n-1,1)}, V)$), or more properly a subspace of the corresponding $H^{p,q}$-structure. To each $w' \in W^1(r, s)$ one can attach a set $S(w')$ of s elements of $W^1(n-1, 1)$ such that $L'_{w'(\lambda)}$ on the left-hand side of (4.5.1) corresponds under the isomorphism to the exterior product $\bigwedge_{w\in S(w')} L_{w(\lambda)}$ on the right-hand side; this is implicit in (4.2.3). For each $w' \in W^1(r, s)$ (resp. $w \in W^1(n-1, 1)$), let $\omega_{w'} \in L'_{w'(\lambda)}$ (resp. $\omega_w \in L_{w(\lambda)}$) be a $\mathscr{K}$-rational $E(\pi_f)$-basis. Define the quadratic periods as in (1.4.3.6):

(4.5.2)

$$Q_{w'}(\pi_f'^{,(r,s)}) = \langle \omega_{w'}, F_\infty \omega_{w'} \rangle_{\mathrm{DR}}; \qquad Q_w(\pi_f^{(n-1,1)}) = \langle \omega_{w'}, F_\infty \omega_{w'} \rangle_{\mathrm{DR}}.$$

(Here F_∞ only operates on $R_{\mathscr{K}/\mathbb{Q}} H(\pi_f^{(n-1,1)}, V)$, $R_{\mathscr{K}/\mathbb{Q}} H(\pi_f'^{,(r,s)}, V)$, so these quadratic periods in $(E(\pi_f) \otimes \mathbb{C})^\times$ are only determined up to $(E(\pi_f) \otimes \mathscr{K})^\times$.) Recalling that Tate twisting leaves de Rham rational structures unchanged, (4.5.1) suggests the following

4.5.3. CONJECTURE. For all $w' \in W^1(r, s)$,

$$Q_{w'}(\pi_f'^{,(r,s)}) \sim_{E(\pi_f)\otimes\mathscr{K}} \prod_{w\in S(w')} Q_w(\pi_f^{(n-1,1)}).$$

4.5.4. For future reference, let $w'_1 \in W^1(r, s)$ be the parameter corresponding to the holomorphic discrete series representation of the $\mathrm{GU}(r, s)$ in our L-packet, so that $Q_{w'_1}(\pi_f'^{,(r,s)})$ is the Petersson square norm of an arithmetic holomorphic automorphic form on $\mathrm{Sh}(D', *')$. Then $S(w'_1) = \{w_1, \ldots, w_s\}$, where w_i corresponds to the component $H_!^{n-1}(\mathrm{Sh}(D, *), \widetilde{V})^{p_i, n-1-p_i}$ in (4.2.1). Thus

(4.5.5)

$$l(w_i) = n - 1 - i, \quad \text{and} \quad H_!^{n-1}(\mathrm{Sh}(D, *), \widetilde{V})^{p_i, n-1-p_i} \cong H_!^{i-1}(\mathscr{W}_{\tau(\Lambda, w_i)})$$

(cf. (2.2.7.5)). Then $Q_{w_i}(\pi_f^{(n-1,1)}) = Q_i(H(\pi_f^{(n-1,1)}, V))$, in the notation of (1.4.3.6), and Conjecture 4.5.3 becomes

4.5.6. CONJECTURE.

$$Q_{w'_1}(\pi_f'^{,(r,s)}) \sim_{E(\pi_f)\otimes\mathscr{K}} \prod_{j=1}^{s} Q_j(H(\pi_f^{(n-1,1)}, V)).$$

It is conceivable that these corollaries to the Tate conjecture, with applications to special values of the L-function (4.4.4), can actually be proved; cf. §6.2 for a discussion of the analogous period relations for the case of split D and D', where the results of §§4.3 and 4.4 are not yet available.

A general discussion of period relations for automorphic forms on Shimura varieties can be found in Ramakrishnan's article [**Ra**].

5. Periods of interior cohomology and L-functions

Let G be a reductive group over $\mathbb{Q}$, and let $r: {}^LG \to \mathrm{GL}(W)$ be a finite-dimensional representation of its L-group. There are two basic methods to prove the analytic continuation and functional equations of the automorphic L-function $L(\pi, s, r)$ of $\pi \in \mathscr{A}(G)$. The method of Langlands-Shahidi derives the analytic properties of $L(\pi, s, r)$ from Langlands' functional equations of Eisenstein series. The second method, whose modern history begins with Hecke, is that of integral representations. Roughly speaking, one finds a group H and a family Π_s, $s \in \mathbb{C}$, of automorphic representations of H, where either $H \subset G$ or $G \subset H$; one then integrates $\varphi \in \pi$ against $\Phi_s \in \Pi_s$ ($\mathrm{Re}(s) \gg 0$) over the smaller of the two adèle groups to obtain a zeta integral $Z(s, \varphi, \Phi)$. The integral $Z(s, \varphi, \Phi)$ is shown to have an analytic continuation and functional equation and, after often laborious computations, one expresses the zeta integral in terms of $L(\pi, s, r)$. Both methods are intrinsically limited to certain classes of r. For integral representations the limits remain to be discovered, and the construction of new ones remains an experimental science. (Surveys of the theory of L-functions can be found in [**GS, R2, Bu**].)

Of the two methods, that of integral representations appears to be better adapted to studying special values, or at least critical values in Deligne's sense. This is because the zeta integrals can frequently be interpreted as period integrals of cohomology classes (see [**Ha1, Ha2**] however, for applications of the Langlands-Shahidi method to relating critical values of different L-functions). Here the term "period" has to be understood broadly, since actual integrals of differentials over cycles occur only exceptionally.

We discuss some typical applications of the method of integral representations to the Deligne conjecture on critical values. Before beginning, it should be noted that this method only rarely provides direct information about L-functions of the motives occurring in the cohomology of Shimura varieties $\mathrm{Sh}(G, X)$. In order to draw a direct connection between an integral representation and a concrete motive in $H^{\cdot}(\mathrm{Sh}((G, X), \widetilde{RV})$, the following conditions must be satisfied:

(5.0.1) The integral has to represent the L-function $L(\pi, s, r_G)$, where r_G is the representation of the L-group discussed in [**BR**];

(5.0.2) Both G and H have to be of hermitian type, or G is of hermitian type and H is a subgroup defining topological cycles on G (modular symbols; see below);

(5.0.3) The integral has to have a cohomological interpretation; i.e. both φ and Φ_s (for critical s) have to define classes in coherent cohomology.

These conditions are extremely restrictive! For example, (5.0.1), which is absolutely essential, is not known for Hilbert modular varieties attached to totally real fields of degree ≥ 4, nor for Siegel modular varieties of large genus. In fact, apart from some low-dimensional cases, the analytic continuation of $L(\pi, s, r_G)$ is only known when $G(\mathbb{R})^{\text{der}}$ is isogenous to $U(r, s) \times \{\text{compact}\}$, $\min(r, s) \leq 1$ (or $\min(r, s) = 2$ when $G = G_D$ as in §4, with D a division algebra), or $\mathrm{SO}(p, 2) \times \{\text{compact}\}$; and of these, (5.0.2) is not known except for the case $U(r, 1) \times \{\text{compact}\}$. Even in these cases, the cohomological interpretation (5.0.3) is a priori given in terms of the wrong motive, and period relations (such as Conjecture 4.5.3) must be invoked in order to relate what is obtained to the Deligne conjecture: see 5.2 and 5.7, below.

In this section, when π is an automorphic representation of G which contributes to the cohomology $H^{\cdot}(\mathrm{Sh}(G, X), \widetilde{RV})$ for some V, we write $H(\pi_f)$ instead of $H^{\cdot}(\pi_f; \widetilde{RV})$ or $H^{\cdot}_!(\pi_f; \widetilde{RV})$. In the examples, $H(\pi_f)$ will be an absolute Hodge motive (i.e., not merely an HdR structure); its realizations will be denoted $H_?(\pi_f)$, $? = B$, DR, λ.

5.1. Hecke L-functions for GL(2). The most familiar integral representation is that of the L-function of an elliptic modular form. For details on what follows, cf. **[Sh3, MSw]**. Let $F(z) = \sum_{n=1}^{\infty} a_n q^n$, $q = e^{2\pi i z}$, be a cusp form on the upper half-plane of weight $k \in \mathbb{Z}$ (say on the Hecke congruence group $\Gamma_0(N)$). Assume F to be a Hecke eigenform, and $a_1 = 1$; then the Mellin transform

$$(5.1.1)\quad D(F, s) = \sum_{n=1}^{\infty} a_n n^{-s}; \qquad (2\pi)^{-s} \cdot \Gamma(s) \cdot D(F, s) = \int_0^{\infty} F(iy) y^{s-1}\, dy$$

has an Euler product. If F is a *new form*, then $D(F, s) = L(F, s)$ is the L-function of the automorphic representation $\pi(F)$ of $\mathrm{GL}(2)_{\mathbb{Q}}$ attached to F, and is also the L-function of the corresponding pure motive $M(\pi(F)) := H(\pi(F)_f)$ attached to F (or to $\pi(F)$) (cf. 2.3.3 and **[J, Sc]**).

Shimura was the first to remark that, for $s \in \mathbb{Z}$, $0 < s < k$, the value of the integral (5.1.1) can be interpreted as a *period* (or linear combination of periods) of the modular form F **[Sh1]**. Here the periods in question are for the cohomology groups introduced by Eichler **[Ei]** and more generally in **[Sh1, Sh3]**. In our case, if Sh is the elliptic modular curve attached to $\Gamma_0(N)$ and V_{k-2} is the $(k-2)$nd symmetric power of the standard representation of GL(2), then the Eichler cohomology groups are the $H^1_!(\mathrm{Sh}, V_{k-2})$ defined in §2, with Hodge type $(k-1, 0) + (0, k-1)$. The *q-expansion principle* **[DR]** identifies F, up to a power of $(2\pi i)$, with a DR-rational $(k-1, 0)$-form in $H^1_!(\mathrm{Sh}, V_{k-2})$ (if the Hecke eigenvalues of F are not in $\mathbb{Q}$ we need to work over a coefficient field); the period integral in (5.1.1) is then literally an integral of F over a Betti rational homology cycle. This is most clearly visible when $k = 2$ and $s = 1$; then (5.1.1) reduces to the integral of the differential form $F(z)\,dz$ over the positive y-axis in the upper half-plane

$\mathfrak{H}$, which defines a (Borel-Moore) homology cycle with closed supports in the modular curve $\mathrm{Sh} = \Gamma_0(N)\backslash\mathfrak{H}$.

The homology cycle in question is a *modular symbol*, which means it is defined by an algebraic subgroup $H = H_{\mathbb{Q}}$ of the group $\mathrm{GL}(2)_{\mathbb{Q}}$ defining the Shimura variety. In this case $H = \{(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix})\}$ is a one-dimensional torus. Let $G^+ \subset \mathrm{GL}(2, \mathbb{R})$ be the connected component containing the identity, and let $K_\infty \subset G^+$ be the stabilizer of the point $i \in \mathfrak{H}$. Then the modular symbol is essentially the image of $H(\mathbb{R}) \cap G^+$ in $\mathfrak{H}$, under the map $G^+ \to G^+/K_\infty \cong \mathfrak{H}$. (Recall that automorphic forms can be identified with cohomology classes on $\mathrm{Sh}(G, X)$ only after choosing a maximal compact subgroup or, equivalently, a base point in X. The base point $i \in \mathfrak{H}$ is a convenient and harmless choice.) The theory of modular symbols for GL(2) was developed by Birch, Manin, Mazur–Swinnerton-Dyer, and others, as an approach to the Birch–Swinnerton-Dyer conjecture and its p-adic generalizations; cf. **[MSw]**.

More generally, let E be any number field, and define the torus $H = H_E \subset \mathrm{GL}(2)_E$ as above. Fix a maximal compact subgroup $K_\infty \subset (\mathrm{GL}(2, E \otimes_{\mathbb{Q}} \mathbb{R})$. Let π be a cuspidal automorphic representation of $\mathrm{GL}(2)_E$ (K_∞-finite vectors). We view H and $G = \mathrm{GL}(2)_E$ as algebraic groups over $\mathbb{Q}$, as usual; let $\|\cdot\|$ denote the idèle norm on H. The Jacquet-Langlands theory attaches an L-function $L(\pi, s)$ to π as the integral over the image $\tilde{c}[H, K_f]$ of $H(\mathbb{Q})\backslash H(\mathbf{A})$ in $G(\mathbb{Q})\backslash G(\mathbf{A})/K_f$ of a well-chosen $F \in \pi$, multiplied by $\|\cdot\|^s$. Here $K_f \subset G(\mathbf{A}_f)$ is some open compact subgroup. Again $\tilde{c}[H, K_f]$ defines a homology class with closed support in the locally symmetric space

$$X(G, K_f) = G(\mathbb{Q})\backslash G(\mathbf{A})/Z_G(\mathbb{R}) \cdot K_\infty K_f; \tag{5.1.2}$$

the image $c[H, K_f]$ of $\tilde{c}[H, K_f]$ is another example of a modular symbol.

When E is totally real, $[E : \mathbb{Q}] = d$, $X(G, K_f)$ is a Shimura variety Sh_E of dimension d, with reflex field $\mathbb{Q}$ (the Hilbert-Blumenthal modular variety). For appropriate $\pi = \pi_\infty \otimes \pi_f$, there exists a finite-dimensional representation $V = V(\pi_\infty)$ of G such that $H(\pi_f)$ is a nonzero direct summand of the (absolute Hodge) motive $H_!^d(\mathrm{Sh}_E, \widetilde{RV})$. One again identifies certain critical values of $L(\pi, s)$ as periods of DR-rational holomorphic automorphic forms over modular symbols; this was worked out in certain cases by Manin **[Ma]**. However, when $d > 1$ $L(\pi, s)$ is *not* the L-function of $H(\pi_f)$! Thus the relation between modular symbols and critical values of L-functions does not automatically have consequences for Deligne's conjecture. A conjectural solution to this problem is explained in 5.2.

When E is not totally real, there is no longer a visible motive of any kind. However, it is still possible to use modular symbols to prove relations between critical values of the twisted L-functions $L(\pi \otimes \chi, s)$, as χ runs through the algebraic Hecke characters of E (or of $\mathrm{GL}(2, \mathbf{A}_E)$) that are compatible with those predicted by Deligne's conjecture, and which permit the definition of p-adic L-functions. This program was carried out in part by P. Kurchanov in the late 1970s **[Ku]**. When E contains a CM field, Hida has recently proved a very general result **[Hi]**.

5.2. Let E be totally real, $[E:\mathbb{Q}]=d$, $G=R_{E/\mathbb{Q}}\mathrm{GL}(2)_E$. Let $\{\sigma_1,\dots,\sigma_d\}$ denote the distinct real embeddings of E. The representations $\pi=\pi_\infty\otimes\pi_f\subset\mathscr{A}_0(G)$ that define motives in the cohomology of Sh_E are those for which π_∞ belongs to the discrete series, and such that the action of $Z_G(\mathbb{R})$ on π_∞ is trivial on the Zariski closure of the global units of E. Such π_∞ are indexed by $(d+1)$-tuples $(k_1,\dots,k_d,r)$, with $k_i\geq 2$, $k_i\equiv r \pmod 2$ for all i. Here k_i correspond to the classical weight at σ_i of the *holomorphic* Hilbert modular forms which occur in the motive, and r is a central character. The classical L-function $L(\pi,s)$, including its Γ-factors, then looks like the L-function of a motive $M(\pi)$ over E, or rank 2 over the field of definition $E(\pi_f)$ of π_f, with the property that, for each i, $M(\pi)_i=M(\pi)_{\sigma_i(E)}$ has Hodge numbers $(k_i-1, 2-r-k_i)$. However, if $V=V(\pi_\infty)$ as above, then the motive $H(\pi_f)\subset H_!^d(\mathrm{Sh}_E,\widetilde{RV})$ is rational over $\mathbb{Q}$, and looks like $\bigotimes_{i=1}^d M(\pi)_i$, the tensor product taken over $E(\pi_f)$.

The factorization of the ℓ-adic realizations $H(\pi_f)_\ell$ of $H(\pi_f)$ is discussed in **[BR]** and Tilouine's talk, to which we refer for what follows. In most cases, $M(\pi)_i$ can be directly constructed in the cohomology (with twisted coefficients) of a *Shimura curve*. Let D_i be a quaternion algebra over E such that

$$D_i\otimes_{\sigma_i(E)}\mathbb{R}\cong M(2,\mathbb{R});$$
$$D_i\otimes_{\sigma_j(E)}\mathbb{R}\cong\mathbb{H}\ \text{(Hamiltonian quaternions)}\quad \text{if } j\neq i.$$

Then the algebraic group $G_i=R_{E/\mathbb{Q}}D_i^\times$ defines a Shimura curve $\mathrm{Sh}(D_i)=\mathrm{Sh}(G_i,X_i)$, where $X_i\cong\mathbb{C}-\mathbb{R}\ (=\mathfrak{H}\cup\overline{\mathfrak{H}})$. The reflex field $E(G_i,X_i)$ is $\sigma_i(E)$. If the local factor π_v of π is special or supercuspidal as a representation of $\mathrm{GL}(2,E_v)$ for every v at which D_i does not split, then there is a representation $\pi_i\in\mathscr{A}(G_i)$ which is *nearly equivalent* to π, in the sense defined in 4.3. Now V is again a representation of G_i, and $H(\pi_{i,f})\subset H^1(\mathrm{Sh}_i,\widetilde{RV})$ is an absolute Hodge motive over E (via σ_i) such that

$$L(H(\pi_{i,f}),s)=L(\pi,s). \tag{5.2.1}$$

Thus $H(\pi_{i,f})$ can be taken as a model for $M(\pi)_i$.

If π_i as above exists for one i, then it exists for all $i\in\{1,\dots,d\}$. As in 4.5, Tate's conjecture predicts that the isomorphism of λ-adic realizations

$$\bigotimes_{i=1}^{d}H_\lambda(\pi_{i,f})_\lambda\cong H_\lambda(\pi_f) \tag{5.2.2}$$

comes from an isomorphism of motives. One can at least ask for an isomorphism of HdR structures, or even of $H^{p,q}$-structures. A related question was posed by Shimura in **[Sh6]**, and taken up again in some of his more recent papers. Each $H_{\mathrm{DR}}(\pi_{i,f})$ contains the class of a $\overline{\mathbb{Q}}$-arithmetic holomorphic form $f_i \in \mathscr{A}(G_i)$. Similarly, $H_{\mathrm{DR}}(\pi_f)$ contains the class of a $\overline{\mathbb{Q}}$-arithmetic Hilbert modular form $f \in \mathscr{A}_0(G)$. Shimura conjectured that the periods of f along homology classes (for example, modular symbols) in Sh_E could be factored as products of the periods of the f_i. Similarly, he conjectured that normalized Petersson inner products satisfy

$$(5.2.3) \qquad \langle f, f \rangle \Big/ \left[\prod_{i=1}^{d} \langle f_i, f_i \rangle\right] \in \overline{\mathbb{Q}}^{\times}.$$

Shimura thus conjectures the existence of an isomorphism of $H^{p,q}$-structures, up to factors in $\overline{\mathbb{Q}}$.

More precise versions of these conjectures (with $\overline{\mathbb{Q}}$ replaced by $\mathbb{Q}$ in (5.2.3), for example) were proposed by Panchishkin and Yoshida in **[Pa, Y]**. That (5.2.2) should correspond to an isomorphism of motives was stressed by Oda in **[O]**.

The relation (5.2.3) of Petersson inner products has been proved in most cases in **[H6]**, using the ideas described in §6. Shimura's conjectured relation between the actual periods, which is necessary in order to apply the known results on the critical values of $L(\pi, s)$ to Deligne's conjecture, seems considerably more delicate.

5.3. Let G be any reductive group over $\mathbb{Q}$, Z_G its center, $K_\infty \subset G(\mathbb{R})$ a maximal compact subgroup, and $H \subset G$ a connected $\mathbb{Q}$-subgroup such that $K_\infty \cap H(\mathbb{R})$ is a maximal compact subgroup of $H(\mathbb{R})$. Let $K_f \subset G(\mathbf{A}^f)$ be an open compact subgroup, and define $X(G, K_f)$ as in (5.1.2). It is known that, at least (but not exclusively) when H is reductive, the map

$$(5.3.1) \qquad X(H, H(\mathbf{A}^f) \cap K_f) \to X(G, K_f)$$

is proper (this is due to Borel and G. Prasad; see **[Ash]** for a proof). When K_f is small enough, $X(H, H(\mathbf{A}^f) \cap K_f)$ is orientable, hence its image under (5.3.1) defines a homology class $c[H, K_f]$ with closed supports. The $c[H, K_f]$ are called (generalized) modular symbols, and their properties have been studied in **[Ash, AB]** and elsewhere.

The relation described above between modular symbols and L-values does not seem to generalize to higher-dimensional Shimura varieties. Rather, periods along modular symbols seem to be related to *residues* of L-functions of cusp forms defined by integral representations. The first example of this phenomenon appears to be due to Petersson **[Pet]**. Let π and π' be two cuspidal automorphic representations of $\mathrm{GL}(2)_{\mathbb{Q}}$, associated to elliptic modular new forms $F(z) = \sum_{n=1}^{\infty} a_n q^n$, $F'(z) = \sum_{n=1}^{\infty} b_n q^n$, of weights k and k', respectively, as in 5.1. We assume the central characters of π and π' to be powers of the idèle norm, for simplicity; then $\mathbb{Q}(\{a_n, b_n\})$ is a totally real number field. Define the Rankin-Selberg convolution

(5.3.2)

$$D(s, F, F') = \sum_{n=1}^{\infty} a_n b_n q^n, \qquad L(\pi\times\pi', s) = L(2s+2-k-k')\cdot D(s, F, F');$$

then $L(\pi \times \pi', s) = L(M(\pi) \otimes M(\pi'), s)$, where the motives $M(\pi)$ and $M(\pi')$ are as in §5.1. Let $\langle\cdot, \cdot\rangle$ denote the Petersson inner product on modular forms of weight k. Petersson observed that $D(s, F, F')$ is entire unless $k = k'$; if $k = k'$ then (in an appropriate normalization)

$$\text{(5.3.3)} \qquad \operatorname{Res}_{s=k} D(s, F, F') = \frac{(4\pi)^k}{(k-1)!}\langle F, F'\rangle .$$

Here, letting $H = \mathrm{GL}(2)_{\mathbb{Q}}$ diagonally embedded in $G = H\times H$, the Petersson inner product should be thought of as the period attached to the cohomology class on G defined by $F \otimes \overline{F}'$ and to the modular symbol for H.

The right-hand side of (5.3.3) vanishes identically unless $\pi \cong \overline{\pi}'$ (the complex conjugate of π', which we are assuming to equal π'). Although this example may appear artificial, this observation suggests that cohomology classes with nontrivial periods along modular symbols tend to have "additional symmetries". (Actually, cuspidal cohomology classes strongly tend to have trivial periods along modular symbols. A very general result of this type is contained in [**AGR**].) Insofar as periods along modular symbols equal residues of L-functions, this is consistent with the Tate conjecture: poles of L-functions correspond to algebraic cycles, which correspond to constraints on Mumford-Tate groups. Langlands' philosophy then leads to the expectation that an automorphic representation π of G with nontrivial periods along modular symbols "comes from" an automorphic representation π' of some group G' which is smaller than G, by Langlands' functoriality (cf. the discussion in 3.4); hence the reference to "additional symmetries".

This vague notion can be made precise in terms of the L-group. In practice, it often turns out that automorphic forms π on G with nontrivial periods along modular symbols attached to H do indeed "come from" π' on G' as above. Construction of π' is carried out either explicitly, by means of theta functions, or implicitly, by means of character identities (the Arthur-Selberg trace formula and Jacquet's relative trace formula).

5.4. Theta correspondence. In its modern formulation, due essentially to Howe, we are given a symplectic vector space $\mathbb{W}$ over a number field E, which we take to be totally real, and two reductive subgroups $G, G' \subset \mathrm{Sp}(\mathbb{W})$ which are *mutual centralizers* in $\mathrm{Sp}(\mathbb{W})$. Such a pair (G, G') is called a *dual reductive pair* (DRP). Here are two examples:

(5.4.1) Let W (resp. V) be a vector space over E with an alternating (resp. symmetric) bilinear form. Then $\mathbb{W} = W\otimes_E V$ is naturally a symplectic vector space and $(\mathrm{Sp}(W), O(V))$ is a DRP in $\mathrm{Sp}(W)$.

(5.4.2) Let $\mathscr{K}/E$ be a totally imaginary quadratic extension and W (resp. V) be a Hermitian (resp. skew-Hermitian) vector space over $\mathscr{K}$. Then $\mathbb{W} = R_{\mathscr{K}/E}(W \otimes_{\mathscr{K}} V)$ is again symplectic; $(U(W), U(V))$ is a DRP in $\mathrm{Sp}(\mathbb{W})$.

For any vector space $\mathbb{W}^+$ over E, the space $\mathscr{S}(\mathbb{W}^+(\mathbf{A}_E))$ of *Schwartz-Bruhat functions* on $\mathbb{W}^+(\mathbf{A}_E)$ consists of functions φ which are locally constant and compactly supported in the non-Archimedean variables, C^∞ in the Archimedean variables, and such that $|D\varphi|$ is bounded for any differential operator in the Archimedean variables with polynomial coefficients. In what follows, we let $\mathbb{W} = \mathbb{W}^+ \oplus \mathbb{W}^-$ be a decomposition of $\mathbb{W}$ as a direct sum of isotropic subspaces. The choice of such a *polarization* introduces a noncanonical element into the theory of theta functions, which makes arithmetic applications more difficult than they should be.

For each $\varphi \in \mathscr{S}(\mathbb{W}^+(\mathbf{A}_E))$ and each additive character $\psi\colon \mathbf{A}_E/E \to \mathbb{C}^\times$, Weil's theory of theta functions [We] defines a *theta kernel* $\theta_\varphi(\mathbf{g})$, $\mathbf{g} \in \mathrm{Sp}(\mathbb{W}, E)\backslash\widetilde{\mathrm{Sp}}(\mathbb{W})$, where $\widetilde{\mathrm{Sp}}(\mathbb{W})$ is a central extension of the adèle group $\mathrm{Sp}(\mathbb{W}, \mathbf{A}_E)$ by $\mathbb{C}^\times$ which naturally splits over $\mathrm{Sp}(\mathbb{W}, E)$. Recall the construction: the *oscillator* (or *Segal-Shale-Weil*) *representation* ω_ψ is a unitary representation of $\widetilde{\mathrm{Sp}}(\mathbb{W})$ on $\mathscr{S}(\mathbb{W}^+(\mathbf{A}_E))$, in the so-called Schrödinger model. Then the sum

$$\theta_\varphi(\mathbf{g}) = \sum_{\lambda \in W^+(E)} (\omega_\psi(\mathbf{g})\varphi)(\lambda), \qquad \mathbf{g} \in \widetilde{\mathrm{Sp}}(\mathbb{W}),$$

converges absolutely to a function on $\mathrm{Sp}(\mathbb{W}, E)\backslash\widetilde{\mathrm{Sp}}(\mathbb{W})$. If (G, G') is a DRP, let $\widetilde{G}$ (resp. $\widetilde{G}'$) be the inverse image of $G(\mathbf{A}_E)$ (resp. $\widetilde{G}'(\mathbf{A}_E)$) in $\widetilde{\mathrm{Sp}}(\mathbb{W})$. One defines the spaces $\mathscr{A}_0(\widetilde{G})$, resp. $\mathscr{A}_0(\widetilde{G}')$, of cusp forms on $G(E)\backslash\widetilde{G}$, resp. $G'(E)\backslash\widetilde{G}'$, without difficulty. In most cases one can identify $\mathscr{A}_0(\widetilde{G}) = \mathscr{A}_0(G)$, and we do so in what follows (we do not identify $\mathscr{A}_0(\widetilde{G}') = \mathscr{A}_0(G')$, although this is also usually possible). Given $F \in \mathscr{A}_0(G)$, one defines its theta-lifting

$$\Theta_\varphi(F)(g') = \int_{G(E)\backslash\widetilde{G}} \theta_\varphi(g \cdot g')F(g)\,dg \in \mathscr{A}_0(\widetilde{G}'). \tag{5.4.3}$$

If $\Pi \in \mathscr{A}_0(G)$ is an irreducible $\widetilde{G}$-submodule, one defines

$$\Theta(\Pi) = \{\Theta_\varphi(F),\ F \in \Pi,\ \text{all } \varphi\} \subset \mathscr{A}_0(\widetilde{G}'). \tag{5.4.4}$$

Then $\Theta(\Pi)$ is a $\widetilde{G}'$-submodule. It has been proved by Howe and Waldspurger (up to possible problems at primes of E dividing 2) that $\Theta(\Pi) \cong (\Pi')^m$, where m is a nonnegative integer, conjectured to equal 1, and Π' is an irreducible representation of $\widetilde{G}'$ [Wa4]. Analogously, given $F' \in \mathscr{A}_0(\widetilde{G}')$, one defines

$$^t\Theta_\varphi(F')(g) = \int_{G'(E)\backslash\widetilde{G}'} \theta_\varphi(g \cdot g')F(g')\,dg' \in \mathscr{A}_0(G). \tag{5.4.5}$$

A fundamental principle, going back (by way of Weil and Siegel) to the analytic theory of representations of integers by quadratic forms, is that the (generalized) Fourier coefficients of $\Theta_\varphi(F)$ can be expressed in terms of *periods* of F along modular symbols attached to subgroups $H \subset G$. For example, if $G = O(V)$, $\dim V = m$, and $G' = \mathrm{Sp}(W)$, $\dim W = 2$ (so $G \cong \mathrm{SL}(2)$), then we may take $H = O(V')$, where $V' \subset V$ is a subspace of dimension $m-1$ (which may depend on the choice of Fourier coefficient). This principle has been exploited by numerous authors to characterize the forms with nontrivial periods along modular symbols. Most striking, perhaps, is the work of Kudla and Millson **[KM]**, which in many cases identifies the Poincaré duals of the modular symbols attached to H as explicit ${}^t\Theta_\varphi(F')$, $F' \in \mathscr{A}_0(\widetilde{G}')$. Here G' is determined by H; the rule of thumb is that G' is small for large H, so such modular symbols generate a very special subspace of homology.

5.5. Petersson's residue formula (5.3.3) has been generalized in many contexts. For example, suppose F is a real quadratic field, and define the Hilbert-Blumenthal modular surface Sh_F as in §5.1. Then Sh_F is the Shimura variety attached to the group $G = R_{F/\mathbb{Q}}\mathrm{GL}(2)_F$. Let $H = \mathrm{GL}(2)_{\mathbb{Q}}$, viewed naturally as a subgroup of G; the corresponding modular symbols $c[H, K_f]$ are algebraic subvarieties of Sh_F, called *Hirzebruch-Zagier cycles*, and their irreducible components are elliptic modular curves. A generalization of (5.3.3) **[HLR**, Satz 3.13] implies (roughly) that the poles of the L-function of $IH^2(\mathrm{Sh}_F)$ correspond to cusp forms on G with nonzero periods along the Hirzebruch-Zagier cycles. This is an important ingredient in the work of Harder-Langlands-Rapoport on the Tate conjecture for Sh_F (cf. Tate's talk).

Generalizations of (5.3.3) to Shimura varieties of low dimension have been proved by Howe–Piatetski-Shapiro, Oda, Rallis, Piatetski-Shapiro and others. In higher dimensions the residues of L-functions tend to be more directly related to theta functions, especially the Siegel-Weil formula and its generalizations by Kudla and Rallis, as well as **[R2, KR2]**.

5.6. Although critical values can rarely be computed directly in terms of periods, there is nevertheless a huge literature dedicated to the critical values of L-functions $L(\pi, s, r)$ whose integral representations satisfy conditions (5.0.2) and (5.0.3) (but not necessarily (5.0.1)). In most cases, π is the representation generated by a holomorphic modular form F, assumed DR rational in the sense described in 3.1.2, and the value at the critical point $s = m$ takes the form

$$L(\pi, m, r) = \alpha \times (2\pi i)^k \times p(\omega) \times \langle F, F\rangle. \tag{5.6.1}$$

Here α is an algebraic number, usually determined up to factors in the coefficient field of the HdR structure $H(\pi_f)$, $k \in \mathbb{Z}$ depends on m and π, $p(\omega)$ is an auxiliary product of CM periods (usually absent, but cf. (5.7.3)), and $\langle\cdot,\cdot\rangle$ is a generalized Petersson inner product. The first such result was proved by Shimura **[Sh4]** for the Rankin-Selberg convolutions (5.3.2). Shimura's methods were generalized to other groups by Shimura, Sturm, Harris, Garrett, Böcherer, Orloff, and others; other (related) approaches to (5.6.1) were found in some cases by Sturm and Zagier.

Formula (5.6.1) does not immediately confirm Deligne's conjecture. In the first place, unless $r = r_G$ the L-function on the left-hand side of (5.6.1) is not $L(H(\pi_f), s)$; moreover, the Petersson inner products are not obviously related to the Deligne period c^+. When $G = \mathrm{GL}(2)_{\mathbf{Q}}$ and r is the symmetric square of the standard representation, Deligne showed in **[D3]** that $c^+(\mathrm{Sym}^2(H(\pi_f)))$ is an elementary multiple of $\langle F, F\rangle$, thus confirming the compatibility of results of Sturm and Zagier with his conjecture. A similar but more elaborate computation, due to Blasius, showed that, when $\pi^1, \dots, \pi^r$ is a family of automorphic representations of $\mathrm{GL}(2)_{\mathbf{Q}}$, associated to holomorphic modular forms, then, for $r > 1$, $c^+(H(\pi^1_f)\otimes\cdots\otimes H(\pi^r_f))$ can be expressed in terms of Petersson inner products; thus Shimura's theorem (for $r = 2$) and results of Garrett and Orloff (for $r = 3$) are compatible with Deligne's conjecture **[Bl2]**. Proposition 1.4.10 above is a more recent result in the same vein, and is applied to unitary Shimura varieties in §5.7, below.

An extension of Shimura's method to certain nonholomorphic coherent cohomology classes on Hilbert-Blumenthal modular varieties is given in **[H5]**. There it is proved that the critical values of the Rankin-Selberg convolution $L(\pi\times\pi', s)$, where π and π' are as in §5.2, are given by

$$(5.6.2)\qquad L(\pi\times\pi', m) = (\text{algebraic number})\times(2\pi i)^k\times\nu^I(\pi)\times\nu^{I'}(\pi').$$

Here $k\in\mathbb{Z}$ and $\nu^I(\pi)$ and $\nu^{I'}(\pi')$ are cohomologically defined invariants, related to the factorization (5.2.3) of the Petersson inner products of holomorphic forms in π and π', respectively.

5.7. Application to L-functions of $\mathrm{Sh}(D, *)$. We return to the notation of §4.4; thus $(D, *)$ and $(D', *')$ have signatures $(n-1, 1)$ and (r, s) respectively. To simplify the exposition, assume n to be *even*. Let $U_D = U(D, *)$, $U_{D'} = U(D', *')$; these groups map to G_D and $G_{D'}$, respectively. The automorphic representation $\pi = \pi^D \subset \mathscr{A}_0(G_D)$ (resp. $\pi' = \pi^{D'} \subset \mathscr{A}_0(G_{D'})$), which contributes to the motive $H(\pi_f^{(n-1,1)}, V)$ (resp. $H(\pi_f'^{,(r,s)}, V)$) pulls back to a representation of U_D (resp. $U_{D'}$), which is not necessarily irreducible; let π_0 (resp. π_0') denote any irreducible constituent of this pullback. Then $L(\Pi_{\mathscr{K}}, x)$, introduced in §4.4 as an L-function on $\mathrm{GL}(n)_{\mathscr{K}}$, can be realized as a Langlands L-function for U_D or $U_{D'}$, namely the standard L-function $L(\pi_0, s, \mathrm{St}) = L(\pi_0', s, \mathrm{St})$. The definition of this L-function is recalled in 6.1 for *split* D, but the general case is identical, since the local factors of $D(\mathbf{A})$ are split almost everywhere. More generally, let

$\chi: \mathscr{K}_{\mathbf{A}}^{\times}/\mathscr{K}^{\times} \to \mathbb{C}^{\times}$ be an algebraic Hecke character whose restriction to the idèles $\mathbf{A}^{\times}$ of $\mathbb{Q}$ is trivial; thus we may assume $k(\chi) = -\lambda(\chi) = k$, for some $k > 0$ in $\mathbb{Z}$ (notation (1.4.6)). Then there are twisted representations $\pi_0 * \chi \subset \mathscr{A}(U_D)$, $\pi_0' * \chi \subset \mathscr{A}(U_{D'})$, such that

$$\text{(5.7.1)} \quad \begin{aligned} L(M_{\mathbb{Q},\lambda}(\Pi) \otimes R_{\mathscr{K}/\mathbb{Q}}M(\chi)_\lambda, s) &= L(\Pi_{\mathscr{K}} \otimes (\chi \circ \det), s - \tfrac{1}{2}(n-1)) \\ &= L(\pi_0 * \chi, s - \tfrac{1}{2}(n-1), \mathrm{St}). \end{aligned}$$

Here $M_{\mathbb{Q},\lambda}(\Pi)$ is as in (4.4.5) and $M(\chi)_\lambda$ is the λ-adic realization of the CM motive $M(\chi)$ of 1.4. Strictly speaking, the Euler factors of $L(\pi_0 * \chi, s, \mathrm{St})$ are known only at unramified primes; we take (5.7.1) as the definition of the remaining Euler factors, or appeal to **[PSR1]**.

Now $L(\pi_0 * \chi, s, \mathrm{St}) = L(\pi_0' * \chi, s, \mathrm{St})$ has integral representations on both U_D and $U_{D'}$, by the doubling method of Piatetski-Shapiro–Rallis and Garrett (**[PSR2]**; this case was worked out in **[G]** and **[Li2]**). Under appropriate hypotheses on χ, the techniques discussed in 5.6 permit us to determine the critical values of $L(\pi_0 * \chi, s, \mathrm{St})$ in terms of periods on $G_{D'}$. Here is the answer. Recall that $H(\pi_f, V)$ (notation 4.3.1(b)) is the sum of $m_0(\pi_f)$-copies of a regular HdR structure over some coefficient field, with Hodge numbers $(a_i + n - i, i - 1 - a_i)$. In order to compare our result with Proposition 1.4.10, let $\chi' = \chi \cdot \|\cdot\|_{\mathbf{A}}^k$, so $k(\chi') = 2k$, $\lambda(\chi') = 0$, $\chi_0' \equiv 1$; the weight $w = n - 1$. Suppose χ' belongs to the sth critical interval for the given HdR structure, as in 1.4.10. Then the center of symmetry of the L-function $L(M_{\mathbb{Q},\lambda}(\Pi) \otimes R_{\mathscr{K}/\mathbb{Q}}M(\chi')_\lambda, s)$ is the point $\frac{1}{2}n$, and the critical values are the set

$$\text{(5.7.2)} \qquad \Sigma = \{m \in \mathbb{Z} \mid \tfrac{1}{2}n \le m \le \min(a_r + s - k, a_s + r + k)\}$$

together with its reflection $n - \Sigma$ in the center of symmetry. Now suppose $m \in \Sigma$. Let $w_1' \in W^1(r, s)$ be the parameter corresponding to the holomorphic discrete series representation as in §4.5, so that $Q_{w_1}(\pi_f'^{,(r,s)})$ is the Petersson square norm of an arithmetic holomorphic automorphic form on $\mathrm{Sh}(D', *')$. Now let $m \in \Sigma$, and suppose moreover that $m > n$; the latter condition is certainly superfluous but it allows us to ignore possible poles of our L-functions. Let $E(\chi) = E(\pi_f) \cdot \mathbb{Q}(\chi)$. Then

5.7.3. Theorem **[H7]**. *There is a constant $C_\infty \in \mathbb{C}^{\times}$, depending only on V, r, and s, such that*

$$\begin{aligned} L(M_{\mathbb{Q},\lambda}(\Pi) \otimes R_{\mathscr{K}/\mathbb{Q}}M(\chi')_\lambda, m) &= L(\pi_0 * \chi, s + k - \tfrac{1}{2}w, \mathrm{St}) \\ &\sim_{E(\chi)\otimes\mathscr{K}} (2\pi i)^{mn - \frac{1}{2}nw + rs} \cdot C_\infty \cdot Q(\chi)^{s - n/2} \cdot Q_{w_1}(\pi_f'^{,(r,s)}). \end{aligned}$$

Let $\beta = (2\pi i)^{rs} \cdot C_\infty$. Comparing 5.7.3 with Conjecture 4.5.6, we find that (5.7.4) *Modulo the Tate conjecture,*

$$L(M_{\mathbb{Q},\lambda}(\Pi) \otimes R_{\mathscr{K}/\mathbb{Q}} M(\chi')_\lambda , m) \\ \sim_{E(\chi)\otimes\mathscr{K}} (2\pi i)^{mn-\frac{1}{2}nw} \cdot \beta \cdot Q(\chi)^{s-n/2} \cdot \prod_{j=1}^{s} Q_j(H(\pi_f^{(n-1,1)}, V)).$$

Note added in proof. At least when the Hodge numbers are sufficiently general—the relation $|a_i - a_{i-1}| \geq 2$ is sufficient—this relation can now be proved without reference to the Tate conjecture, at least up to factors in $\overline{\mathbb{Q}}$. The method is a variant of that described in §6.2, below.

It is easy to see that, apart from the indeterminacy introduced by possible multiples in $\mathscr{K}^\times$ in the formula, this is precisely what is predicted by Deligne's conjecture, as interpreted in Proposition 1.4.10, provided $\beta \in \mathbb{Q}^\times$. Garrett has recently announced computations suggesting that this is indeed the case, but it remains to be checked that his normalizations are compatible with those used to prove 5.7.3.

5.8. It should be mentioned that noncritical values do not fit so well into this picture. This may be due in part to lack of imagination, but there also seem to be intrinsic barriers to extending the methods. Especially, the methods discussed in §6 completely break down outside the critical range.

6. Applications of the Siegel-Weil formula

Occasionally, an integral representation which does not satisfy (5.0.2) and (5.0.3) can be related to a period integral on another pair of groups (G', H'), which does satisfy (5.0.2) and (5.0.3), usually by means of a theta correspondence. For example, in **[Sh6]**, Shimura obtained (in most cases) an expression for the left-hand side of (5.6.2) in terms of the Shimura varieties attached to multiplicative groups of quaternion algebras, a formula quite different from that of **[H5]**. Waldspurger has used a similar idea to prove in some cases that, when π' is attached to a CM motive and m is at the center of the critical strip the algebraic factor α in (5.6.2) is essentially a *square* in the coefficient field (cf. **[Wa3]**; the result is stated there only for the ratio of two L-values, but Waldspurger's methods can be modified to yield the stronger assertion). In **[HK]** Kudla and I have proved a similar result for the L-function attached to three elliptic modular forms, whose integral representation was discovered by Garrett.

These examples are all ultimately related to the *Siegel-Weil formula* and its extensions beyond the range of absolute convergence. We will be discussing another, still largely conjectural application of the Siegel-Weil formula, in this case to the standard L-functions for unitary groups. The most general extension of the Siegel-Weil formula is due to Kudla and Rallis. Its application to special values of L-functions is based on the *doubling method*, which arose out of work of Rallis on the theta correspondence, and which neatly exhibits the interconnection between special values of L-functions, period relations, and the arithmetic of the theta correspondence.

6.1. Let E be a totally real number field, $\mathcal{K}$ a CM quadratic extension of F. We return to the notation of 5.4, and consider dual reductive pairs (G, G') of one of the following types:

$$G = \mathrm{Sp}(W), \quad G' = O(\mathcal{V}), \quad \text{as in } (5.4.1); \quad \dim_E W = 2n, \quad \dim_E \mathcal{V} = m; \tag{6.1.1}$$

$$G = U(W), \quad G' = U(\mathcal{V}), \quad \text{as in } (5.4.2); \quad \dim_{\mathcal{K}} W = 2n, \quad \dim_{\mathcal{K}}(\mathcal{V}) = m. \tag{6.1.2}$$

In (6.1.1) we assume that the determinant of $\mathcal{V}$ is a square in E, so that the quadratic Dirichlet character associated to $\mathcal{V}$ is trivial; in both cases we assume m to be even. These hypotheses can be relaxed but the results are then more complicated to state.

Let $\Gamma = \mathrm{Gal}(\mathcal{K}/E)$. In (6.1.1) (resp. (6.1.2)) the L-group ${}^LG$ may be taken to be $\mathrm{SO}(2n+1, \mathbb{C})$ (resp. $\mathrm{GL}(2n, \mathbb{C}) \rtimes \Gamma$, with the nontrivial element of Γ acting on $\mathrm{GL}(m, \mathbb{C})$ by an outer automorphism); let

$$\mathrm{St}\colon {}^LG \to \mathrm{GL}(2n+1, \mathbb{C}) \quad (\text{resp. } \mathrm{St}\colon {}^LG \to \mathrm{GL}(4n, \mathbb{C}))$$

be the natural representation of $\mathrm{SO}(2n+1, \mathbb{C})$ (resp. the representation of $\mathrm{GL}(2n, \mathbb{C}) \rtimes \Gamma$ induced from the natural representation of $\mathrm{GL}(2n, \mathbb{C})$).

Let $\pi \subset \mathcal{A}_0(G_1)$, $\check{\pi}$ its contragredient. The local Euler factors $L_v(\pi, s, \mathrm{St})$ have been defined provisionally at all places, and the partial L-function $L^{\mathrm{fin}}(\pi, s, \mathrm{St}) := \prod_{v \text{ finite}} L_v(\pi, s, \mathrm{St})$ satisfies a functional equation of the form

$$L^{\mathrm{fin}}(\pi, s, \mathrm{St}) = \varepsilon^{\mathrm{fin}}(\pi, s) \cdot \gamma_\infty(\pi, s) \cdot L^{\mathrm{fin}}(\pi, 1-s, \mathrm{St}). \tag{6.1.3}$$

Here the Archimedean factor $\gamma_\infty(\pi, s)$ and the ε-factor $\varepsilon^{\mathrm{fin}}(\pi, s)$ have the usual form but are not yet completely understood. See **[PSR1**, §3**]** for a more detailed account in case (6.1.1), and **[Li2]** for an analysis of (6.1.2).

Choose $f \in \pi$, $\check{f} \in \check{\pi}$. We denote the theta lift from $\mathcal{A}_0(G)$ to $\mathcal{A}(G')$ by ${}^t\Theta_\varphi$, φ variable as in 5.4. Let $\langle \cdot, \cdot \rangle_G$, $\langle \cdot, \cdot \rangle_{G'}$, be the normalized L_2-inner products, relative to Tamagawa measure. Let $s_0 = \frac{1}{2}(m-2n)$ in case (6.1.1), $s_0 = \frac{1}{2}(m-2n+1)$ in case (6.1.2). Define a Dirichlet series

$$d_n(s) = \zeta_E(s+n+\tfrac{1}{2}) \cdot \prod_{i=1}^{n} \zeta_E(2s+2i-1) \quad (\text{case } (6.1.1))$$
$$= \prod_{0 \le j < n} \zeta_E(2s+2n-2j) \prod_{0 < j \le n} L(2s+2n-2j+1, \varepsilon_{\mathcal{K}}) \quad (\text{case } (6.1.2),$$

where $\zeta_E(s)$ is the Dedekind zeta function of E and $L(s, \varepsilon_{\mathcal{K}})$ is the L-function attached to the nontrivial Dirichlet character $\varepsilon_{\mathcal{K}}: \Gamma \to \{\pm 1\}$.

Assume $s_0 \geq \frac{1}{2}$; let $r = 1$ if $s_0 > \frac{1}{2}$, and $r = 2$ if $s_0 = \frac{1}{2}$. *Rallis' inner product formula* is the general name for identities of the form:

(6.1.4)
$$\langle {}^t\Theta_{\varphi_1}(f), {}^t\Theta_{\varphi_2}(\check{f})\rangle_{G'} = r\cdot\langle f, \check{f}\rangle_G \cdot L(\pi, s_0, \mathrm{St})\cdot d_n(s_0)^{-1}\cdot \prod_{v\in S} Z_v(\varphi_1, \varphi_2, f, \check{f}).$$

Here S is a finite set of bad places, including Archimedean places, and $Z_v(\varphi_1, \varphi_2, f, \check{f})$ is a normalized zeta integral for the Piatetski-Shapiro–Rallis doubling method **[PSR2]**.

When m is sufficiently large relative to n ($m > 4n+2$ in (6.1.1); $m > 4n$ in (6.1.2)), such an identity follows from Weil's generalization of the Siegel mass formula **[We]** and from the Piatetski-Shapiro–Rallis integral representation for $L(\pi, s_0, \mathrm{St})$; in this case it is worked out in some generality by **[Li2]**, following and extending the original ideal of Rallis **[R1]**.

As s_0 approaches the central point $\frac{1}{2}$ (or as m approaches $2n$), the function $L(\pi, s, \mathrm{St})$ may acquire poles at $s = s_0$, and the formula (6.1.4) must obviously be modified. In a remarkable series of papers, (especially **[KR1, KR2]**) Kudla and Rallis have extended the Siegel-Weil formula beyond the range considered by Weil, in the process obtaining precise information about the location of possible poles of $L(\pi, s, \mathrm{St})$ and extending the identity (6.1.4), appropriately modified, down to the central point. As of this writing their results are nearly complete in case (6.1.1), and rather less so in case (6.1.2) (cf. **[R2]** for a partial account of their results).

We will only be concerned with case (6.1.2), when $m = 2n$, so $s_0 = \frac{1}{2}$ is the center of symmetry. At this point the L-function has no pole, and (6.1.4) needs no modification. The formula (6.1.4), whose proof remains to be written down in this case, can be regarded as a heuristic device. The term $d_n(s_0)^{-1} \cdot \prod_{v\in S} Z_v(\varphi_1, \varphi_2, f, \check{f})$ is essentially elementary (although the actual evaluation of $Z_v(\varphi_1, \varphi_2, f, \check{f})$ is not known in general when v is a real place). Thus the formula (6.1.4) relates the special value $L(\pi, \frac{1}{2}, \mathrm{St})$ to the Petersson norms of forms in π in ${}^t\Theta(\pi)$.

Note added in proof. Kudla has recently discovered a variant of formula (6.1.4) which calculates the derivative at the center of symmetry of $L(\pi, s, \mathrm{St})$ whenever the value vanishes. His method should clarify and generalize the Gross-Zagier formula, which relates the derivatives of L-functions of elliptic modular forms to height pairings.

6.2. We wish to apply formula (6.1.4) when the Archimedean component π_∞ of π belongs to the discrete series. For each place v of E, let $G_v = G(E_v)$, $G'_v = G'(E_v)$. When v is Archimedean and $K_v \subset G_v$ is a chosen maximal compact subgroup, a $(\mathrm{Lie}(G_v), K_v)$-module will be referred to, by abuse of language, as an *admissible representation* of G_v; the standard definition is used at non-Archimedean places. Then for any v we can define $\mathrm{Rep}(G_v)$ as the set of equivalence classes of irreducible admissible representations of G_v.

Recall that, as an abstract adèlic representation, ${}^t\Theta(\pi)$ is completely determined (except possibly in residue characteristic 2) by π. More precisely, if v is a place of E, there are local correspondences ${}^t\Theta_v$: $\mathrm{Rep}(G_v) \to \mathrm{Rep}(G'_v)$. These may be multivalued in residue characteristic 2; we assume Howe's conjecture that this is not the case (this is known at least for unramified representations). Then if π is the restricted tensor product $\bigotimes_v \pi_v$, where v runs through places of E and $\pi_v \in \mathrm{Rep}(G_v)$, then ${}^t\Theta(\pi) \cong \bigotimes_v {}^t\Theta_v(\pi_v)$.

When v is real and π_v is in the discrete series, then ${}^t\Theta_v(\pi_v)$ has been determined explicitly Li in **[Li1]**. When ${}^t\Theta_v(\pi_v) \neq 0$, it belongs to the discrete series. Moreover, let χ_v vary among one-dimensional representations of G_v. Then for given G'_v, the set of χ_v for which ${}^t\Theta_v(\pi_v \otimes \chi_v) \neq 0$ can be determined precisely. We work this out in a special case.

Fix a place v, and suppose W has signature $(n-1, 1)$ and $\mathscr{V}$ has signature (r, s) at v. We may identify $G \cong U(D_0, *)$, $G' \cong U(D_0, *')$, as in §4, with D_0 the *split* algebra $M(n)_\chi$ and $*$, $*'$ appropriate involutions of the second kind. Then the discrete series L-packets of G_v and G'_v correspond to representations V of $\mathrm{GL}(n, \mathbb{C})$, with highest weight λ, and may be parametrized as in §4.5. In particular, the discrete series L-packet of $G_v \cong U(n-1, 1)$ has n members, denoted π_{w_j}, $j = 1, \dots, n$, as in §4.5.4, so that π_{w_j} corresponds to coherent cohomology in degree $j-1$ (4.5.5). Let V be such a representation, with highest weight $\Lambda = (a_1 \geq a_2 \geq \cdots \geq a_n)$ with $a_i + a_{n+1-i} = 0$ for all i, as in §4.2. Suppose $\pi_v = \pi_{w_s(\lambda)}$ is in the corresponding L-packet. Index characters of G_v or G'_v by integers: $\{\chi_k(g) = \det(g)^{-k},\ k \in \mathbb{Z}\}$. Then Li's computation **[Li1, Theorem 6.2]** can be paraphrased as follows:

6.2.1. LEMMA. *The theta lifting* ${}^t\Theta_v(\pi_{w_s(\lambda)} \otimes \chi_k)$ *vanishes unless* $2k$ *is in the* $(r+1)$*st critical interval for the HdR structure with Hodge numbers* $(a_i + n - i, n - 1 + i - a_i)$. *In the latter case we have, in the notation of* §4.5,

$${}^t\Theta_v(\pi_{w_s(\lambda)} \otimes \chi_k) = [\pi_{w'_1(\lambda)}]^* \otimes \chi_{-k},$$

the representation corresponding to antiholomorphic forms.

Now we return to the global setting. For simplicity we assume $E = \mathbb{Q}$, so $\mathscr{K}$ is imaginary quadratic, and we assume the hypotheses of §4.5; we also assume Λ regular, so that interior and cuspidal cohomology coincide. We also assume that the results of Kottwitz described in §4 can be extended to stable discrete series L-packets for the untwisted unitary groups G and G'; this is generally assumed to be possible and can probably be worked out under appropriate local hypotheses on π_f. Let $\pi = \pi_{w_s(\lambda)} \otimes \pi_f$, and let

$\chi: G(\mathbf{A})/G(\mathbb{Q}) \to \mathbb{C}^\times$ be a character with Archimedean component χ_k. Let

$$\pi'_f = \bigotimes_{v \text{ finite}} {}^t\Theta_v(\pi_v \otimes \chi_v) \otimes \chi_v .$$

Choose a polarization $W \otimes_{\mathscr{K}} \mathscr{V} = \mathbb{W}^+ \oplus \mathbb{W}^-$ as in §5.4, and let $\mathscr{S}_\infty = \mathscr{S}(\mathbb{W}^+(\mathbb{R}))$, $\mathscr{S}_f = \mathscr{S}(\mathbb{W}^+(\mathbf{A}^f))$, $\mathscr{S}_f(\overline{\mathbb{Q}}) \subset \mathscr{S}_f$ the $\overline{\mathbb{Q}}$-subspace of $\overline{\mathbb{Q}}$-valued functions. Define

$$q_s = n - 1 - p_s, \quad \mathscr{W}_s = \mathscr{W}_{\tau(\Lambda, w_s)}, \quad \mathscr{W}' = \mathscr{W}_{\tau(\Lambda, w'_1)},$$

in the notation of 4.5.4; thus $\mathscr{W}_s$ and $\mathscr{W}'$ are automorphic vector bundles on the Shimura varieties $\mathrm{Sh}(D_0, *)$ and $\mathrm{Sh}(D_0, *')$, respectively. Let

$$H_!^{s-1}(\mathscr{W}_s)(\pi_f^{(n-1,1)}) = H(\pi_f^{(n-1,1)}, V) \cap H_!^{n-1}(\mathrm{Sh}(D_0, *), \tilde{V})^{p_s, q_s}$$

in the notation of 4.5, and define $H_!^0(\mathscr{W}')(\pi_f'^{,(r,s)}) \subset H(\pi_f'^{,(r,s)}, V)$ analogously to be the lowest step in the Hodge filtration. We identify the elements $f \in H_!^{s-1}(\mathscr{W}_s)(\pi_f^{(n-1,1)})$ with vectors in π via (3.2.4.3), and make the analogous identification for $H_!^0(\mathscr{W}')(\pi_f'^{,(r,s)})$; however, we identify the elements of $H_!^0(\mathscr{W}')(\pi_f'^{,(r,s)})$ with *antiholomorphic* automorphic forms by replacing the given complex structure with its complex conjugate. Then, for an appropriate choice of $\varphi_\infty \in \mathscr{S}_\infty$, the theta lifting defines a homomorphism

$$\begin{gathered} {}^t\Theta_\chi : \mathscr{S}_f \otimes H_!^{s-1}(\mathscr{W}_s)(\pi_f^{(n-1,1)}) \to H_!^0(\mathscr{W}')(\pi_f'^{,(r,s)}), \\ {}^t\Theta_\chi(\varphi_f \otimes f) = [{}^t\Theta_{\varphi_\infty \otimes \varphi_f}(f \otimes \chi)] \otimes \chi . \end{gathered}$$

This homomorphism commutes with the action of $G'(\mathbf{A}^f)$ on both sides. Since we are assuming $m_0(\pi_f) = m' = 1$, it follows easily from Schur's lemma that there exists a constant $c(\pi, \chi) \in \mathbb{C}$, well defined up to $\overline{\mathbb{Q}}^\times$, such that

(6.2.2)

$${}^t\Theta_\chi(\mathscr{S}_f(\overline{\mathbb{Q}}) \otimes H_!^{s-1}(\mathscr{W}_s)(\pi_f^{(n-1,1)})(\overline{\mathbb{Q}})) = c(\pi, \chi) \cdot H_!^0(\mathscr{W}')(\pi_f'^{,(r,s)})(\overline{\mathbb{Q}}).$$

In (6.2.2), one may easily replace $\overline{\mathbb{Q}}$ by $E(\pi_f) \cdot E(\chi) \cdot \mathbb{Q}^{\mathrm{ab}}$, where $\mathbb{Q}^{\mathrm{ab}}$ is the maximal abelian extension of $\mathbb{Q}$, but the descent to $E(\pi_f) \cdot E(\chi)$ is not automatic, since the action of $G'(\mathbf{A}^f)$ on $\mathscr{S}_f$ is a priori defined only over $\mathbb{Q}^{\mathrm{ab}}$.

Write $Q_s(\pi) = Q_s(H(\pi_f^{(n-1,1)}, V))$, in the notation of (1.4.3.6). In what follows, if $a, b \in \mathbb{C}$, we write $a \sim_{\overline{\mathbb{Q}}} b$ if $ab = 0$ or if $a/b \in \overline{\mathbb{Q}}^\times$. Combining (6.2.2) with (6.1.4), observing that the terms $Z_v(\varphi_1, \varphi_2, f, \check{f})$ for v finite can be assumed algebraic, and recalling (4.5.2), we obtain the relation

(6.2.3)
$$\begin{aligned} &|c(\pi, \chi)|^2 \cdot Q_{w'_1}(\pi_f'^{,(r,s)}) \\ &\quad \sim_{\overline{\mathbb{Q}}} Q_s(\pi) \cdot L(\pi \otimes \chi, \tfrac{1}{2}, \mathrm{St}) \cdot d_n(\tfrac{1}{2})^{-1} \cdot Z_\infty(\varphi_1, \varphi_2, f, \check{f}). \end{aligned}$$

Recall that ${}^t\Theta_\chi$ vanishes identically unless $2k$ belongs to the $(r+1)$st critical interval for the HdR structure $H(\pi_f^{(n-1,1)}, V)$. In view of Theorem 5.7.3, the relation (6.2.3) reduces Conjecture 4.5.6 for the split algebra D_0 to

(a) a statement about the rationality of the numbers $c(\pi, \chi)$ (provided the relation (6.2.2) can be descended to $E(\pi_f) \otimes \mathscr{K})$, and

(b) the nonvanishing Conjecture 1.5.2.

As explained in the remarks following (5.7.4), (a) and (b) would have consequences for Deligne's conjecture for the motives in question. Of these two steps, (a), which is the subject of work in progress with Kudla, seems to be attainable, whereas (b) looks very hard.

When $n = 2$, but E is an arbitrary totally real field, the above constructions have been carried out successfully in **[H6]**, and they form the basis for the proof of the period relations (5.2.3) conjectured by Shimura.

References

[AJ] J. Adams and J. Johnson, *Endoscopic groups and packets of non-tempered representations*, Compositio Math. **64** (1987), 271–309.

[A] J. Arthur, *Unipotent automorphic representations: Conjectures*, Astérisque **171–172** (1989), 13–71.

[AA] L. Clozel and J. S. Milne, *Automorphic forms, Shimura varieties, and L-functions*. I, II, Proceedings of the Conference held at the University of Michigan, Ann Arbor, Michigan, July 6–16, 1988, Perspectives in Mathematics, vols. 10, 11, Academic Press, Boston, MA, 1990.

[Ash] A. Ash, *Non-square integrable cohomology of arithmetic groups*, Duke Math. J. **47** (1980), 435–449.

[AB] A. Ash and A. Borel, *Generalized modular symbols*, in **[Luminy]**.

[AGR] A. Ash, D. Ginzburg, and S. Rallis, *Vanishing periods of cusp forms over modular symbols*, Math. Annalen **296** (1993), 709–724.

[AMRT] A. Ash, D. Mumford, M. Rapoport, and Y.-S. Tai, *Smooth compactification of locally symmetric varieties*, Lie Groups: History, Frontiers and Applications, vol. 4, Math. Sci. Press, Brookline, MA, 1975.

[BB] W. L. Baily, Jr. and A. Borel, *Compactification of arithmetic quotients of bounded symmetric domains*, Ann. of Math. **84** (1966), 442–528.

[BBD] A. Beilinson, J. Bernstein, and P. Deligne, *Faisceaux pervers*, Astérisque, no. 100 (1982).

[Bl1] D. Blasius, *On the critical values of Hecke L-series*, Ann. of Math. **124** (1986), 23–63.

[Bl2] ——, *Critical values of certain tensor product L-functions*, Invent. Math. **90** (1987), 181–188.

[BHR] D. Blasius, M. Harris, and D. Ramakrishnan, *Coherent cohomology, limits of discrete series, and Galois conjugation*, Duke Math. J. (in press).

[BR] D. Blasius and J. Rogawski, *Zeta-functions of Shimura varieties*, these Proceeedings, vol. 2, pp. 525–571.

[Bo] A. Borel, *Stable real cohomology of arithmetic groups*, Ann. Sci. École Norm. Sup. **7** (1974), 235–272.

[BS] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Comment. Math. Helv. **48** (1973), 436–491.

[BW] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, Ann. of Math. Stud. no. 94, Princeton Univ. Press, Princeton, NJ, 1980.

[Bu] D. Bump, *The Rankin-Selberg method: A survey*, Number Theory, Trace Formulas, and Discrete Groups (Symposium in honor of Atle Selberg, Oslo, 1987), Academic Press, New York, 1989.

[Cl] L. Clozel, *Représentations galoisiennes associées aux représentations automorphes autoduales de* GL(n), Inst. Hautes Études Sci. Publ. Math. **73** (1991), 97–145.

[D1] P. Deligne, *Equations différentielles à points singuliers réguliers*, Lecture Notes in Math., vol. 163, Springer-Verlag, Berlin and New York, 1970.

[D2] ——, *Théorie de Hodge*. II, Inst. Hautes Études Sci. Publ. Math. **44** (1974), 5–77.

[D3] ——, *Valeurs de fonctions L et périodes d'intégrales*, Proc. Sympos. Pure Math. **33**, part 2 (1979), 313–346.

[D4] ——, *Le groupe fondamental de la droite projective moins trois points*, Publ. Res. Inst. Math. Sci. **16** (1989), 79–297.

[DMOS] P. Deligne, J. S. Milne, A. Ogus, and K.-Y. Shih, *Hodge cycles, motives, and Shimura varieties*, Lecture Notes in Math., vol. 900, Springer-Verlag, Berlin and New York, 1982.

[DR] P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973, pp. 143–316.

[Ei] M. Eichler, *Eine Verallgemeinerung der Abelschen Integrale*, Math. Z., vol. 67 (1957), 267–298.

[F] G. Faltings, *On the cohomology of locally symmetric hermitian spaces*, Lecture Notes in Math., vol. 1029, Springer-Verlag, Berlin and New York, 1984, pp. 55–98.

[FC] G. Faltings and C.-L. Chai, *Degeneration of Abelian varieties*, Springer-Verlag, Berlin and New York, 1990.

[Fr] J. Franke, *Harmonic analysis in weighted L^2-spaces* (manuscript, 1991).

[G] P. Garrett, *Integral representations of Eisenstein series and L-functions*, Number Theory, Trace formulas, and Discrete Groups (Symposium in honor of Atle Selberg, Oslo, 1987), Academic Press, New York, 1989.

[GS] S. Gelbart and F. Shahidi, *The analytic properties of automorphic L-functions*, Perspectives in Mathematics, Academic Press, New York, 1988.

[GSc] C. Goldstein and N. Schappacher, *Conjecture de Deligne et Γ-hypothèse de Lichtenbaum sur les corps quadratiques imaginaires*, C. R. Acad. Sci. Paris **296** (1983), 615–618.

[Ha1] G. Harder, *Eisenstein cohomology of arithmetic groups: The case* GL_2, Invent. Math. **89** (1987), 37–118.

[Ha2] ——, *Some results on the Eisenstein cohomology of arithmetic subgroups of* GL_n, in **[Luminy]**.

[Ha3] ——, *Arithmetische Eigenschaften von Eisensteinklassen, die modulare Konstruktion von gemischten Motiven und von Erweiterungen endlicher Galois Modulen*, manuscript, 1991.

[HLR] G. Harder, R. P. Langlands, and M. Rapoport, *Algebraische Zykeln auf Hilbert-Blumenthal Flächen*, J. Reine Angew. Math. **366** (1986), 53–120.

[H1] M. Harris, *Arithmetic vector bundles and automorphic forms on Shimura varieties*. I, Invent. Math. **82** (1985), 151–189; II, Compositio Math. **60** (1986), 323–378.

[H2] ——, *Functorial properties of toroidal compactification of locally symmetric varieties*, Proc. London Math. Soc. **59** (1989), 1–22.

[H3] ——, *Automorphic forms and the cohomology of vector bundles on Shimura varieties*, in [**AA**, Vol. II, pp. 41–91].

[H4] ——, *Automorphic forms of $\bar\partial$-cohomology type as coherent cohomology classes*, J. Differential Geom. **32** (1990), 1–63.

[H5] ——, *Period invariants of Hilbert modular forms*. I, in **[Luminy]**.

[H6] ——, *L-functions of 2×2 unitary groups and factorization of periods of Hilbert modular forms*, J. Amer. Math. Soc. **6** (1993), 637–720.

[H7] ——, *L-functions and periods of polarized regular motives* (manuscript, 1991).

[HK] M. Harris and S. Kudla, *The central critical value of a triple product L-function*, Ann. of Math. **133** (1991), 605–672.

[HZ] M. Harris and S. Zucker, *Boundary cohomology of Shimura varieties*. I (to appear in Ann. Sci. École Norm. Sup.); II (Invent. Math., in press).

[Hi] H. Hida, *On the critical values of L-functions of* GL(2) *and* GL(2) × GL(2), manuscript, 1992.

[J] U. Jannsen, *Mixed motives and algebraic K-theory*, Lecture Notes in Math., vol. 1400, Springer-Verlag, Berlin and New York, 1990.

[Ka] N. Katz, *p-adic interpolation of real analytic Eisenstein series*, Ann. of Math. **104** (1976), 459–517.

[K] R. Kottwitz, *On the lambda-adic representations associated to some simple Shimura varieties*, Invent. Math. **108** (1992), 653–665.

[KM] S. Kudla and J. Millson, *Geodesic cycles and the Weil representation. I, Quotients of hyperbolic space and Siegel modular forms*, Compositio Math. **45** (1982), 207–271.

[KR1] S. S. Kudla and S. Rallis, *On the Weil-Siegel formula*, J. Reine Angew. Math. **387** (1987), 1–68; II: *The isotropic convergent case*, ibid. **391** (1987), 65–84.

[KR2] _____, *Poles of Eisenstein series and L-functions* (Festschrift in honor of I. I. Piatetski-Shapiro), Israel Math. Conf. Proc. **3** (1990), 81–110.

[Ku] P. F. Kurchanov, *Local measures connected with Jacquet-Langlands cusp forms over fields of CM type*, Math. USSR Sb. **36** (1980), 449–467.

[L] R. P. Langlands, *Automorphic representations, Shimura varieties, and motives: Ein Märchen*, Proc. Sympos. Pure Math. **33** (1979), Part II, 205–246.

[Li1] J.-S. Li, *Theta lifting for unitary representations with non-zero cohomology*, Duke Math. J. **61** (1990), 913–937.

[Li2] _____, *Non-vanishing theorems for the cohomology of certain arithmetic quotients*, J. Reine Angew. Math. **428** (1992), 177–217.

[Lo] E. Looijenga, *L^2-cohomology of locally symmetric varieties*, Compositio Math. **67** (1988), 3–20.

[LR] E. Looijenga and M. Rapoport, *Weights in the local cohomology of a Baily-Borel compactification*, Proc. Sympos. Pure Math. **53** (1991).

[Luminy] J.-P. Labesse and J. Schwermer, *Cohomology of arithmetic groups and automorphic forms* (Proceedings, Luminy/Marseille 1989), Lecture Notes in Math., vol. 1447, Springer-Verlag, Berlin and New York, 1990, pp. 155–202.

[Ma] Yu. I. Manin, *Non-Archimedean integration and p-adic Hecke-Langlands L-series*, Russian Math. Surveys **31** (1976), 5–54.

[MM] Y. Matsushima and S. Murakami, *On vector bundle values, harmonic forms, and automorphic forms on symmetric spaces*, Ann. of Math. **78** (1963), 365–416.

[60] B. Mazur and H. P. F. Swinnerton-Dyer, *Arithmetic of Weil curves*, Invent. Math. **25** (1974), 1–61.

[Mi] J. S. Milne, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, in [**AA**, vol. I, pp. 283–414].

[Mu] D. Mumford, *Hirzebruch's proportionality theorem in the non-compact case*, Invent. Math. **42** (1977), 239–272.

[O] T. Oda, *Periods of Hilbert modular surfaces*, Progress in Math., no. 19, Birkhaüser, Boston, 1982.

[OS] T. Oda and J. Schwermer, *Mixed Hodge structures and automorphic forms for Siegel modular varieties of degree two*, Math. Ann. **286** (1990), 481–509.

[Pa] A. A. Panchishkin, *Motives over totally real fields and p-adic L-functions*, manuscript, 1990.

[Pe] H. Petersson, *Über die Berechnung der Skalarprodukte ganzer Modulformen*, Comment. Math. Helv. **22** (1949), 168–199.

[PSR1] I. I. Piatetski-Shapiro and S. Rallis, *ε-factor of representations of classical groups*, Proc. Nat. Acad. Sci. U.S.A. **83** (1986), 4589–4593.

[PSR2] _____, *L-functions for the classical groups*, Lecture Notes in Math, vol. 1254, Springer-Verlag, Berlin and New York, 1987, pp. 1–52.

[P1] R. Pink, *Arithmetical compactification of mixed Shimura varieties*, Bonner Math. Schriften, vol. 209, Math. Inst. Univ. Bonn, Bonn, 1990.

[P2] _____, *On ℓ-adic sheaves on Shimura varieties and their higher direct images in the Baily-Borel compactification*, manuscript, 1991.

[R1] S. Rallis, *L-functions and the oscillator representation*, Lecture Notes in Math., vol. 1245, Springer-Verlag, Berlin and New York, 1987.

[R2] _____, *Poles of standard L-functions*, Proc. Internat. Congress Math. 1990, vol. 2, Springer-Verlag, Berlin, 1991, pp. 833–845.

[RA] D. Ramakrishnan, *Problems arising from the Tate and Beilinson conjectures in the context of Shimura varieties,* in **[AA,** Vol. II, pp. 227–252].

[Sa1] M. Saito, *Modules de Hodge polarisables,* Res. Inst. Math. Sci. Publ. **24** (1988), 849–995.

[Sa2] ——, *Mixed Hodge modules and admissible variations,* C. R. Acad. Sci. Paris **309** (1989), 351–356.

[Sa3] ——, *Mixed Hodge modules,* Res. Inst. Math. Sci. Publ. **26** (1990), 221–333.

[SS] L. Saper and M. Stern, *L_2-cohomology of arithmetic varieties,* Ann. of Math. **132** (1990), 1–69.

[Sv] G. Savin, *Limit multiplicities of discrete series,* Invent. Math. **95** (1989), 149–159.

[Sch] W. Schmid, *Variations of Hodge structure: The singularities of the period mapping,* Invent. Math. **22** (1973), 211–319.

[Sc] A. J. Scholl, *Motives for modular forms,* Invent. Math. **100** (1990), 419–430.

[Schw] J. Schwermer, *Cohomology of arithmetic groups, automorphic forms, and L-functions,* in **[Luminy]**.

[Sh1] G. Shimura, *Sur les intégrales attachées aux formes automorphes,* J. Math. Soc. Japan **11** (1959), 291–311.

[Sh2] ——, *Moduli of abelian varieties and number theory,* Proc. Sympos. Pure Math. **IX** (1966), 312–332.

[Sh3] ——, *Introduction to the arithmetic theory of automorphic functions,* Publ. Math. Soc. Japan, no. 11, Iwanami Shoten/Princeton Univ. Press, 1971.

[Sh4] ——, *The special values of the zeta functions associated with cusp forms,* Comm. Pure Appl. Math. **29** (1976), 783–804.

[Sh5] ——, *The critical values of certain zeta functions associated with modular forms of half-integral weight,* J. Math. Soc. Japan **33** (1981), 649–672.

[Sh6] ——, *Algebraic relations between critical values of zeta functions and inner products,* Amer. J. Math. **105** (1983), 253–285.

[V] D. Vogan, Jr., *Representations of real reductive Lie groups,* Progress in Math., no. 15, Birkhaüser, Boston, 1981.

[VZ] D. A. Vogan, Jr. and G. Zuckerman, *Unitary representations with non-zero cohomology,* Compositio Math. **53** (1984), 51–90.

[W1] J.-L. Waldspurger, *Correspondance de Shimura,* J. Math. Pures Appl. **60** (1980), 1–132.

[W2] ——, *Quelques propriétés arithmétiques de certaines formes automorphes sur* GL(2), Compositio Math. **54** (1985), 121–171.

[W3] ——, *Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie,* Compostio Math. **54** (1985), 173–242.

[W4] ——, *Démonstration d'une conjecture de dualité de Howe dans le cas p-adique,* $p \neq 2$, Festschrift in honor of I. I. Piatetski-Shapiro, Israel Math. Conf. Proc., no. 2, Weizmann, Jerusalem, 1990, pp. 267–324.

[We] A. Weil, *Sur la formule de Siegel dans le théorie des groupes classiques,* Acta Math. **113** (1965), 1–87.

[Y] H. Yoshida, *On the zeta functions of Shimura varieties and periods of Hilbert modular forms,* manuscript, 1992.

[Z1] S. Zucker, *Locally homogeneous variations of Hodge structure,* Enseign. Math. **27** (1981), 243–276.

[Z2] ——, *L^2-cohomology of warped products and arithmetic groups,* Invent. Math. **70** (1982), 169–218.

Brandeis University, Waltham, MA

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Galois Representations Congruent to Those Coming from Shimura Varieties

J. TILOUINE

0. Introduction

This paper presents a review of various methods of constructing the compatible system $\sigma(\pi) = (\sigma_\ell(\pi))_\ell$ of ℓ-adic Galois representations associated to an arithmetic automorphic representation π, by using congruences. Here, "associated to π" means "having the same L-function" and the notion of an arithmetic π is defined below. In some sense, these methods are palliatives to the absence of a direct construction involving the cohomology of a Shimura variety (probably after some ingenious use of Langlands's functoriality). Such an approach would provide better information than that encapsulated in a compatible system of ℓ-adic representations. Namely, it would give a (Grothendieck) motive M_π with the same L-function as π. The possibility of attaching a motive to an arithmetic automorphic representation has been raised by Langlands in **[17]** (see precise formulations for $\mathrm{GL}(n)$ in **[6]** and **[21]**). We want first to state in a relatively precise way the part of this conjecture we shall deal with. Let F be a number field, $F_{\mathbf{A}}$, F_f, F_∞ be the ring of adèles, resp. finite adèles, resp. adèles at infinity, of F. Let G be a reductive group defined over F. In the sequel, we assume that G splits over F, and that its dual group $\widehat{G}$ is provided with a faithful rational representation r into $\mathrm{GL}(n, \mathbb{C})$ (in fact, later, G will be $\mathrm{GL}(2)_{/F}$ or $\mathrm{GSp}(4)_{/\mathbb{Q}}$). An irreducible admissible representation π of $G(F_{\mathbf{A}})$ can be decomposed into $\pi_f \otimes \pi_\infty$, where π_f, resp. π_∞, is an irreducible admissible representation of $G(F_f)$, resp. of $G(F_\infty)$. For any Archimedean place of F, let F_v be the completion of F at v.

DEFINITION 0.1. An automorphic representation $\pi = \pi_f \otimes \pi_\infty$ is called *arithmetic* (algebraic, in **[6]**, or motivic, in **[21]**) if there exist (A_0)-type Hecke characters $\varphi_1, \dots, \varphi_n$ of F such that for any Archimedean place

1991 *Mathematics Subject Classification*. Primary 11Fxx.
This paper is in final form and no version of it will be submitted for publication elsewhere.

v of F, the representation $r \circ \sigma(\pi_v)\colon \overline{F}_v^\times \to \mathrm{GL}(n, \mathbb{C})$ obtained by composing r with the restriction to $\overline{F}_v^\times = \mathbb{C}^\times$ of the representation of the Weil group $\sigma(\pi_v)\colon W_{F_v} \to {}^L G$, (see Borel, [3]) associated to π_v is conjugated to $\mathrm{diag}(\varphi_{i,v})_{i=1,\dots,n}$.

REMARK. This definition depends on r; so does, in fact, the notion of automorphic L-function; but in our cases, r will be the obvious one. □

Let V be the $\mathbb{C}$-vector space of the representation π_f of $G(F_f)$; the vector space ${}^\sigma V = V \otimes_{\mathbb{C},\sigma^{-1}} \mathbb{C}$ defines another representation of $G(F_f)$, denoted by σ_{π_f}.

DEFINITION 0.2. An admissible irreducible automorphic representation π is said to be *defined over a field* $E \subset \mathbb{C}$ if any automorphism σ of $\mathbb{C}$ that fixes E "fixes π_f", that is, ${}^\sigma\pi_f \simeq \pi_f$.

See [6, §3] for details on this notion. In the case where π is arithmetic, one conjectures that the minimal E is a number field; this is proven if π is regular [6, Theorem 3.13].

CONJECTURE 0.3. To every arithmetic automorphic representation π on G over F defined over the number field E, one can attach a motive M defined over F, with multiplication by E, so that the L-functions of π (up to some twist of π) and M coincide.

This conjecture implies

CONJECTURE 0.4. For π as above, there exists a strictly compatible system $\sigma(\pi) = (\sigma_\ell(\pi))_\ell$ of continuous representations

$$\sigma_\ell(\pi)\colon \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}(n, E \otimes \mathbb{Q}_\ell)$$

such that $\sigma(\pi)$ and π have the same L-function (up to some twist).

For the definition of a strictly compatible system of "-adic" representations, see [25, 1.2.3].

In very few cases, Conjecture 0.3 is proven (see below), but in most cases, even Conjecture 0.4 is still out of reach. We want to explain how one can construct, for some specific G, F, and π, the system $\sigma(\pi)$ (or sometimes only the p-adic representation $\sigma_p(\pi)$ for some primes p), although one is unable to construct the motive M_π. Let us formulate the general scheme of construction. It relies on the following principle whose statement is deliberately vague.

CONGRUENCE PRINCIPLE. *Let p be a rational prime. Let π and π' be arithmetic automorphic representations such that $\sigma_p(\pi)$ and $\sigma_p(\pi')$ exist. If π and π' are p-congruent, then $\sigma_p(\pi)$ and $\sigma_p(\pi')$ are p-congruent.*

The various meanings to be given to the expression "p-congruent" will be explained below. Then the proof runs in three steps. Fix G, F, and π.

STEP 1. Prove the existence of $\sigma(\pi')$ for a large class $\mathscr{C}$ of arithmetic automorphic representations π' of $G(F_{\mathbb{A}})$.

STEP 2. Find infinitely many representations π' in $\mathscr{C}$ that are p-congruent to π.

STEP 3. Patch together the $\sigma_p(\pi')$'s (which are mutually p-congruent by the Congruence Principle) and check that the resulting representation is the desired $\sigma_p(\pi)$.

There are two kinds of p-congruences: the horizontal ones (modulo infinitely many distinct primes) and the vertical ones (modulo p^n, $n = 1, \dots$). We order the ideas we want to review into three categories according to the type of p-congruences they involve.

I. Horizontal over $\operatorname{Spec}\mathbb{Z}$: Deligne-Serre, 1974 **[9]**, Rogawski-Tunnell, 1983 **[23]**.

II. Horizontal over $\operatorname{Spec}\Lambda$, $\Lambda = \mathbb{Z}_p[[T]]$: Hida, 1986 **[14]**, 1988 **[16]**, and Wiles, 1988 **[29]**.

III. Vertical: R. Taylor, 1988 **[27]**, 1990 **[28]**.

A last word concerning this approach. In II and III, a process of p-adic limit is involved when performing Step 3; therefore, even if the $\sigma_p(\pi')$'s do come from motives $M_{\pi'}$ as in our examples—in which case one says that the p-adic Galois representation $\sigma_p(\pi')$ is "geometric", nothing can be inferred concerning the geometricity of the limit $\sigma_p(\pi)$. We shall recall in §II an example found by Mazur and Wiles in 1986 **[20]**, showing that a p-adic limit of geometric $\sigma_p(\pi')$'s may not be Hodge-Tate. The question remains however whether something like a formal motive (a pro-object of a conjectural category of motives over Artinian $\mathbb{Z}_p$-algebras) could correspond in general to the representations constructed by this p-adic limit process.

NOTATION 0.5. When we consider an automorphic form f corresponding to the automorphic representation π, we write $\rho_p(f)$ instead of $\sigma_p(\pi)$.

I. Spec $\mathbb{Z}$ horizontal p-congruences

1.1. Deligne-Serre theorem. Take $G = \mathrm{GL}(2)_{/\mathbb{Q}}$. Recall that any cuspidal automorphic representation $\pi = \pi_f \otimes \pi_\infty$ such that π_∞ is holomorphic of weight $k \geq 1$ is arithmetic (cf. Clozel **[6**, p. 91**]**). Let π be as above, and moreover π_∞ in the limit of the holomorphic discrete series (that is, π_∞ has weight one). We are going to apply to π the three Steps 1–3 in order to obtain $\sigma(\pi)$. A cuspidal automorphic representation π' with π'_∞ in the holomorphic discrete series (resp. in the limit of the holomorphic discrete series) corresponds to a classical cusp eigenform f' of weight $k \geq 2$ (resp. $k = 1$), level N, character ε. Let $(b_n)_{n\geq 1}$ be the sequence of eigenvalues of f' for the Hecke operators T_n $(n = 1, 2, \dots)$. It is well known that the field $E = \mathbb{Q}((b_n)_n)$ generated by the eigenvalues b_n $(n = 1, 2, \dots)$ is a number field; it is the minimal field of definition for the representation π_f. The first step is carried out by Deligne's theorem **[8]** (completed by Langlands, Deligne, and Carayol at primes dividing the level).

THEOREM 1.1.1. *For any cuspidal π' in the holomorphic discrete series, there exists $\sigma(\pi')$ such that $L(\pi', s - (k-1)/2) = L(\sigma(\pi'), s)$. That is, for* *any prime ℓ, there exists a continuous representation*

$$\sigma_\ell(\pi'):\ \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, E \otimes \mathbb{Q}_\ell)$$

that is unramified outside ℓN, and such that for any prime $q \neq \ell$

$$\det(I_2 - \mathrm{Frob}_q X\,;\, \sigma_\ell(\pi')^I q) = 1 - b_q X + \varepsilon(q) q^{k-1} X^2.$$

In fact, one even knows that $\sigma(\pi')$ is geometric, that is, Conjecture 0.3 holds: Jannsen has proven that there is a motive $M_{\pi'}$ in the category $\mathscr{M}_{\mathrm{AH}}$ (the morphisms come from absolute Hodge cycles; cf. **[10]**), and Scholl has proven that $M_{\pi'}$ exists in the category of Grothendieck motives (the morphisms come from algebraic cycles, cf. **[24]**). Recall that the motive is cut off from the cohomology of the modular curve $X_1(N)$ (or, more precisely, from that of the desingularized Kuga-Sato variety $\mathbf{E}^{k-2}$ above $X_1(N)$) by means of Hecke correspondences. The motive $M_{\pi'}$ is defined over $\mathbb{Q}$, has rank two over E, and is pure of weight $k-1$.

For π of weight 1, one looks for an Artin motive M_π or, in other words, for an Artin representation $\sigma(\pi)$: $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, \mathbb{C})$ with the prescribed L-function. The second step of the construction can be summarized by the following diagram:

$$\begin{array}{lccccc}
\text{char.} = 0 & f & \xrightarrow{①} & g & \tilde{g} & \xrightarrow{③} \ \rho_p(\tilde{g}) \\
\text{reduction} & \downarrow & & \downarrow & \nearrow ② & \downarrow \\
\text{modulo}\,p & \overline{f} & = & \overline{g} & & \overline{\rho}_p(\tilde{g}).
\end{array}$$

DIAGRAM 1

EXPLANATION. The bar above a symbol means reduction mod p. We start with the holomorphic eigenform f, of weight 1, level N, character ε, eigenvalues $(a_n)_n$, corresponding to π. Put $E = \mathbb{Q}((a_n)_n)$. Choose an arbitrary rational prime number p that splits in E. There exists a modular form θ_p of level 1 and weight $n > 1$, whose Fourier coefficients belong to $\mathbb{Z}$ and are congruent to 1 modulo p (for instance, the Eisenstein series E_n for $n \equiv 0 \bmod (p-1)$. ① consists in multiplying f by θ_p in order to increase the weight. The resulting form $g = g_p = f\theta_p$ has weight $k = n+1 > 1$, and although it is no longer an eigenform, its reduction modulo p, $\overline{g}$, coincides with the mod p eigenform $\overline{f}$. Let $h_k(N, \mathbb{Z})$ be the ring generated by the Hecke operators T_n, $n = 1, 2, \ldots$, viewed as endomorphisms of the space of weight k, level N cusp forms. The mod p eigenform $\overline{g}$ gives rise to a character $\lambda_{\overline{g}}$: $h_k(N, \mathbb{Z}) \to \mathbb{F}_p$, $T_n \mapsto \overline{a}_n$. ② amounts to the Deligne-Serre lifting lemma, which states that such a character lifts to characteristic zero: there exists a cusp eigenform $\tilde{g}$ of weight k and level N, with eigenvalues $(b_n)_n$ in a finite extension of $\mathbb{Z}$, such that $\overline{b}_n = \overline{a}_n$ for any $n \geq 1$. ③ is nothing but Deligne's Theorem 1.1.1: since $\tilde{g}$ has weight $k \geq 2$, there exists a p-adic Galois representation $\rho_p(\tilde{g})$ with prescribed ramification set and characteristic polynomial. Reducing $\bmod p$, we find, for any such p, a Galois representation

$$\overline{\rho}_p(\widetilde{g})\colon \operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \operatorname{GL}(2, \mathbb{F}_p)$$

unramified outside Np and such that for any prime q relatively prime to Np,

$$\det(I_2 - \operatorname{Frob}_q X\,;\, \overline{\rho}_p(\widetilde{g})) = 1 - \overline{a}_q X + \overline{\varepsilon}(q) q X^2.$$

Therefore we found the desired class of π' of Step 2: they are the representations corresponding to the $\widetilde{g}_p$ for p varying, and $\sigma_p(\pi') = \rho_p(\widetilde{g})$.

REMARK. Instead of a form θ_p of level 1 and weight $n > 1$, one could use a form of level p and weight 1, so that $g = f\theta_p$ would be of weight 2. In that case, in order to produce $\rho_p(\widetilde{g})$, one can replace the powerful result of Deligne by the easier Eichler-Shimura relations for Hecke correspondences on the Jacobian of the modular curve. This trick has been used by Koike **[11]** and is susceptible of generalizations to the Siegel modular case (cf. Blasius and Ramakrishnan **[1]**).

Step 3 is the simplest kind of patching one can dream: one shows that for p sufficiently large

(i) $\overline{\rho}_p(\widetilde{g})$ lifts to $\operatorname{GL}(2, \mathbb{C})$ (because by an analytic argument, one sees that the image of $\overline{\rho}_p(\widetilde{g})$ has bounded order, hence is prime to p, for p large);

(ii) this lifting is constant (up to conjugacy).

This is the desired $\sigma(\pi)$. □

1.2. Rogawski-Tunnell theorem. In the case studied by Rogawski and Tunnell, G is GL(2) over a totally real field F of degree d. Recall that any cuspidal automorphic representation $\pi = \pi_f \otimes (\bigotimes_{v|\infty} \pi_v)$ such that for any $v|\infty$, π_v is holomorphic of weight $k_v \geq 1$ and for any v, w above ∞, $k_v \equiv k_w \bmod 2$, is arithmetic (see **[6,** 1.2.3]). We introduce a condition $(^*)$ by requiring

(i) nothing if d is odd;

(ii) that there exists a finite place v of F such that π'_v is square integrable, if d is even.

Let π as above be such that for any $v|\infty$, π_v is in the limit of the holomorphic discrete series (that is, has weight $k_v = 1$). Under the assumption $(^*)$, Rogawski and Tunnell attach to π an Artin representation $\sigma(\pi)\colon \operatorname{Gal}(\overline{F}/F) \to \operatorname{GL}(2, \mathbb{C})$ (that is, an Artin motive) with the prescribed L-function, by following closely the scheme of proof of Deligne-Serre. The first step, i.e. the fact that there is a "large class" $\mathscr{C}$ of representations π' for which $\sigma(\pi')$ exists, is provided by Carayol's theorem **[5]**.

THEOREM 1.2.1. *For any cuspidal automorphic representation $\pi' = \pi'_f \otimes (\bigotimes_{v|\infty} \pi'_v)$ such that for any $v|\infty$, $k_v \geq 2$ and $k_v \equiv k_w \bmod 2$, satisfying the condition $(^*)$, the system $\sigma(\pi')$ exists.*

Then, Steps 2 and 3 are very similar (although more technical) to those of Deligne-Serre. In particular, given an arbitrary prime p, the argument of increasing the weight of f by multiplying by a θ_p congruent to $1 \bmod p$, is possible thanks to a construction of θ_p by N. Katz using the irreducibility of the Hilbert moduli scheme over $\mathbb{Z}$. An important point in §2 is to show that assumption $(^*)$ is preserved by congruences, so that the forms $\tilde{g}$ one obtains do admit a Galois representation $\rho_p(\tilde{g})$. □

1.3. Comments. Before considering the second kind of p-congruences, some remarks are in order. The *lifting* process in Step 3 above was adequate because we were looking for an Artin representation, that is, a representation with finite image. In the Hilbert modular case, this was particular to the weight 1 case. For higher weights, since the expected Galois representations must have infinite image, it is natural to replace *liftings* by *deformations* from characteristic p to characteristic zero, of a $\bmod p$ Galois representation. This leads to consider a different kind of p-congruences. In fact, the starting point is due to B. Mazur, who observed around 1985 that Hida's theory of analytic families of (ordinary) eigenforms **[13, 14]** provides an adequate tool to deal with deformations of Galois representations.

II. Spec Λ horizontal p-congruences

II.1. Main result of Hida theory. Let $G = \mathrm{GL}(2)$ over a totally real field F; p a rational prime. Fix an embedding $\overline{\mathbb{Q}} \subset \overline{\mathbb{Q}}_p$. For the sake of simplicity, we write down the formulas only for $F = \mathbb{Q}$, although everything is similar for any F (see Hida **[14]** and Wiles **[29]**). A cusp eigenform f with eigenvalues $(a_n)_n$ is p-ordinary if a_p is a p-adic unit.

REMARK. When $F \neq \mathbb{Q}$, one must think of the level N and of the indexes n of the Fourier coefficients of f, as integral ideals of F; p, however, is a rational prime.

Let $\mathscr{O}$ be a finite extension of $\mathbb{Z}_p$ containing the a_n's. Let $\Lambda = \mathscr{O}[[T]]$ be the Iwasawa algebra over $\mathscr{O}$.

THEOREM II.1.1. *Let f be a p-ordinary cusp eigenform of type (k, N, ε), $k \geq 2$. There exists an analytic family of p-ordinary cusp eigenforms (abbreviated as AFOCE; it is also called Λ-adic eigenform) of level N, that is a formal power series in q, with coefficients in Λ (or in a finite extension thereof, assumed to be Λ for simplicity):*

$$\mathbf{F} = \sum_{n\geq 1} A_n(T) q^n, \qquad A_n(T) \in \Lambda,$$

such that for any $\ell \geq 2$, $\mathbf{F}_\ell = \sum_{n\geq 1} A_n((1+p)^\ell - 1)q^n$ is a cusp eigenform of type $(\ell, \mathrm{l.c.m.}(N, p)$, some character), and

$$\mathbf{F}_k = \begin{cases} f & \text{if } p|N, \\ f_0 & \text{if } (p, N) = 1, \end{cases}$$

where f_0 is defined as follows. Let

$$1 - a_p X + \varepsilon(p) p^{k-1} X^2 = (1 - \alpha_p X)(1 - \beta_p X),$$

with $|\alpha_p|_p = 1$, $|\beta_p|_p < 1$; *then,* $f_0(z) = f(z) - \beta_p f(pz)$.

IDEA OF THE PROOF. It is a simple application of the Deligne-Serre lifting lemma (namely, the going-down theorem for finite flat algebras) to the big p-adic ordinary Hecke algebra. The main work is to establish the properties of this algebra. Fix $k \geq 2$; define the big p-adic Hecke algebra $\mathbf{h}(N)$ of level N as the inverse limit of the finite free $\mathscr{O}$-algebras $h_k(Np^r; \mathscr{O}) = h_k(Np^r; \mathbb{Z}) \otimes \mathscr{O}$. The transition map $\varphi_{r,s}$ $(r > s > 0)$ is induced by the inclusion of level Np^s forms into level Np^r forms. One can prove (cf. **[15,** §11]) that $\mathbf{h}(N)$ does not depend on $k \geq 2$. Its ordinary part $\mathbf{h}^0(N)$ is defined as the largest quotient of $\mathbf{h}(N)$ in which the image of T_p is invertible. Take the inverse limit χ of the well-known characters $\langle\ \rangle_k \colon (\mathbb{Z}/Np^r\mathbb{Z})^\times \to h_k(Np^r; \mathscr{O})$ given on a prime q by $q^{1-k}((T_q)^2 - T_{q^2})$. Twisting χ by the kth power of the cyclotomic character we get a continuous character

$$\langle\ \rangle \colon \operatorname*{inv.lim}_{r \geq 1} (\mathbb{Z}/Np^r\mathbb{Z})^\times \to \mathbf{h}(N);$$

hence, in particular, a structure of Λ-algebra on $\mathbf{h}(N)$ and on $\mathbf{h}^0(N)$. The twist is necessary because of the way we stated Theorem II.1.1.

THEOREM II.1.2. *$\mathbf{h}^0(N)$ is a finite free Λ-algebra.*

Assume then, for example, that $(p, N) = 1$. The form f_0 gives rise to a character $\lambda_{f_0} \colon \mathbf{h}_k(Np; \mathscr{O}) \to \mathscr{O}$, $T_n \mapsto a_n$ if $(p, n) = 1$, and $T_p \mapsto \alpha_p$. Compose it with the projection $\mathbf{h}^0(N) \to \mathbf{h}_k(Np; \mathscr{O})$. By flatness of $\mathbf{h}^0$ over Λ (Theorem II.1.2) and by the going-down theorem, the resulting character lifts to a finite extension of Λ, say Λ itself: $\lambda_{\mathbf{F}} \colon \mathbf{h}^0 \to \Lambda$, $T_n \mapsto A_n(T)$. □

II.2. Deformations and pseudo-representations. Let $\mathscr{L}$ be the fraction field of Λ. For any prime q of F prime to Np, put $\langle q \rangle = q((T_q)^2 - T_{q^2})$. The relation between Hida theory and the deformations of Galois representations is provided by the following

THEOREM II.2.1 (Hida **[14,** Theorem 3]). *Let* $\mathbf{F}$ *be an AFOCE such that* (*) *holds for any weight* $\ell \geq 2$ (*we say* $\mathbf{F}$ *satisfies* (*)); *then, there exists a continuous representation*

$$\rho_p(\mathbf{F}) \colon \operatorname{Gal}(\overline{F}/F) \to \operatorname{GL}(2, \mathscr{L})$$

unramified outside Np *and such that for any prime* q *relatively prime to* Np,

$$\det(I_2 - \operatorname{Frob}_q X; \rho_p(\mathbf{F})) = 1 - A_q(T) X + q^{-1} \lambda_{\mathbf{F}}(\langle q \rangle) X^2.$$

REMARK. One can show that if (*) holds for some $k \geq 2$, it holds for all weights ≥ 2.

SKETCH OF PROOF. We follow the elegant proof of Wiles [29], because it involves a very important idea, which has proved to be very fruitful lately, that of pseudo-representation (of degree 2).

Let A be a ring, G a group, and $c \in G$ an element of order two.

DEFINITION II.2.2. A *pseudo-representation* of G defined over A is a quadruple $\tau = (a, d, t, x)$; $a, d, t: G \to A$, $x: G \times G \to A$ satisfying the conditions

(i) $2a_{g,g'} = a_g a_{g'} + x_{g,g'}$, $2d_{g,g'} = d_g d_{g'} + x_{g',g}$;
(ii) $a_g = t_g + t_{cg}$, $d_g = t_g - t_{cg}$;
(iii) $t_1 = 2$, $t_c = 0$, $x_{g,c} = x_{c,g} = 0$;
(iv) $x_{g,g'}x_{h,h'} = x_{g,h'}x_{h,g'}$, $4x_{gh,g'h'} = a_g a_{h'} x_{h,g'} + a_{h'} d_h x_{g,g'} + a_g d_{g'} x_{h,h'} + d_h d_{g'} x_{g,h'}$.

Define the *trace* of τ by $\mathrm{tr}(\tau) = t$, and the *determinant* by $\det(\tau)(g) = a_g d_g - x_{g,g}$.

COMMENTS. (i) If G and A have a topology, all the maps are assumed to be continuous.

(ii) If $\rho: G \to \mathrm{GL}(2, A)$, $\rho = (\alpha, \beta, \gamma, \delta)$ is a representation defined over A, with $\rho(c) = (1, 0, 0, -1)$, it gives rise to an obvious pseudo-representation $\tau = (a, d, t, x)$, where $a = 2\alpha$, $d = 2\delta$, $t = \alpha + \delta$, $x_{g,g'} = 4\beta_g\gamma_{g'}$.

(iii) Conversely, if A is a **field** of characteristic different from 2, any pseudo-representation τ corresponds to a representation ρ such that $\rho(c) = (1, 0, 0, -1)$.

KEY POINT. By definition, a pseudo-representation is determined by its trace. In particular, if $A' \subset A$ is a subring of A, if τ is defined over A and $\mathrm{tr}(\tau)$ takes values in A', then τ is defined over A'.

This is the main difference with the notion of representation. It can be reformulated as the

PATCHING LEMMA II.2.3 (Wiles [29]). *Let $(\mathfrak{A}_i)_{i\geq 1}$ be a sequence of ideals in A such that $\bigcap_i \mathfrak{A}_i = \{0\}$. For each i, let τ_i be a pseudo-representation defined over $A/\mathfrak{A}_i$. Assume there exists a dense subset Σ of G and a function $t: \Sigma \to A$ such that for any i, $t \bmod \mathfrak{A}_i = \mathrm{tr}(\tau_i)$. Then, there exists a pseudo-representation τ defined over A such that for any i, $\tau \bmod \mathfrak{A}_i = \tau_i$.*

Apply this lemma to $G = \mathrm{Gal}(\overline{F}/F)$, $A = \mathbf{\Lambda}$, $\mathfrak{A}_\ell = ((1+T) - (1+p)^\ell)$ for $\ell \geq 2$, $\Sigma = (\mathrm{Frob}_q;\, q$ prime of F, $(q, Np) = 1\}$, $t: \Sigma \to \mathbf{\Lambda}$ defined by $q \mapsto A_q(T)$. Start from $\mathbf{F}$. For any $\ell \geq 2$, since $\mathbf{F}_\ell$ satisfies $(^*)$, $\rho_p(\mathbf{F}_l)$ exists. Let $\tau_\ell = \tau_p(\mathbf{F}_\ell)$ be the corresponding pseudo-representation. It is defined over $\mathscr{O} = \mathbf{\Lambda}/\mathfrak{A}_\ell$. These τ_ℓ satisfy the patching condition of Lemma II.2.3; therefore, they glue into $\tau(\mathbf{F})$ defined over $\mathbf{\Lambda}$, which gives rise to the desired representation over $\mathscr{L}$. □

II.3. Mazur-Wiles's example. Let $\mathbf{F}$ be an AFOCE. By construction, for any integer $\ell \geq 2$, $\rho_p(\mathbf{F}) \bmod \mathfrak{A}_\ell$ is isomorphic to the Deligne representation attached to $\mathbf{F}_\ell$. However, for $\ell = 1$, as observed by Mazur-Wiles, something strange may happen. Consider the unique AFOCE $\mathbf{\Delta}$ of level 1 interpolating Ramanujan Δ-function in weight 12. Recall that Mazur-Wiles **[20]** have observed that if $p \neq 11, 23, 691$ and $\tau(p) \not\equiv 0 \bmod p$, the Hida representation

$$\rho_p(\mathbf{\Delta})\colon \operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \operatorname{GL}(2, \mathbb{Z}_p[[T]])$$

is full (that is, its image contains $\operatorname{SL}(2, \mathbb{Z}_p[[T]])$). It is a consequence of the well-known result of **[26]**, together with a group-theoretic lemma of N. Boston **[20,** Appendix]. The same holds for the reduction ρ_1 of $\rho_p(\mathbf{\Delta})$ in weight 1. Since ρ_1 has infinite image, we already see that it does not come from a (pure) motive (it should be an Artin motive). It is even believed that ρ_1 does not fit in a compatible system of ℓ-adic representations. In fact, although $\det \rho_1$ is a finite order character, the algebraicity of $\operatorname{tr}\rho_1(\mathrm{Frob}_q) \in \mathbb{Z}_p$ (for $q \neq p$) is not clear. The main pathology of ρ_1 is that for $p = 13, 17, 19$, its restriction $\rho_{1,p}$ to $D_p = \operatorname{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ is not Hodge-Tate. This is equivalent to saying $\rho_{1,p}$ is not semisimple. It can be seen as follows. Let V be the space of ρ_1; it is unramified outside p; recall that $\rho_p(\mathbf{\Delta})$ is p-ordinary (see **[14, 16]**). It implies there is a D_p-equivariant exact sequence

$$0 \to \mathbb{Q}_p(\omega^{11}) \to V \to \mathbb{Q}_p \to 0,$$

where the $\mathbb{Q}_p$ on the right is unramified (Frob_p acts on it by $\tau(p)$, which is not of finite order in $\mathbb{Z} \subset \mathbb{Z}_p$), and the first is finitely ramified (I_p acts by ω^{11}). If V were semisimple (thus, Hodge-Tate of type $(0, 0)$ over $\mathbb{Q}_p(\zeta_p)$), the action of inertia would factor through $\operatorname{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. There should exist an **infinite** Galois extension $L/\mathbb{Q}$ unramified over $\mathbb{Q}(\zeta_p)$. On the other hand, for $p \leq 19$, Odlyzko has given a bound for the degree of a Galois extension $L/\mathbb{Q}$ unramified over $\mathbb{Q}(\zeta_p)$. Contradiction. □

II.4. Wiles's theorem. Let us see now how Wiles, using Theorem II.2.1, was able to remove the assumption $(^*)$ for the existence of $\rho_p(f)$, at least for p-ordinary f's. Assume $d = [F : \mathbb{Q}]$ is even, and the cusp eigenform f is p-ordinary but does not satisfy $(^*)$. By Theorem I.1.1, there exists an AFOCE, say $\mathbf{F}$, which interpolates f. The key of the construction, namely the creation of a congruence between $\mathbf{F}$ and another AFOCE $\mathbf{G}$ that satisfies $(^*)$ appeared also in Ribet's ICM talk in Warsaw **[22]**. It stems from the remark that if g is a form of level Nq (q prime, $(q, N) = 1$), which is q-new and whose Nebentypus is q-trivial, the corresponding automorphic representation π' is special at q if the weight k is ≥ 2, and is supercuspidal at q if $k = 1$; in particular, π' satisfies $(^*)$. The process can be summarized by the following diagram:

$$\tilde{\mathbf{h}}^0(Nq) \nearrow \tilde{\mathbf{h}}^0(Nq)^{q\text{-new}} \searrow \Lambda, \quad \tilde{\mathbf{h}}^0(Nq) \searrow \mathbf{h}^0(N) \xrightarrow{\lambda_{\mathbf{F}}} \Lambda \qquad\qquad \tilde{\mathbf{h}}^0(Nq) \nearrow \tilde{\mathbf{h}}^0(Nq)^{q\text{-new}} \searrow \Lambda/I(q), \quad \tilde{\mathbf{h}}^0(Nq) \searrow \mathbf{h}^0(N) \xrightarrow{\lambda_{\mathbf{F}}} \Lambda/I(q).$$

Diagram 2

EXPLANATION. q is any prime in F relatively prime to Np, and $\tilde{\mathbf{h}}(Nq)$ is the subalgebra of $\mathbf{h}(Nq)$ generated by the Hecke operators T_n for $(n, q) = 1$. In both diagrams, the left arrows are the obvious homomorphisms deduced from the inclusion of level N forms, resp. q-new level Nq forms, into level Nq forms. The incomplete diagram on the left means that, because of the multiplicity one theorem (proven by Hida in this context **[14**, Theorem 3]), it is impossible to find a character making it commutative (it would correspond to a Λ-adic eigenform which is both q-new and q-old). On the other hand, such a character exists modulo $I(q)$, where

$$\Lambda/I(q) \simeq \mathbf{h}^0(Nq)^{q\text{-new}} \otimes_{\tilde{\mathbf{h}}^0(Nq)} \tilde{\mathbf{h}}^0(N) \otimes_{\tilde{\mathbf{h}}^0(N), \lambda_{\mathbf{F}}} \Lambda$$

is the so-called module of congruences between $\mathbf{F}$ and q-new Λ-adic eigenforms. Let $w(q) = \lambda_{\mathbf{F}}(T_q^2 - \langle q \rangle(1 + \mathrm{Norm}(q)^{-1}))$; it belongs to $I(q)$.

Wiles shows that

(W1) when q varies, the divisor of $I(q)w(q)^{-1}$ is bounded.

(W2) The set $\{w(q);\ q \text{ prime, } (q, Np) = 1\}$ has an unbounded set of height one prime divisors. This fact is an application of Chebotarev theorem to the Brylinski-Labesse representation (of degree 2^d; cf. **[4]**).

This carries out Step 2 in Wiles's method:

FACT II.4.1. *There exist infinitely many triples* $(q_i, P_i, \mathbf{G}_i)_i$, *where* q_i *is a prime in* F, $(q_i, Np) = 1$, P_i *is a height one prime in* Λ, $\mathbf{G}_i$ *is an AFOCE that is* q_i*-new (with trivial Nebentypus at* q_i*), such that* $\lambda_{\mathbf{F}} \equiv \lambda_{\mathbf{G}_i} \bmod P_i$.

Then, Step 3 is a simple application of the patching lemma: since $\mathbf{G}_i$ is special at q_i, it has a Hida representation $\rho_p(\mathbf{G}_i)$, hence a pseudo-representation $\tau_p(\mathbf{G}_i)$ defined over Λ. Reduce it $\bmod P_i$. Note that $\bigcap_i (P_i) = \{0\}$. One can apply the patching lemma with the map $t\colon \{\mathrm{Frob}_\ell\} \to \Lambda$, $\mathrm{Frob}_\ell \mapsto A_\ell(T)$ (after a trivial modification, because for each i, $\mathrm{tr}(\tau_i) \equiv t$ only on $\Sigma \setminus \{\mathrm{Frob}_{q_i}\}$). One gets $\tau_p(\mathbf{F})$, hence $\tau_p(f)$, by specializing to the weight k. This gives $\rho_p(f)$ over the field of fractions of $\mathscr{O}$. $\square$

REMARK. If f has weight $(1, \dots, 1)$, that is, for any $v|\infty$, π_v is in the limit of the holomorphic discrete series, f is obviously ordinary at any prime p; therefore Wiles's method gives, in particular, a generalization of the Rogawski-Tunnell theorem by dropping assumption $(^*)$ for the existence of $\sigma(\pi)$ in the weight one case.

II.5. Comments. Finally, in 1988, R. Taylor **[27]** got rid of any assumption about p in the Hilbert modular case (in weight (k_v), $k_v \geq 2$) thus constructing $\sigma(\pi)$ for any automorphic cusp π such that

(i) for each $v|\infty$, π_v is in the holomorphic discrete series;

(ii) the corresponding weights k_v $(v|\infty)$ have the same parity.

In fact, all the necessary ingredients were essentially contained in Wiles's paper; Taylor's observation was that they can be brought into play without resorting to Hida families, i.e., by dealing only with the form f instead of introducing $\mathbf{F}$. This requires, however, another kind of p-congruences, namely the vertical ones; it has already proved to be successful in the $\mathrm{GSp}(4, \mathbb{Q})$-case, as we shall see.

III. Vertical p-congruences

III.1. Hilbert modular case in general. Let us see how Taylor carries out Steps 2 and 3 for a general Hilbert cusp eigenform f of weight $(k_v)_{v|\infty}$, $k_v \geq 2$, $k_v \equiv k_w \bmod 2$. Let us first draw the exact analogue of Diagram 2 for Hecke algebras of finite level, with $\lambda_f \colon \mathbf{h}_k(N; \mathscr{O}) \to \mathscr{O}$, $T_n \mapsto a_n$ replacing $\lambda_{\mathbf{F}}$.

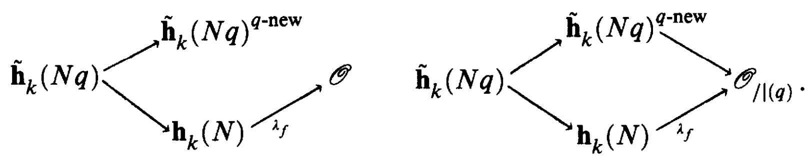


DIAGRAM 3

Step 2 consists in two estimates. Let $\ell(q)$ be the length of the congruence module between f and q-new eigenforms, and let $w(q) = a_q^2 - \varepsilon(q)q^{k-1}(1 + \mathrm{Norm}(q)) \in \mathscr{O}$. The first estimate is that $\ell(q) - \mathrm{ord}(w(q))$ is bounded in q. Let ϖ be a uniformizer in $\mathscr{O}$. The second estimate, again an application of the Chebotarev theorem to the Brylinski-Labesse representation, states that for any $i \geq 1$, there exists a prime q (outside any given finite set of primes) such that $w(q) \equiv 0 \bmod \varpi^i$. From this, we conclude

FACT III.1.1. *There exists a sequence* $(q_i, \lambda_i)_{i \geq 1}$ *of pairs, where* q_i *is a prime in* F *prime to* Np, *and* $\lambda_i \colon \tilde{\mathbf{h}}_k(Nq_i; \mathscr{O})^{q_i\text{-new}} \to \mathscr{O}/(\varpi^i)$ *is a character* $\bmod\, \varpi^i$, *such that* $\lambda_i \equiv \lambda_f \bmod \varpi^i$.

REMARK. The main difference between horizontal and vertical congruences is that in the vertical case, the Deligne-Serre lifting lemma (i.e., the going-down theorem) does not apply, i.e., the characters Λ_i $(i > 1)$ above do not necessarily lift to characteristic zero. Here is a simple counterexample.

Take $A = \{(x, y) \in \mathbb{Z}_p \times \mathbb{Z}_p ; x \equiv y \bmod p\} \subset \mathbb{Z}_p \times \mathbb{Z}_p$. This is a local algebra with only two characteristic zero characters, namely $(x, y) \mapsto x$ and $(x, y) \mapsto y$. Let $\varepsilon = (p, 0)$; one has $A = \mathbb{Z}_p[\varepsilon] = \mathbb{Z}_p[X]/X(p - X)$. The maximal ideal is $\mathfrak{M} = (p, \varepsilon)$, hence $A/\mathfrak{M}^2 = \mathbb{Z}_p[X]/(p, X)^2$. So, $\mathrm{Hom}_{\mathrm{alg}}(A, \mathbb{Z}/p^2\mathbb{Z}) \simeq \mathbb{Z}/p\mathbb{Z}$ by $\varphi \mapsto \xi$, where $\varphi(\varepsilon) = p\xi$. If $p > 2$, there is a character $\bmod p^2$ that does not lift. Because of this remark, the patching process of Step 3 is performed in a way slightly different from Wiles's. In fact, let $\Sigma = \{\mathrm{Frob}_\ell ; \ell \text{ prime}, (\ell, Np) = 1\}$, $\tilde{\mathbf{h}}_k(Nq; \mathscr{O})^{q\text{-new}}$ embeds into a product $\prod \mathfrak{R}_i$ of local $\mathbb{Q}_p$-algebras. One shows by using Carayol's Theorem I.2 that there exists a pseudo-representation τ on $\prod \mathfrak{R}_i$ such that for any $g = \mathrm{Frob}_\ell \in \Sigma$, $\ell \neq q$,

$$\mathrm{tr}(\tau)(g) = T_\ell \quad \text{and} \quad \det(\tau)(g) = \langle \ell \rangle_k \, \mathrm{Norm}(\ell)^{k-1}.$$

This pseudo-representation is defined over $\mathbf{h}_k(Nq; \mathscr{O})^{q\text{-new}}$; therefore, one can push it forward by λ_i. By this process, one gets a sequence of pseudo-representations $(\tau_i)_i$ over $\mathscr{O}/(\varpi^i)$ satisfying the patching condition. They patch into $\tau_p(f)$ which gives rise to the desired representation $\rho_p(f)$ over the field of fractions of $\mathscr{O}$. □

COMMENT. This p-adic construction does not show that the restriction to the local Galois group $D_v = \mathrm{Gal}(\overline{F}_v/F_v)$ (v dividing p) of the representation $\rho_p(f)$ is Hodge-Tate. However, the existence of the motive underlying the system $(\rho_p(f))_p$ has been directly proven, after Taylor's result, by Blasius-Rogawski **[2]**, using Langlands's functoriality. It implies that $\rho_p(f)|D_v$ is Hodge-Tate. □

III.2. The $\mathrm{GSp}(4, \mathbb{Q})$ **case.** Take $G = \mathrm{GSp}(4, \mathbb{Q})$, $\pi = \pi_f \otimes \pi_\infty$, where π_∞ is in the limit of the holomorphic discrete series. Taylor **[28]** has obtained a good approximation of Conjecture 0.4 for such a π (see below for a precise statement); his method is very similar to the Hilbert modular one. Step 1 is provided by the following theorem.

Let $\pi' = \pi'_f \otimes \pi'_\infty$ be a cuspidal automorphic representation such that π'_∞ is in the holomorphic discrete series. Let S be the ramification set of π'_f (the set of rational primes q such that there is no nonzero $\mathrm{GSp}(4, \mathbb{Z}_q)$-fixed vector in π'_q). For any $q \notin S$, let $P_q(X)$ be the degree 4 polynomial such that $P_q(q^{-s})^{-1}$ is the q-Euler factor of $L(\pi', s - w/2)$ (for some integer w; for its definition see **[28,** §3.1]).

THEOREM III.2.1 (Shimura, Deligne, Chai-Faltings). *For any π' as above, there exists a fixed number field $E_{\pi'}$ such that for any $q \notin S$, $P_q(X) \in E_{\pi'}[X]$. Moreover, let $\alpha_q, \beta_q, \gamma_q, \delta_q \in \overline{\mathbb{Q}}$ be the reciprocal roots of $P_q(X)$. There exist an integer $m \geq 4$ and a system $\tilde{\sigma}(\pi')$ of ℓ-adic representations*

$$\tilde{\sigma}_\ell(\pi') \colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(m, \overline{\mathbb{Q}}_\ell)$$

such that for any $q \notin S$, the eigenvalues of $\tilde{\sigma}(\pi')(\mathrm{Frob}_q)$ are in $\{\alpha_q, \beta_q, \gamma_q, \delta_q\}$.

COMMENTS. It is conjectured that $m = 4$ and $P_q(X)$ is exactly the characteristic polynomial of $\tilde{\sigma}(\pi')(\mathrm{Frob}_q)$. □

THEOREM III.2.2 (R. Taylor). *For $\pi = \pi_f \otimes \pi_\infty$ with π_∞ in the limit of the holomorphic discrete series, the number field E_π as above exists, and for any $\ell \neq 2, 3$, $\tilde{\sigma}_\ell(\pi)$ as above exists.*

BRIEF SKETCH OF PROOF. First, one sees directly the existence of the number field E_π. Then, Step 2, as in Deligne-Serre, relies on the trick of increasing the weight in order to be brought back to the holomorphic discrete series, for which Theorem III.2.1 applies, except that, in order to get vertical p-congruences, one has to replace θ_p (defined by Blasius and Ramakrishnan [1], using Chai and Faltings results [7]) by its p^nth power, which is congruent to $1 \bmod p^{n+1}$. Then, for Step 3, as in III.1, one has to patch pseudo-representations $\bmod \varpi^i$ into a pseudo-representation defined over $\mathscr{O}$ (the p-adic completion of E_π). For this, since we deal with m-dimensional representations, one needs a generalization of the notion of pseudo-representation to any degree. Taylor defines it as a single map $T: G \to A$ satisfying axioms inspired from the invariant-theoretical characterization of the trace on $\mathrm{GL}(m)$ over an algebraically closed field of characteristic zero (see **[28,** §1]). With this definition, the patching lemma applies (§1.3, Example 2) and we get the desired representation. □

COMMENTS. (i) One does not know whether it is Hodge-Tate.

(ii) This existence theorem is used in a recent work by Harris, Soudry, and Taylor **[12, 30]** to prove the existence of a compatible system of ℓ-adic representations attached to an automorphic representation on GL(2) over an imaginary quadratic field.

REFERENCES

1. D. Blasius and D. Ramakrishnan, *Maass forms and Galois representations*, Galois Group over $\mathbb{Q}$, Academic Press, New York, 1990.
2. D. Blasius and J. Rogawski, *Galois representations for Hilbert modular forms*, Bull. Amer. Math. Soc. (N. S.) **21** (1989), 65–69.
3. A. Borel, *Automorphic L-functions*, Proc. Sympos. Pure Math. **33** (1979), part 2, pp. 27–61.
4. J.-L. Brylinski and J.-P. Labesse, *Cohomologie d'intersection et fonctions L de certaines variétés de Shimura*, Ann. Sci. École Norm. Sup. (4) **17** (1984), 361–412.
5. H. Carayol, *Sur les représentations l-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. École Norm. Sup. (4) **19** (1986), 409–468.
6. L. Clozel, *Motifs et formes automorphes: Applications du principe de fonctorialité*, Automorphic Forms, Shimura Varieties, and L-Functions, Academic Press, New York, 1990.
7. C. L. Chai and G. Faltings, *Degeneration of abelian varieties*, Ergeb. Math. Grenz. (3) no. 22, Springer-Verlag, Berlin, 1990.
8. P. Deligne, *Formes modulaires et représentations l-adiques*, Sém. Bourbaki, 1968–69, exposé no. 355.
9. P. Deligne and J.-P. Serre, *Formes modulaires de poids* 1, Ann. Sci. École Norm. Sup. (4) **7** (1974), 507–530.
10. U. Jannsen, *Mixed motives and algebraic K-theory*, Lecture Notes in Math., vol. 1400, Springer-Verlag, Berlin and New York, 1990.
11. M. Koike, *Congruences between cusp forms and linear representations of the Galois group*, Algebraic Number Theory (Internat. Sympos., Kyoto, 1976), Japan Soc. Promotion Sci., Tokyo, 1977, pp. 109–116.
12. M. Harris, D. Soudry, and R. Taylor, *l-Adic representations associated to modular forms over imaginary quadratic fields*, Invent. Math. **112** (1993), 377–411.
13. H. Hida, *Iwasawa modules attached to congruences of cusp forms*, Ann. Sci. École Norm. Sup. (4) **19** (1986), 214–273.

14. ——, *Galois representations into* $GL_2(\mathbb{Z}_p[[X]])$ *attached to ordinary cusp forms*, Invent. Math. **85** (1986), 546–613.
15. ——, *On p-adic Hecke algebras for* GL(2) *over totally real fields*, Ann. of Math. **128** (1988), 295–384.
16. ——, *Nearly ordinary Hecke algebras and Galois representations of several variables*, Amer. J. Math., Special Issue (1989).
17. R. P. Langlands, *Automorphic forms, Shimura varieties and motives: ein Märchen*, Proc. Sympos. Pure Math. **33** (1979), part 2, 205–246.
18. B. Mazur, *Deformations of Galois representations*, Galois Groups over $\mathbb{Q}$, Adv. Stud. Pure Math., no. 16, Academic Press, New York, 1990.
19. B. Mazur and J. Tilouine, *Représentations galoisiennes, différentielles de Kähler et conjectures principales*, Inst. Hautes Études Sci. Publ. Math., no. 71 (1990), 65–103.
20. B. Mazur and A. Wiles, *On p-adic analytic families of Galois representations*, Compositio Math. **59** (1986), 231–264.
21. D. Ramakrishnan, *Pure motives and automorphic forms*, these Proceedings, vol. 2, pp. 411–446.
22. K. Ribet, *Congruence relations between modular forms*, Proc. Internat. Cong. Math. (August 16–23, 1983, Warsaw), PWN and North-Holland, 1984.
23. J. Rogawski and J. Tunnell, *On Artin L-functions associated to Hilbert modular forms of weight* 1, Invent. Math. **74** (1983), 1–42.
24. A. Scholl, *Motives for modular forms*, Invent. Math. **100** (1990), 419–430.
25. J.-P. Serre, *Abelian* ℓ*-adic representations and elliptic curves*, Benjamin, 1968.
26. H. P. F. Swinnerton-Dyer, *On* ℓ*-adic representations and congruences for coefficients of modular forms*, Modular Functions of One Variable, III, Lecture Notes in Math., vol. 350, Springer-Verlag, Berlin, 1973, pp. 1–50.
27. R. Taylor, *On Galois representations associated to Hilbert modular forms*, Invent. Math. **98** (1989), 265–280.
28. ——, *Galois representations associated to Siegel modular forms of low weight*, Duke Math. J. **63** (1991), 281–332.
29. A. Wiles, *On nearly ordinary* λ*-adic representations associated to modular forms*, Invent. Math. **94** (1988), 529–573.

Added in proof.

30. R. Taylor, *l-Adic representations associated to modular forms over imaginary quadratic fields*, II, preprint.

EQUIPE D'ARITHMÉTIQUE ET GÉOMÉTRIE ALGÉBRIQUE, FACULTÉ DES SCIENCES D'ORSAY, FRANCE

Proceedings of Symposia in Pure Mathematics
Volume 55 (1994), Part 2

Report on mod ℓ Representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$

KENNETH A. RIBET

In memory of Kenneth F. Ireland

1. Introduction

Let $N \geq 1$ and $k \geq 2$ be integers. Let $\Gamma_1(N)$ be the group

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}(2, \mathbf{Z}) \;\middle|\; c \equiv 0,\ a \equiv d \equiv 1 \bmod N \right\},$$

and let $S = S_k(\Gamma_1(N))$ be the complex vector space of cusp forms of weight k on $\Gamma_1(N)$. There is a standard action $d \mapsto \langle d \rangle \in \mathrm{Aut}\, S$ of the group $(\mathbf{Z}/N\mathbf{Z})^*$ on S. In addition, S comes equipped with a family of Hecke operators T_n $(n \geq 1)$; the T_n commute with each other and with the "diamond bracket" operators $\langle d \rangle$. These operators are traditionally written on the right; for example, the n^{th} Hecke operator is normally written $f \mapsto f|T_n$.

An *eigenform* in S is a nonzero cusp form $f \in S$ which is an eigenvector for each of the operators T_n and $\langle d \rangle$. In particular, if f is an eigenform, then there is a Dirichlet character ε defined mod N such that $f|\langle d \rangle = \varepsilon(d) f$ for $d \in (\mathbf{Z}/N\mathbf{Z})^*$. The functional equation satisfied by a weight-k cusp form implies that this "Nebentypus" character has the same parity as k: $\varepsilon(-1) = (-1)^k$. The eigenvalues a_n of the T_n acting on f are algebraic integers and generate together a finite extension of $\mathbf{Q}$.

Let $\overline{\mathbf{Q}}$ be an algebraic closure of the field of rational numbers, and consider the Galois group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Suppose that $f \in S$ is an eigenform, and let E be a number field containing the associated eigenvalues a_n and $\varepsilon(d)$. A construction of Deligne **[10]** attaches to f a family of continuous

1991 *Mathematics Subject Classification.* Primary 11F11, 11R32; Secondary 11F33, 11G18.
This papèr is in final form and no version of it will be submitted for publication elsewhere.
The author was supported in part by NSF Grant #DMS 88-06815. He wishes to thank both the Institut des Hautes Études Scientifiques and the Université de Paris 7 for gracious hospitality during the period when this article was begun. Thanks are due to Fred Diamond, Bas Edixhoven, and Barry Mazur for helpful comments. The referee deserves special thanks for a thorough reading of the manuscript and for many useful comments and suggestions.

representations

$$\rho_\lambda\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, E_\lambda);$$

here, λ runs over the set of prime ideals of the integer ring of E, and E_λ denotes the completion of E at λ. Each representation ρ_λ is irreducible; it is characterized up to isomorphism by the identities

$$\operatorname{trace}(\rho_\lambda(\operatorname{Frob}_p)) = a_p, \qquad \det(\rho_\lambda(\operatorname{Frob}_p)) = \varepsilon(p)p^{k-1}$$

for prime numbers p that are prime both to N and to the norm of λ. The symbol Frob_p denotes an arithmetic Frobenius element for p in $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. (For a survey of this material, the reader may wish to consult **[37]**.)

The reduction of ρ_λ mod λ is a certain semisimple representation

$$\bar{\rho}_\lambda\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F}_\lambda),$$

where $\mathbf{F}_\lambda$ is the residue field of λ. Its construction involves a choice: one chooses a model of ρ_λ that has values in $\mathbf{GL}(2, \mathscr{O}_\lambda)$, where $\mathscr{O}_\lambda$ is the ring of integers of E_λ, and then forms the "naive reduction" of this model, i.e., the composite of ρ_λ and the canonical map $\mathbf{GL}(2, \mathscr{O}_\lambda) \to \mathbf{GL}(2, \mathbf{F}_\lambda)$. The *semisimplification* of this naive reduction depends only on ρ_λ. It is the desired representation $\bar{\rho}_\lambda$.

Consider the representations $\bar{\rho}_\lambda$ obtained for varying N, k, f, but with $\mathbf{F}_\lambda$ having a fixed characteristic. To assemble them, fix a prime number ℓ, and choose a place v dividing ℓ of the field of algebraic numbers in $\mathbf{C}$. Let $\mathbf{F}$ be the residue field of v, so that $\mathbf{F}$ is an algebraic closure of its prime field $\mathbf{F}_\ell$. For each algebraic number field E contained in $\mathbf{C}$, v induces a prime λ on E, together with a canonical inclusion $\mathbf{F}_\lambda \hookrightarrow \mathbf{F}$. In particular, if E is the field generated by the eigenvalues attached to an eigenform f and λ is the prime of E induced by v, the representation $\bar{\rho}_\lambda$ may be viewed as taking values in $\mathbf{GL}(2, \mathbf{F})$.

Suppose that $\rho\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ is a continuous semisimple representation. When can we expect that ρ is isomorphic to one of the $\bar{\rho}_\lambda$? Identify ε with a character defined on $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, and let $\tilde{\chi}$ be the ℓ-adic cyclotomic character. Recall that the identity $\det(\rho_\lambda(\operatorname{Frob}_p)) = \varepsilon(p)p^{k-1}$ on Frobenius elements implies the formula

$$\det \rho_\lambda = \varepsilon \tilde{\chi}^{k-1} \tag{1.1}$$

for $\det \rho_\lambda$. On reducing this identity, one obtains $\det \bar{\rho}_\lambda = \bar{\varepsilon}\chi^{k-1}$, where χ is the mod ℓ cyclotomic character and $\bar{\varepsilon}$ is the reduction of ε mod λ (i.e., mod v). In particular, let $c \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ be a complex conjugation. Then $\det \bar{\rho}_\lambda(c) = \bar{\varepsilon}(c)\chi(c)^{k-1}$. Now $\chi(c) = -1$, and furthermore $\varepsilon(c)$ is another name for $\varepsilon(-1)$. Using the formula $\varepsilon(-1) = (-1)^k$, we obtain $\det \bar{\rho}_\lambda(c) = -1$. In other words, $\bar{\rho}$ must be an *odd* representation if it is to be isomorphic to some $\bar{\rho}_\lambda$.

Serre **[49]** has proposed that this parity condition represents a sufficient, as well as a necessary, condition for a representation $\rho\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$

to arise from some eigenform. The *reducible* odd representations ρ arise from Eisenstein series and will be neglected in the following discussion. For *irreducible* representations, we have the

(1.2) SERRE CONJECTURE **[49,** $(3.2.3_?)$**]**. *Let* $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ *be an irreducible, odd representation. Then* ρ *is isomorphic to the representation* $\bar{\rho}_\lambda$ *associated to some eigenform in one of the spaces* $S_k(\Gamma_1(N))$.

The associated **Refined Conjecture [49,** $(3.2.4_?)$**]** asserts that ρ is a representation $\bar{\rho}_\lambda$ associated to an eigenform in a *specific* space $S_{k(\rho)}(\Gamma_1(N(\rho)))$. The invariants $k(\rho)$ and $N(\rho)$ are defined by local properties of ρ. More precisely, as we recall in §2, $N(\rho)$ is an integer prime to ℓ that depends only on the restrictions of ρ to decomposition groups in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for each prime number $p \neq \ell$. Similarly, $k(\rho)$ is an integer ≥ 2 that depends only on the restriction of ρ to a decomposition group for ℓ in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. (See **[49, §2]**.)

The formulation of **[49,** $(3.2.4_?)$**]** requires that ρ arise from an eigenform $f \in S_{k(\rho)}(\Gamma_1(N(\rho)))$ whose character ε is also predicted in advance. This extra requirement means simply that ε can be chosen to be of order prime to ℓ. Subsequently, however, Serre found examples for $\ell \leq 3$ showing that the requirement cannot always be satisfied **[50]**; his examples concern the two-dimensional space $S_2(\Gamma_1(13))$. On the other hand, Carayol [7, Proposition 3] and Serre have each shown that the situation for $\ell \geq 5$ is much more favorable: the representations $\bar{\rho}_\lambda$ arising from a given space $S_k(\Gamma_1(N))$ all arise from eigenforms whose associated character has prime-to-ℓ order.

More precisely, Carayol proves the following result [7, Proposition 3]:

(1.3) THEOREM. *Assume that* $\ell \geq 5$ *and that* $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ *is an odd irreducible representation. Suppose that* ρ *arises from an eigenform* $f \in S_k(\Gamma_1(N))$ *and that* ε *is the Dirichlet character that is associated to* f. *Let* ε' *be a character on* $(\mathbf{Z}/N\mathbf{Z})^*$ *that is congruent to* $\varepsilon \bmod v$. *Then* ρ *arises from an eigenform* $f' \in S_k(\Gamma_1(N))$ *whose Nebentypus character is* ε'.

In view of this result and the counterexamples for $\ell = 2$ and $\ell = 3$, we will restrict attention to the invariants $N(\rho)$ and $k(\rho)$ in discussing Serre's conjectures.

The Refined Conjecture has a number of striking consequences, which are catalogued in §4 of **[49]**. Among these are Fermat's Last "Theorem," the conjecture of Taniyama that all elliptic curves over $\mathbf{Q}$ are modular, and variants of Taniyama's conjecture concerning other motives of rank 2.

To bridge the gap between the Serre Conjecture and its refinement, we formulate the following

(1.4) MOTIVATING PROBLEM. *Suppose that* $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ *is an irreducible representation that arises from some space of cusp forms* $S_k(\Gamma_1(N))$. *Does* ρ *arise from an eigenform in* $S_{k(\rho)}(\Gamma_1(N(\rho)))$, *where* $k(\rho)$ *and* $N(\rho)$ *are the invariants assigned to* ρ *by Serre?*

An affirmative solution to (1.4) would imply an equivalence between Serre's conjecture and its refinement. At the time of this writing, one is very close to a complete solution of (1.4), at least when ℓ is an odd prime. The solution involves the work of a large number of mathematicians, including N. Boston, H. Carayol, F. Diamond, B. Edixhoven, G. Faltings, B. H. Gross, B. Jordan, H. W. Lenstra, Jr., R. Livné, B. Mazur, J.-P. Serre, and the author.

The goal of this article is twofold. On the one hand, we shall describe what is known about (1.4). On the other, we shall present the following new result:

(1.5) THEOREM. *Let $\ell \geq 3$ be a prime, and let ρ: $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ be irreducible. Suppose that ρ arises from an eigenform f of weight two and trivial character on $\Gamma_1(M) \cap \Gamma_0(p)$, where p is prime to ℓM. Assume that ρ is unramified at p. Then ρ arises from a weight-two eigenform with trivial character on $\Gamma_1(M)$.*

The new feature of this theorem is that the level M is not required to be prime to ℓ. In fact, M can be divisible by an arbitrarily high power of ℓ. As it stands, (1.5) applies only to forms with trivial character on $\Gamma_1(M) \cap \Gamma_0(p)$, i.e., to forms on $\Gamma_0(Mp)$. However, it is very likely that the proof we give for (1.5) will generalize without difficulty to cover eigenforms on $\Gamma_1(M) \cap \Gamma_0(p)$ whose Nebentypus characters are *arbitrary* characters mod M. The discussion below shows that (1.4) will be solved for primes $\ell \geq 5$ as soon as Theorem 1.5 is generalized to this case.

One can also contemplate a variant of (1.5) that applies to forms of weight k on $\Gamma_1(M) \cap \Gamma_0(p)$, where k satisfies $2 \leq k \leq \ell+1$ and M is required to be prime to ℓ. Such a variant, which accepts arbitrary characters mod M, will again lead to a solution of (1.4). It seems very likely that the methods of Faltings, Jordan, and Livné **[16, 24]** will lead to a variant of this type.

In light of the substantial progress made in relating Serre's conjecture to its refinement, it is striking that there is little to report concerning the conjecture itself **[49**, 3.2.3$_?$**]**, aside from the numerical evidence presented in §5 of **[49]**.

2. Stripping powers of ℓ from the level

Let ρ: $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ be an irreducible odd representation, as above, and let $N(\rho)$ and $k(\rho)$ be the invariants assigned to ρ by Serre. Recall **[49**, §1**]** that $N(\rho)$ is a product $\prod p^{n(p,\rho)}$ extended over the set of prime numbers $p \neq \ell$. The integer $n(p, \rho)$ is defined by restricting ρ to a decomposition group D_p for p in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Consider the ramification subgroup G_0 of D_p, along with the higher ramification groups G_i ($i > 0$). Let V be a two-dimensional $\mathbf{F}$-vector space affording the representation ρ. For each $i \geq 0$, let V_i be the subspace of V consisting of those $v \in V$ that are fixed by all elements of G_i. Then

$$n(p, \rho) = \sum_{i \geq 0} \frac{1}{(G_0 : G_i)} \dim V/V_i.$$

In particular, $N(\rho)$ is prime to ℓ, so Serre's refined conjecture implies the weaker statement that ρ arises from some space of cusp forms $S_k(\Gamma_1(N))$, with N prime to ℓ. Let us say simply that ρ arises from a subgroup Γ of $\mathbf{SL}(2, \mathbf{Z})$ if it arises from the space of weight-k cusp forms on Γ, for *some* $k \geq 2$. Then the weak statement implied by Serre's conjecture is that ρ arises from the group $\Gamma_1(N)$ for some N that is prime to ℓ. We shall now show (at least when $\ell \geq 3$) that any ρ that is modular in the sense that it arises from some $\Gamma_1(N)$ with N not necessarily prime to ℓ is automatically modular in the apparently stronger sense that it arises from $\Gamma_1(N)$ for some N that is prime to ℓ. This is certainly a "well-known" fact, which is close to results that are already in the literature, cf. [**49,** *Remarque*, p. 195].

(2.1) THEOREM. *Assume $\ell \geq 3$. Suppose that ρ arises from $\Gamma_1(M)$, where M is the product $N\ell^\alpha$ with $(N, \ell) = 1$. Then ρ arises from $\Gamma_1(N)$.*

We prove (2.1) using the concrete techniques of Serre [**46,** §3], [**47,** Theorem 5.4], and Queen [**35,** §3]. An alternative method would be based on results of N. Katz: see the appendices of [**25**] and the discussions in [**20,** §1] and [**21,** §1]. (Thanks are due to Fred Diamond for pointing out these references.) Note that the assumption $\ell \geq 3$ is made principally for convenience; it should be possible to analyze the case $\ell = 2$ without great difficulty.

Our proof consists of several independent steps. Before beginning, we shall normalize the eigenforms that occur in the proof. Recall that if f satisfies $f|T_n = a_n f$ for all $n \geq 1$, then we may scale f by a constant so that its Fourier coefficients are the eigenvalues a_n. We will assume that all eigenforms are scaled in this way, i.e., that they are "normalized eigenforms." We represent these eigenforms as series $\sum_{n \geq 1} a_n q^n$; the variable q initially represents $e^{2\pi i\tau}$, where τ is an element of the complex upper half-plane, but rapidly becomes a formal variable. In particular, we say that two forms are congruent mod ℓ or mod v if their Fourier series are formally congruent. Similarly, the mod v reduction of an eigenform f is meant to be the formal power series $\sum \bar{a}_n q^n$, where "‾" denotes the reduction map mod v.

STEP 1. *The representation ρ arises from $\Gamma_0(\ell^r) \cap \Gamma_1(\ell N)$ for some $r \geq 0$.*

When $\ell \geq 5$, the assertion follows from (1.3). In that case, we can take $r = \alpha$. It can be proved in an elementary manner, for all $\ell \geq 3$, by the following argument.

Suppose that ρ arises from an eigenform f in $S_k(\Gamma_1(M))$ and that κ is the associated Dirichlet character mod M. We assume that $k \geq 2$ and that α is positive (i.e., that M is divisible by ℓ). Decompose κ as a product

$\varepsilon\eta\omega^i$, where ε has conductor dividing N, η has ℓ-power order and ℓ-power conductor, and ω is the "Teichmüller" character, i.e., that character of conductor ℓ and order $\ell-1$ that is congruent to the identity function mod v. The exponent i is naturally an integer mod $\ell-1$. Because η has, in particular, odd order, we may write it in the form ξ^{-2}, where ξ is a character of ℓ-power order. The cusp form $f\otimes\xi$ is a form of type $(k, \varepsilon\omega^i)$ (i.e., weight k and character $\varepsilon\omega^i$) which gives rise to ρ. If $r \geq \alpha$ is large enough so that ℓ^r is divisible by the square of the conductor of ξ, then $f\otimes\xi$ is again of level $N\ell^r$.

In what follows, we fix an $r > 0$ so that ρ arises from $\Gamma_0(\ell^r)\cap\Gamma_1(\ell N)$.

STEP 2. *ρ arises from* $\Gamma_0(\ell^r)\cap\Gamma_1(N)$.

We assume as above that ρ arises from an eigenform of type $(k, \varepsilon\omega^i)$ on the group $\Gamma_0(\ell^r)\cap\Gamma_1(\ell N)$. Take the exponent i to be a positive integer. The Eisenstein series

$$G := L(1-i, \omega^{-i})/2 + \sum_{n=1}^{\infty}\left(\sum_{d|n}\omega^{-i}(d)d^{i-1}\right)q^n$$

is then of type (i, ω^{-i}) on $\Gamma_0(\ell)$, as proved (for example) in [**46,** Lemme 10]. Thus fG is on the group $\Gamma_0(\ell^r)\cap\Gamma_1(N)$, if r is positive. On the other hand, it is well known that the order at v of $L(1-i, \omega^{-i})$ is negative; this follows from the Kummer congruences and the Von Staudt theorem. (See, e.g., [**22,** §3.4].) In other words, if c is the constant term of G, then $E := c^{-1}G$ is a form with v-integral coefficients such that $E \equiv 1 \mod v$. Consequently, the product fE, viewed mod v, is a nonzero eigenform with the same eigenvalues as f. At the same time, fE is on the group $\Gamma_0(\ell^r)\cap\Gamma_1(N)$. By a well-known lemma [**12,** Lemme 6.11], we may find an eigenform on $\Gamma_0(\ell^r)\cap\Gamma_1(N)$ whose eigenvalues are congruent to those of f.

STEP 3. *ρ arises from* $\Gamma_0(\ell)\cap\Gamma_1(N)$.

Assume that ρ arises from an eigenform $f = \sum a_n q^n$ on $\Gamma_0(\ell^r)\cap\Gamma_1(N)$, with $r > 1$. Let K be a finite Galois extension of $\mathbf{Q}$ containing the a_n, and let $\sigma \in \mathrm{Gal}(K/\mathbf{Q})$ be a Frobenius element for v. Thus $\sigma a \equiv a^\ell \mod v$ for all elements a of the ring of integers of K. The series $\sum \sigma^{-1}a_n q^n$ is the Fourier expansion of a normalized eigenform $\sigma^{-1}f$ of the same weight as f. We wish to show that f is congruent mod v to a cusp form of some weight on $\Gamma_0(\ell^{r-1})\cap\Gamma_1(N)$. Consider $g := (\sigma^{-1}f)^\ell|U$, where U is the ℓ^{th} Hecke operator T_ℓ. On the one hand, g is a form on $\Gamma_0(\ell^{r-1})\cap\Gamma_1(N)$ (see, e.g., [**29,** Lemma 1]). On the other hand, the Fourier expansion of g is congruent mod v to $\sum(\sigma^{-1}a_n)^\ell q^n$, which in turn is congruent to the Fourier expansion of f.

STEP 4. *ρ arises from $\Gamma_1(N)$.*

To prove this, we use the argument given by Serre in [**46**, §3.2], replacing the operators W, U, V, ... by their analogues for forms on $\Gamma_1(N)\cap\Gamma_0(\ell)$. More precisely, we let U be the ℓ^{th} Hecke operator, as above, and let V be the operator $\sum a_n q^n \mapsto \sum a_n q^{\ell n}$, which takes forms on $\Gamma_1(N)$ to forms of the same weight on $\Gamma_1(N)\cap\Gamma_0(\ell)$. Further, we let W be the operator $V_\ell^{N\ell}$ as defined in §1 of [**29**]. Thus W is given by the matrix $\left(\begin{smallmatrix} \ell x & y \\ N\ell z & \ell \end{smallmatrix}\right)$, where x, y, and z are integers, and we have $\ell x - Nyz = 1$. This operator is the inverse of the operator denoted W_Q in [**1**]. Indeed, by [**1**, Proposition 1.1], one has $W = \varepsilon(\ell)W_Q$ on the space of forms with character ε on $\Gamma_1(N)\cap\Gamma_0(\ell)$. (We consider ε as a Dirichlet character mod N, which permits us to evaluate it at ℓ.) At the same time, W^2 is multiplication by $\varepsilon(\ell)$, according to Lemma 2 of [**29**].

Next, suppose that F is a form of weight w and character ε on $\Gamma_1(N)\cap\Gamma_0(\ell)$. Then [**29**, Lemma 3] implies that the form

$$\mathrm{Tr}(F) := F + \varepsilon^{-1}(\ell)\ell^{1-w/2}F|W|U$$

is a form of weight w on $\Gamma_1(N)$. (Apply [**29**, Lemma 3] to $F|W$.) Further, if G is a form of weight w and character ε on $\Gamma_1(N)$, one has $G|W = \ell^{w/2}\varepsilon(\ell)G|V$ [**1**, Proposition 1.5].

If $\ell > 3$, let E be the normalized Eisenstein series of weight $\ell-1$ on $\mathbf{SL}(2,\mathbf{Z})$, so that we have $E \equiv 1 \mod \ell$. Similarly, if $\ell = 3$, let E be the normalized Eisenstein series E_4 of weight four. Let a denote the weight of E. As in [**46**], we introduce the form

$$g := E - \ell^{a/2}E|W = E - \ell^a E|V.$$

One has $g \equiv 1 \mod \ell$; furthermore, $g|W$ is divisible by a positive power of ℓ (in fact, by $\ell^{1+a/2}$).

Let f be an eigenform giving rise to ρ, as above, and consider $\mathrm{Tr}(fg^i)$, where i is a positive integer. The modular form $\mathrm{Tr}(fg^i)$ is a form on $\Gamma_1(N)$. On the other hand, a calculation similar to that of §3.2 of [**46**] shows that $\mathrm{Tr}(fg^i) \equiv f \mod \ell$ for sufficiently large i. □

A partial converse to (2.1) is the following result.

(2.2) THEOREM. *Suppose that ρ arises from $S_k(\Gamma_1(N))$, with N prime to ℓ and $2 \le k \le \ell+1$. Assume that $\ell > 3$ or that $N > 3$. Then ρ arises from $S_2(\Gamma_1(N\ell))$.*

For $\ell \ge 5$, the theorem is proved by Ash and Stevens [**2**, Theorem 3.5a] by methods involving parabolic cohomology. Another approach, which originated with Serre, is worked out in Gross's article [**17**, Proposition 9.3]. Although Gross's prime p (which corresponds to our ℓ) is arbitrary, he assumes that the level N satisfies $N \ge 4$. According to comments in §10 of **[17]**, Proposition 9.3 of **[17]** extends to the case where $N \leq 2$ and k is even.

3. Adjustment of the weight

Let $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ be an irreducible representation, and let $k(\rho)$ be the "weight" attached to ρ by Serre **[49,** §2]. Edixhoven **[15,** §4] proved that if ρ arises from a cusp form of some weight, then ρ arises from a cusp form of weight $k(\rho)$. More precisely, **[15,** Theorem 4.3] gives the following result.

(3.1) THEOREM. *Let N be prime to ℓ, and assume that ρ arises from an eigenform $f \in S_k(\Gamma_1(N))$ with character ε. Then ρ arises from an eigenform $f' \in S_{k(\rho)}(\Gamma_1(N))$ with character ε. The integers k and $k(\rho)$ are congruent* mod $\ell-1$; *moreover, we have $k \geq k(\rho)$ provided that ℓ is odd.*

The proof of (3.1) relies on two points which merit independent discussion. First, let D_ℓ be a decomposition group for ℓ in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and let I_ℓ be the inertia subgroup of D_ℓ. Suppose that

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}\, V \approx \mathbf{GL}(2, \mathbf{F})$$

arises from an eigenform $f = \sum a_n q^n$ of weight k and character ε on $\Gamma_1(N)$, with N prime to ℓ. The proof of (3.1) requires information about the restriction of ρ to D_ℓ, in the case where k satisfies $2 \leq k \leq \ell$.

Specifically, suppose that this inequality is satisfied and that f is ordinary in the sense that $a_\ell \not\equiv 0 \mod v$. A theorem of Deligne (proved in **[17]**) states that V has an unramified quotient V/V_0 on which Frob_ℓ acts by multiplication by a_ℓ (or, more precisely, by the image of a_ℓ in $\mathbf{F}$). The action of D_ℓ on V_0 is determined by this information, since the determinant of ρ may be computed from k and ε. On the other hand, suppose that f is supersingular, i.e., not ordinary. Let ψ and ψ' be the two fundamental characters $I_\ell \to \mathbf{F}$ of level 2, as defined in **[44,** §1.7]. A theorem of Fontaine states that the restriction of ρ to I_ℓ is the direct sum of two one-dimensional representations, on which I_ℓ operates by the $(k-1)^{\text{st}}$ powers of ψ and ψ'. This theorem was proved by Fontaine in an exchange of letters with Serre in 1979; a new proof was given by Edixhoven in **[15]**.

Secondly, Edixhoven's proof requires B. H. Gross's result about companion forms **[17]**, which conversely may be viewed as a special case of (3.1). Before stating Gross's result, we recall that there is an operator $\theta = q\frac{d}{dq}$ on mod ℓ modular forms for $\Gamma_1(N)$ when N is prime to ℓ. This operator maps forms of weight k to forms of weight $k+\ell+1$. It was considered initially by Serre and Swinnerton-Dyer **[45, 53]** for forms on $\mathbf{SL}(2, \mathbf{Z})$ and then constructed by Katz **[26]** for forms of full level N, where $N \geq 3$. It is introduced in **[17,** §4] in the case of forms on $\Gamma_1(N)$, with $N \geq 4$. According to Edixhoven **[15,** §2.1], mod ℓ modular forms of weight k on $\Gamma_1(N)$ are $\mathbf{GL}(2, \mathbf{Z}/n\mathbf{Z})$-invariant sections of the line bundle $\underline{\omega}^k$ on the curve X. Here $n \geq 3$ is prime to ℓ, and X is the modular curve parametrizing elliptic curves with a $(\Gamma_1(N), \Gamma(n))$-structure. The integer N is assumed only to be prime to ℓ. Using this fact, Edixhoven constructs θ for modular forms of weight k on $\Gamma_1(N)$.

Consider now an ordinary form f as above, whose weight k satisfies $2 \le k \le \ell-1$. Suppose in addition that the exact sequence of $\mathbf{F}[D_\ell]$-modules

$$0 \to V_0 \to V \to V/V_0 \to 0$$

is *split*. The inertia group I_ℓ acts on V_0 via the character χ^{k-1}, since $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts on $\det V$ via the product $\varepsilon\chi^{k-1}$ and since I_ℓ acts trivially on V/V_0. Consider the representation $\rho' := \rho \otimes \chi^{1-k}$ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Locally at ℓ, it is the direct sum of $V_0 \otimes \chi^{1-k}$ and $(V/V_0) \otimes \chi^{1-k}$. In particular, it is an extension of the unramified line $V_0 \otimes \chi^{1-k}$ by a line on which I_ℓ acts via the character $\chi^{1-k} = \chi^{\ell+1-k-1}$. This suggests that ρ' arises from an eigenform of weight $k' := \ell+1-k$ and character ε on $\Gamma_1(N)$; indeed, Serre's recipe for weights sets $k(\rho') = k'$. It is clear that ρ' is modular of some weight, because twisting representations by χ corresponds to applying the operator θ on modular forms. Hence (3.1) asserts, in particular, that ρ' arises from an eigenform f' of weight k'. The construction of the "companion" f' to f is the main result of [**17**]. (If $k = \ell$ and ρ is split locally at ℓ, then Gross still produces a companion form, of weight 1, in certain cases. For ℓ odd, the remaining cases of weight 1 are treated in a recent article by Coleman and Voloch [**9**].) This completes the discussion of Theorem 3.1.

As above, we shall use the phrase "ρ arises from $\Gamma_1(N)$" to indicate that ρ arises from an eigenform in $S_k(\Gamma_1(N))$ for some $k \ge 2$.

(**3.2**) COROLLARY. *Suppose that ρ arises from $\Gamma_1(N)$, with N prime to ℓ. Then there exists a power χ^i of the* mod ℓ *cyclotomic character χ such that $\rho \otimes \chi^i$ arises from $S_k(\Gamma_1(N))$ for some $k \le \ell+1$.*

PROOF. Replacement of ρ by any of its twists $\rho \otimes \chi^i$ ($i = 0, 1, 2, \dots$) corresponds to applying the operator $\theta = q\frac{d}{dq}$ on modular forms i times. Hence, ρ arises from a given $\Gamma_1(N)$ if and only if $\rho_i := \rho \otimes \chi^i$ arises from the same group. The definition of $k(\rho)$ is such that a suitable twist ρ_i satisfies the inequality

$$2 \le k(\rho_i) \le \ell+1.$$

(See [**15**, Theorem 3.4], and the discussion following the statement of that theorem.) By Theorem 3.1, ρ_i arises from $S_{k(\rho_i)}(\Gamma_1(N))$. □

Let ρ: $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ be an irreducible representation. Consider the following four sets of positive integers:

$\mathcal{N}_1 := \{ N \text{ prime to } \ell \mid \rho \text{ arises from } S_{k(\rho)}(\Gamma_1(N)) \}$,

$\mathcal{N}_2 := \{ N \text{ prime to } \ell \mid \rho \text{ arises from } \Gamma_1(N) \}$,

$\mathcal{N}_3 := \{ N \text{ prime to } \ell \mid \rho \text{ arises from } \Gamma_1(N\ell^\alpha) \text{ for some } \alpha \geq 0 \}$,

$\mathcal{N}_4 := \{ N \text{ prime to } \ell \mid \rho \text{ arises from } S_2(\Gamma_1(N\ell^2)) \}$.

(3.3) THEOREM. *If $\ell \geq 5$, then the four sets of integers $\mathcal{N}_i(\rho)$ are equal.*

PROOF. The equality of $\mathcal{N}_1$ and $\mathcal{N}_2$ is guaranteed by Theorem 3.1. That $\mathcal{N}_2$ and $\mathcal{N}_3$ are equal follows from Theorem 2.1. Since it is clear that $\mathcal{N}_4 \subseteq \mathcal{N}_3$, it remains only to show that $\mathcal{N}_3 \subseteq \mathcal{N}_4$.

We will show, equivalently, that $\mathcal{N}_2 \subseteq \mathcal{N}_4$. Assume that ρ arises from $\Gamma_1(N)$, and choose $i \geq 0$ such that $\rho \otimes \chi^i$ arises from $S_k(\Gamma_1(N))$, with $2 \leq k \leq \ell+1$. By (2.2), $\rho \otimes \chi^i$ arises from $S_2(\Gamma_1(N\ell))$. To see that $N \in \mathcal{N}_4$, we view the twisting operator $\otimes\chi^{-i}$ as a Dirichlet twist on modular forms. Such a twist changes the level of a modular form but not its weight. From [**1**, Proposition 3.1], we find that $\rho = (\rho\otimes\chi^i) \otimes \chi^{-i}$ arises from $S_2(\Gamma_1(N\ell^2))$. □

In case $\ell \geq 5$, we will write simply $\mathcal{N}(\rho)$ for the common value of the four sets $\mathcal{N}_i(\rho)$.

4. The levels of ρ and of f

Suppose that an irreducible representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ is "modular" in the sense that it arises from an eigenform of some weight and some level. Problem (1.4) requires that we find an eigenform of weight $k(\rho)$ and level $N(\rho)$ that gives rise to ρ. Because of Edixhoven's theorem (Theorem 3.1), the weight aspect of (1.4) disappears. In fact, suppose that $\ell \geq 5$. Then Theorem 3.3 translates (1.4) into the following question:

(4.1) $$N(\rho) \overset{?}{\in} \mathcal{N}(\rho).$$

In other words, our goal is to show that ρ is modular of level $N\ell^\alpha$, for some $\alpha \geq 0$. With (4.1) in mind, we are led to examine the set $\mathcal{N}(\rho)$ more closely. The following theorem was proved by Carayol [**7**] and, independently, by Livné [**30**, Proposition 0.1]:

(4.2) THEOREM. *Suppose that ρ arises from $\Gamma_1(N)$. Then Serre's invariant $N(\rho)$ divides N.*

In attacking problem (4.1), one begins with an eigenform f giving rise to ρ. If N is the level of f, we have $N(\rho)|N$. Since $N(\rho)$ is, by definition, prime to ℓ, we have $N(\rho)|N'$, where N' is the prime-to-ℓ part of N. In case this divisibility is strict, the aim is to "lower" N by replacing it by a divisor. For the most part, the strategy is to do this "locally": we consider a prime $p \neq \ell$ that divides $N/N(\rho)$ and seek to replace N by N/p^i for some positive integer i.

As a preliminary step, we should certainly replace f by the *newform* (or primitive form) associated to f. This is an eigenform whose system of λ-adic representations coincides with the system (ρ_λ) attached to f and whose level is minimal among all such eigenforms. The level of the newform attached to f is a divisor of N. (For the theory of newforms, see for instance the account in **[29]** or **[34,** §4.6].) According to a theorem of Deligne, Langlands **[28]**, and Carayol **[6]**, the conductor of the system (ρ_λ) coincides with N once this replacement is made.

Assume, then, that f is a newform. Theorem 4.2 is actually a consequence of a much more refined comparison of $N(\rho)$ with (the prime-to-ℓ part of) N. This comparison is again due to Carayol [7] and Livné **[30]**. We will discuss the analysis of Carayol and Livné, under the simplifying assumption $\ell \geq 5$. In view of (1.3), this assumption permits us to assume that the character ε that is associated to f has prime-to-ℓ order and thereby facilitates the discussion.

Consider that representation ρ_λ that is attached to f and the prime λ induced by the chosen place v of $\overline{\mathbf{Q}}$. Then ρ is the reduction of ρ_λ in the sense that is outlined above. Since ρ is irreducible, ρ is in fact the "naive" reduction of any model of ρ_λ over $\mathcal{O}_\lambda$: no semisimplification is required.

Fix a prime $p \neq \ell$, and let $D_p \subset \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ again be a decomposition group for p. Let I_p be the inertia subgroup of D_p, and let $e = \mathrm{ord}_p(N/N(\rho))$. Thus, e is a nonnegative integer, according to (4.2).

The inequality $e \geq 0$ can be interpreted conceptually. Consider ρ_λ and ρ as homomorphisms

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(V_\lambda), \qquad \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}\, V,$$

where V_λ and V are two-dimensional vector spaces over E_λ and $\mathbf{F}$ respectively. Then one has

$$e = \dim_{\mathbf{F}} V^{I_p} - \dim_{E_\lambda} V_\lambda^{I_p}, \tag{4.3}$$

where V^{I_p} is the space of I_p-invariants in V and $V_\lambda^{I_p}$ is space of I_p-invariants in V_λ. In particular, we have $e > 0$ if and only if

$$\dim_{\mathbf{F}} V^{I_p} > \dim_{E_\lambda} V_\lambda^{I_p}.$$

Carayol and Livné have classified the situations in which we have $e > 0$. We will now describe their classification. For each situation that arises in the classification, we will summarize what is known about the question of showing that ρ is modular of level (dividing) N/p, i.e., that ρ arises from $\Gamma_1(N/p)$. Our summary appeals to a number of results which are proved only later in this article (for example, Theorems 5.1, 1.5, and 8.1). The material in this section may thus be viewed as a motivation for the theorems in §§5–8.

The first step in the classification is to distinguish cases according to the behavior of the component at p of the automorphic representation attached to f. Recall that the newform f is associated to an automorphic representation $\pi_{\mathbf{A}}$ of $\mathbf{GL}(2, \mathbf{A})$, where $\mathbf{A}$ is the adèle ring of $\mathbf{Q}$. Let π_p be the component of $\pi_{\mathbf{A}}$ at p, so that π is an admissible representation of $\mathbf{GL}(2, \mathbf{Q}_p)$. This representation may be classified as a principal series representation, a special representation, or a cuspidal representation of $\mathbf{GL}(2, \mathbf{Q}_p)$.

The Langlands correspondence attaches to π a λ-adic representation $\rho_{\lambda,\pi}$ of the Weil group of $\mathbf{Q}_p$. (Strictly speaking, $\rho_{\lambda,\pi}$ is defined only up to isomorphism. See **[54,** §4] for a discussion of the relation between λ-adic representations of the Weil group of $\mathbf{Q}_p$ and complex representations of the corresponding Weil-Deligne group.) By the main result of **[6]**, $\rho_{\lambda,\pi}$ is isomorphic to the restriction of ρ_λ to the Weil group of $\mathbf{Q}_p$. Therefore, properties of π are mirrored by the local behavior of ρ_λ at p. The restriction of ρ_λ to D_p is: reducible and semisimple when π is a principal series representation, reducible but not semisimple when π is special, and irreducible when π is cuspidal.

We now discuss the inequality $e > 0$ in each of these three cases.

The principal series case. In this case, which was treated by Carayol, the representation π arises from an (unordered) pair of Grossencharacters of $\mathbf{Q}_p^*$, say α and β. Identify these characters with characters of D_p, using the reciprocity map of local class field theory. Then $\rho_\lambda|_{D_p}$ is the direct sum $\alpha \oplus \beta$.

(4.4) PROPOSITION [7]. *Assume that $e > 0$. Then there is a Dirichlet character ϕ of conductor p and ℓ-power order such that the newform associated to $f \otimes \phi$ has level dividing N/p. In particular, ρ is modular of level N/p.*

PROOF. One of α and β is ramified but has unramified reduction mod λ. After possibly permuting α and β, we may assume that the restriction of α to I_p is nontrivial but has ℓ-power order. On the other hand, because of formula (1.1), and the assumption that the order of ε is prime to ℓ, the restriction of $\det \rho_\lambda$ to I_p has prime-to-ℓ order. (Note that $\tilde{\chi}$ is ramified only at ℓ.) Thus $\alpha\beta|_{I_p}$ has prime-to-ℓ order. It follows from this that, on I_p, α is a power of β, since β is the product $\alpha^{-1} \cdot \alpha\beta$ of characters of relatively prime orders. Thus the kernel of $\alpha^{-1}|_{I_p}$ contains that of $\beta|_{I_p}$, and it follows that the conductor of $\alpha^{-1}\beta$ is at most that of β. (Incidentally, we will write the conductor multiplicatively: the conductor of a character of D_p is a power of p, and the conductor of the trivial character of D_p is $1 = p^0$.)

Since $\alpha|_{I_p}$ has prime-to-p order, α coincides on I_p with some power ω^i of the Teichmüller character ω which appeared in the proof of (2.1). Note that ω is a character that is defined naturally on $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Let $\phi = \omega^{-i}$, and consider the twist $f \otimes \phi$, which corresponds to the λ-adic representation $\rho_\lambda \otimes \phi$. Since ϕ has ℓ-power order, the reductions of $\rho_\lambda \otimes \phi$

and ρ_λ coincide. On the other hand, $(\rho_\lambda \otimes \phi)|_{D_p}$ is a direct sum of the *unramified* $\alpha\phi$ and the character $\beta\phi$. This latter character coincides with $\beta\alpha^{-1}$ on I_p. Therefore, its conductor is a divisor of the conductor of β. At the same time, the conductor of $\alpha\phi$, which is 1, is a strict divisor of the conductor of α. Therefore, the power of p dividing the conductor of $\rho_\lambda\phi$ is smaller than the power of p dividing the conductor of ρ_λ.

In the language of cusp forms, this means that the level of the newform associated to $f \otimes \phi$ is divisible by a smaller power of p than the level N of f. The two levels differ only at p, since the character ϕ is ramified only at p. □

The special case. Here, we have $\pi = \alpha \otimes \mathrm{sp}$, where sp is a "standard" special representation and α is a Grossencharacter of $\mathbf{Q}_p^*$, which we identify as above with a character of D_p. We consider that sp has been normalized so that $\rho_\lambda|_{D_p}$ is an extension of the one-dimensional representation with character α by the one-dimensional representation with character $\alpha\tilde{\chi}$. (The fact that $\rho_\lambda|_{D_p}$ has this form, when π is special, was proved by Langlands [**28**].) The semisimplification of $\rho|_{D_p}$ then has the form $\bar{\alpha} \oplus \bar{\alpha}\chi$, where $\bar{\alpha}$ is the reduction of α mod λ. Also, $\varepsilon = \alpha^2\tilde{\chi}^{2-k}$. Finally, the conductor of $\alpha \otimes \mathrm{sp}$ is the square of the conductor of α if α is ramified, and p^1 if not.

(4.5) Proposition [7]. *Assume that $e > 0$ and that α is ramified. Then there is a Dirichlet character ϕ of conductor p and ℓ-power order such that the newform associated to $f \otimes \phi$ has level dividing N/p. In particular, ρ is modular of level N/p.*

Proof. Since $e > 0$, we have $\dim_{\mathbf{F}} V^{I_p} > 0$, so that $\bar{\alpha}$ is unramified. Thus $\alpha|_{I_p}$ has ℓ-power order and coincides with some power ω^i of ω. It follows that the conductor of α is p. As above, we set $\phi = \omega^{-i}$. The representation $\pi \otimes \phi$ then has conductor p, whereas π itself has conductor p^2. Hence replacing f by the newform associated to $f \otimes \phi$ replaces N by N/p. □

In the situation of Proposition 4.5, $\bar{\varepsilon}$ is unramified at p, since the determinant of ρ is unramified at p. Since we have assumed that the order of ε is prime to ℓ, this implies that ε is unramified at p. Hence f is in fact a newform on the group $\Gamma_1(M) \cap \Gamma_0(p^2)$, where $M = N/p^2$ is prime to p, and the character ϕ is quadratic.

In the case where π is special but Proposition 4.5 does *not* apply, α is an unramified character, so that ε, in particular, is unramified. The newform f then has level Mp, with M prime to p, and f is a cusp form on the group $\Gamma_1(M) \cap \Gamma_0(p)$. We have $e > 0$ if and only if the representation ρ is unramified. When this is the case, $N(\rho)$ is prime to p, and one must show that ρ is represented by an eigenform of level M. Here is a statement of the problem to be solved (cf. [**7**, Conjecture B]):

(4.6) PROBLEM. *Suppose that $\ell \geq 5$ and that p is a prime distinct from ℓ. Assume that ρ arises from an eigenform f of weight k on $\Gamma_1(M) \cap \Gamma_0(p)$, where M is prime to p, and that ρ is unramified at p. Show that ρ arises from an eigenform on $\Gamma_1(M)$.*

Note that π is automatically special in all nontrivial cases of (4.6). Indeed, let f be an eigenform on $\Gamma_1(M) \cap \Gamma_0(p)$. If the newform associated to f has level dividing M, then in particular ρ arises from an eigenform on $\Gamma_1(M)$. If, to the contrary, the level of the newform associated to f is divisible by p, then π is special.

We now describe the extent to which (4.6) has been treated. A first remark, which is a consequence of the discussion concerning the four sets $\mathcal{N}_i(\rho)$, is that it suffices to solve (4.6) in *either* of the following situations:

(a) $k = 2$ and M is of the form $N\ell^\alpha$ with N prime to ℓ;

(b) M is prime to ℓ and $2 \leq k \leq \ell+1$.

One can assume $\alpha \leq 2$ in case (a), but this does not seem to be helpful.

(4.7) MAZUR'S PRINCIPLE. *Suppose that $\ell \geq 5$ and that $p \neq \ell$ is a prime such that $p \not\equiv 1 \mod \ell$. Assume that ρ arises from an eigenform f of weight k on $\Gamma_1(M) \cap \Gamma_0(p)$, with M prime to p, and that ρ is unramified at p. Then ρ arises from an eigenform on $\Gamma_1(M)$.*

This theorem was proved by Mazur in a letter to J.-F. Mestre [**31**] in the special case where $k = 2$ and f has trivial character, i.e., is a cusp form on $\Gamma_0(Mp)$. Mazur's proof was presented by the author in [**41**, §6]. In proving the theorem more generally, we may apply the analysis above to work either in case (a) or in case (b). We choose to work in situation (a), i.e., in the case of forms of weight $k = 2$. This is essentially the case we have already treated in [**41**, §6]; the main difference is that we now allow the Nebentypus character of f to be nontrivial. (This character is naturally defined mod M, because f is a form on $\Gamma_1(M) \cap \Gamma_0(p)$.) For the convenience of the reader, we establish Mazur's principle, in case (a), in §8. (In §8, we require only that ℓ be odd: we do not exclude the case $\ell = 3$. Also, we use slightly different notation: the prime q of §8 plays the role of p in (4.7).)

After Mazur formulated his principle in 1985, problem 4.6 was studied extensively when $k = 2$ and f is a cusp form on $\Gamma_0(Mp)$. Here, one wishes to show that ρ arises from a weight-two eigenform on $\Gamma_0(M)$. In [**41**], the author proved this result when M is prime to ℓ. Next, in a joint article [**32**], Mazur and the author established a "multiplicity one" theorem which implies, by the techniques of [**41**], that Theorem 1.5 holds when ℓ, but not ℓ^2, divides M. Then, in a recent note [**43**], the author proved Theorem 1.5 when M is "exactly divisible" by a prime q, different from p and ℓ, at which ρ is ramified. The argument given in [**43**] relies heavily on a recent theorem of N. Boston, H. W. Lenstra, Jr., and the author [**4**], but avoids appeal to the "multiplicity one" principle which appears in **[41]**. Theorem 1.5, which will be proved in §7, may be viewed as the most recent link in this chain.

Theorem 1.5 falls short of solving (4.6) because of the requirement that the character of f be trivial. However, as already mentioned in §1, the author expects that the proof of (1.5) will extend without difficulty to the case where the character associated to f is an arbitrary character on $(\mathbf{Z}/M\mathbf{Z})^*$.

Problem 4.6 has been studied from perspective (b) by Jordan and Livné **[24]**, who expect to treat this case (at least) in the case where f has trivial character and the weight k satisfies the inequality $2 \leq k < \ell$. Their arguments are an adaptation to weight k of arguments given in **[41]**. In particular, the arguments of Jordan and Livné rely on a weight-k "multiplicity-one" principle, which will be treated in a forthcoming article of Faltings and Jordan **[16]**. It is quite possible that the arguments of §7 can be adapted to their situation, thereby obviating the necessity of using a weight-k multiplicity-one principle.

The cuspidal case. In this case, the restriction of ρ_λ to D_p is irreducible. Assume that $e > 0$. Then, according to Carayol **[7,** §§1.1–1.2], $\rho_\lambda|_{D_p}$ is the two-dimensional representation induced from an abelian character $\xi: H \to \overline{\mathbf{Q}}^*$, where H is the subgroup of index two in D_p that corresponds to the unramified quadratic extension of $\mathbf{Q}_p$. Moreover, the reduction of ξ mod v is unramified. The level N is divisible exactly by p^2; i.e., it is divisible by p^2, but not by p^3. The Serre invariant $N(\rho)$ is divisible at most by p. One wishes to show that ρ arises from an eigenform on $\Gamma_1(N/p)$ of some weight, i.e., that ρ arises from $\Gamma_1(N/p)$.

For this, we can (and will) assume that the weight k of f satisfies $2 \leq k \leq \ell+1$.

By Theorem 5.1, there are infinitely many prime numbers q, prime to N and congruent to -1 mod ℓ, such that ρ arises from a weight-k eigenform on the group $\Gamma_1(N) \cap \Gamma_0(q)$ for which the associated newform has level divisible by q. (One takes q to be an "auxiliary prime" as defined in §5.) Fix such a prime, together with an eigenform F satisfying the given condition. Let $\pi_{\mathbf{A}}$ be the representation of $\mathbf{GL}(2, \mathbf{A})$ associated with F. In the language of automorphic representations, the condition on F ensures that the component at q of $\pi_{\mathbf{A}}$ is a special representation.

Consider the twisted form of $\mathbf{GL}(2)$ which arises from the quaternion algebra over $\mathbf{Q}$ which is ramified exactly at p and at q. Since the components at both p and q of $\pi_{\mathbf{A}}$ are discrete series representations, we can conclude that $\pi_{\mathbf{A}}$ arises by the Jacquet-Langlands correspondence from an automorphic representation on this twisted form. Using this correspondence and a lemma which generalizes (1.3), Carayol proves that ρ arises from an eigenform on $\Gamma_1(N/p) \cap \Gamma_0(q)$. Note that, since ℓ is odd, the congruence $q \equiv -1$ mod ℓ implies that one has $q \not\equiv +1$ mod ℓ. Using this fact, together with Mazur's Principle (4.7), we conclude, as desired, that ρ is modular of level N/p.

5. Diamond's theorem

This section concerns the result of F. Diamond [**13**] that was alluded to above. We establish an analogue (Theorem 5.1) of a result proved by the author [**38, 42**] in the case of weight-two forms on $\Gamma_0(N)$. This analogue is a simple variant of Theorem 1 of [**13**]. More precisely, we shall focus attention on some intermediate results obtained by Diamond during the course of the proof of [**13**, Theorem 1] and then derive Theorem 5.1 from these.

Results like Theorem 5.1 have been useful both in "level lowering" (as in §4) and in other contexts. For example, [**5**] applies the theorem of [**42**] to the study of deformations of mod ℓ representations arising from weight-two modular forms.

Let $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ be an irreducible mod ℓ representation arising from $S_k(\Gamma_1(N))$, with N prime to ℓ. Consider the following two conditions on a prime number q, assumed to be prime to $N\ell$:

I. The representation ρ arises from a weight-k eigenform on $\Gamma_1(N)\cap\Gamma_0(q)$ for which the associated newform has level divisible by q.
II. The characteristic polynomial of $\rho(\mathrm{Frob}_q)$ is of the form $(T-a)(T-qa)$, with $a \in \mathbf{F}^*$.

(**5.1**) THEOREM (Diamond). *Assume that* k *satisfies* $2 \le k \le \ell+1$. *Then conditions* I *and* II *are equivalent.*

We first make some comments concerning the two conditions. Let $\mathscr{S}$ be the space $S_k(\Gamma_1(N)\cap\Gamma_0(q))$, and let $\mathscr{S}^{q\text{-new}}$ be the "q-new" subspace of $\mathscr{S}$, defined, for example, as the kernel of the natural trace map from $\mathscr{S}$ to the direct sum of two copies of $S_k(\Gamma_1(N))$. Condition I means simply that ρ arises from an eigenform in the space $\mathscr{S}^{q\text{-new}}$.

Next, we should point out that Condition II is satisfied in the case where q is an *auxiliary prime.* Let σ be the three-dimensional representation $\rho \times \chi$, where the mod ℓ cyclotomic character χ is regarded as a one-dimensional representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ over $\mathbf{F}$. An auxiliary prime is a prime number q, prime to ℓN, such that $\sigma(\mathrm{Frob}_q)$ is conjugate to the matrices $\sigma(c)$, with $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ a complex conjugation. If q is such a prime, the characteristic polynomial of $\rho(\mathrm{Frob}_q)$ coincides with that of the $\rho(c)$. Since ρ is an odd representation, $\rho(c)$ is conjugate to $\left(\begin{smallmatrix}-1 & 0\\ 0 & 1\end{smallmatrix}\right)$ and thus has characteristic polynomial $(T-1)(T+1)$. Because $\chi(\mathrm{Frob}_q) = \chi(c) = -1$, we have $q \equiv -1 \pmod{\ell}$. Thus the characteristic polynomial of $\rho(\mathrm{Frob}_q)$ may be written $(T-a)(T-qa)$, where a can be taken to be either $+1$ or -1, so that condition II is satisfied.

By the Čebotarev Density Theorem, the set of prime numbers for which c and Frob_q map to conjugate elements in the image of $\rho \times \chi$ has positive density. In particular, there are an infinite number of auxiliary primes.

Also, we should point out that the implication "Condition I $\Rightarrow$ Condition II" is a direct consequence of Langlands's theorem ([**28**], see also [**6**]). Indeed, suppose that I is satisfied, and let f be a newform in $\mathscr{S}^{q\text{-new}}$ which gives rise to ρ. Let π be the admissible representation of $\mathbf{GL}(2, \mathbf{Q}_q)$ which is associated to f. The fact that f is associated with $\Gamma_1(D) \cap \Gamma_0(q)$, for some $D|N$, implies that π is a special representation $\alpha \otimes \mathrm{sp}$, with α unramified. As discussed above, a theorem of Langlands states that $\rho_\lambda|_{D_q}$ is an extension of the one-dimensional representation with character α by the one-dimensional representation with character $\alpha\tilde{\chi}$. The semisimplification of $\rho|_{D_q}$ has the form $\bar{\alpha}\oplus\bar{\alpha}\chi$, where $\bar{\alpha}$ is the reduction of α mod λ. Applying $\bar{\alpha}\oplus\bar{\alpha}\chi$ to $\mathrm{Frob}_q \in D_q$, we obtain the matrix $\left(\begin{smallmatrix} a & 0 \\ 0 & qa \end{smallmatrix}\right)$, with $a = \alpha(\mathrm{Frob}_q)$. The characteristic polynomial of this matrix is $(T-a)(T-qa)$. Hence condition I implies condition II.

We now turn to the implication "Condition II $\Rightarrow$ Condition I," which is the essential content of Theorem 5.1. We first compare eigenforms in $S_k(\Gamma_1(N))$ and in $\mathscr{S}$. Suppose that ρ arises from a weight-k eigenform $f = \sum c_n e^{2\pi i n\tau}$ on $\Gamma_1(N)$, and let ε be the character of $(\mathbf{Z}/N\mathbf{Z})^*$ associated with f. (For obvious reasons, we must refrain from writing $q = e^{2\pi i\tau}$.) The forms f and $f' := \sum c_n(e^{2\pi i n\tau})^q$ may be considered as elements of $\mathscr{S}$; indeed, this construction defines the standard inclusion of $S_k(\Gamma_1(N))\oplus S_k(\Gamma_1(N))$ into $\mathscr{S}$ as a space of oldforms. Both forms f and f' in $\mathscr{S}$ are eigenvectors for the diamond operators $\langle d\rangle$ with $d \in (\mathbf{Z}/N\mathbf{Z})^*$ and the Hecke operators T_n with $(n, q) = 1$; the eigenvalues are the same as those for f when considered as an element of $S_k(\Gamma_1(N))$. (Some authors would write $\langle d\rangle_N$ for the operator $\langle d\rangle$ to emphasize that only the group $(\mathbf{Z}/N\mathbf{Z})^*$, rather than $(\mathbf{Z}/Nq\mathbf{Z})^*$, is operating. The group $(\mathbf{Z}/Nq\mathbf{Z})^*$ operates through its quotient $(\mathbf{Z}/N\mathbf{Z})^*$ because we are considering cusp forms on $\Gamma_0(q)\cap\Gamma_1(N)$.) However, neither of the forms f, f' is (in general) an eigenvector for the q^{th} Hecke operator $U := T_q$ on $\mathscr{S}$.

In order to obtain a true eigenform in $\mathscr{S}$, we consider the two roots α and β of the polynomial $T^2 - c_q T + q^{k-1}\varepsilon(q)$. A computation shows that the two forms $f - \alpha f'$ and $f - \beta f'$ are eigenvectors for U on $\mathscr{S}$, with eigenvalues β and α, respectively.

Consider the Hecke algebra $\mathbf{T} = \mathbf{T}_{\mathscr{S}}$ associated with $\mathscr{S}$, i.e., the subring of $\mathrm{End}\,\mathscr{S}$ generated by the T_n with $n \geq 1$. We have $\langle d\rangle \in \mathbf{T}$ for all $d \in (\mathbf{Z}/N\mathbf{Z})^*$. Indeed, if r is a prime number that is prime to qN, then $T_r^2 - T_{r^2} = \langle r\rangle r^{k-1}$, so that $r^{k-1}\langle r\rangle \in \mathbf{T}$. By taking two distinct r's that map to d mod N, we find $\langle r\rangle \in \mathbf{T}$.

The action of $\mathbf{T}$ on $g := f - \beta f'$ is given by a homomorphism $\varphi\colon \mathbf{T} \to \mathbf{C}$, which is characterized by the formulas $\varphi(\langle d\rangle) = \varepsilon(d)$, $\varphi(U) = \alpha$, and $\varphi(T_n) = c_n$, the latter valid for n prime to q. We may view φ as a homomorphism $\mathbf{T} \to \mathscr{O}$, where $\mathscr{O}$ is the integer ring of a suitable number field.

Our fixed place $v|\ell$ of $\overline{\mathbf{Q}}$ induces a prime $\lambda|\ell$ of $\mathscr{O}$, together with an inclusion $\mathscr{O}/\lambda \hookrightarrow \mathbf{F}$. On composing φ with the map $\mathscr{O} \to \mathbf{F}$, we obtain a ring homomorphism $\bar{\varphi}: \mathbf{T} \to \mathbf{F}$. This map depends on f and on (α, β) as an ordered pair.

Assume now that condition II is satisfied and fix $a \in \mathbf{F}^*$ as in the condition. Then $T^2 - c_q T + q^{k-1}\varepsilon(q) \equiv (T-a)(T-qa) \bmod \lambda$. After permuting the roots α and β if necessary, we have $\alpha \equiv a \bmod \lambda$. Setting

$$\eta = U^2 - q^{k-2}\langle q\rangle,$$

we find from condition II that $\bar{\varphi}(\eta) = 0$. Indeed, one has $\bar{\varphi}(q\eta) = qa^2 - \bar{\varepsilon}(q)q^{k-1}$, but $\varepsilon(q)q^{k-1}$ is congruent to qa^2 in view of condition II. In other words, the kernel $\mathfrak{m}$ of $\bar{\varphi}$ is a maximal ideal of $\mathbf{T}$ that contains η.

The subspace $\mathscr{S}^{q\text{-new}}$ of $\mathscr{S}$ is stable under all Hecke operators T_n and $\langle d\rangle$. Thus $\mathscr{S}^{q\text{-new}}$ cuts out a quotient $\mathbf{T}_{q\text{-new}}$ of $\mathbf{T}$: the image of $\mathbf{T}$ in $\operatorname{End}\mathscr{S}^{q\text{-new}}$. One verifies condition I by proving that $\mathfrak{m}$ arises by pull-back from a maximal ideal $\overline{\mathfrak{m}}$ of $\mathbf{T}_{q\text{-new}}$, i.e., that $\mathfrak{m}$ lies in the support of the $\mathbf{T}$-module $\mathscr{S}^{q\text{-new}}$. It is well known that this property of $\mathfrak{m}$ implies that ρ arises from an eigenform in $\mathscr{S}$; see **[17,** Proposition 9.3] and **[12,** Lemma 6.11], and the proofs of these results. Therefore, to prove that condition II implies condition I, it suffices to verify

(5.2) Proposition. *Suppose that ℓ is prime to N and that k satisfies $2 \le k \le \ell+1$. Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}$ associated with $\mathscr{S} = S_k(\Gamma_1(N)\cap\Gamma_0(q))$, where q is a prime number not dividing $N\ell$. Assume that $\mathfrak{m}$ divides ℓ (i.e., $\mathfrak{m}$ contains ℓ) and that $\mathfrak{m}$ contains $\eta = U^2 - q^{k-2}\langle q\rangle$. Then $\mathfrak{m}$ arises by pull-back from the q-new quotient of $\mathbf{T}$.*

This proposition is implicit in Diamond's work, although it is not explicitly stated in **[13]**.

To prove the proposition, we recall some results from **[13]**, using somewhat different notation. Let L be the parabolic cohomology group with $\mathbf{Z}_\ell$-coefficients that is associated with $\mathscr{S}$, so that L is a free $\mathbf{Z}_\ell$-module of rank $2\dim_{\mathbf{C}}\mathscr{S}$. To construct L, we let Γ be the image of $\Gamma_1(N)\cap\Gamma_0(q)$ in $\mathbf{PSL}(2, \mathbf{Z})$ and consider the parabolic cohomology group

$$M := H^1_{\mathrm{par}}(\Gamma, \operatorname{Sym}^{k-2}(\mathbf{Z}_\ell \oplus \mathbf{Z}_\ell)).$$

This group is a finitely generated $\mathbf{Z}_\ell$-module which is not necessarily torsion free; L is defined as the image of M in $M \otimes \mathbf{Q}_\ell$, i.e., the largest torsion-free quotient of M. We let $X = X_1 \oplus X_2$ be the direct sum of two copies of the corresponding parabolic cohomology group made with $\Gamma_1(N)$ in place of $\Gamma_1(N)\cap\Gamma_0(q)$. Diamond considers the degeneracy map $X \hookrightarrow L$ that corresponds to the inclusion of $S_k(\Gamma_1(N)) \oplus S_k(\Gamma_1(N))$ in $\mathscr{S}$ on the level of modular forms. We will denote this map by α. (There is no further need to refer to the roots α and β that occurred in the discussion above.) The map α becomes $\mathbf{T}$-equivariant when one endows X with its natural action of $\mathbf{T}$. (In this action, the T_n with n prime to q act "diagonally" on X, while U acts as the two-by-two matrix

$$\left(\begin{array}{cc} T_q & q^{k-1} \\ -\langle q\rangle & 0 \end{array}\right).)$$

In particular, the image of α is stable under $\mathbf{T}$.

Using results of Ihara and the hypotheses on ℓ, Diamond proves that the cokernel Y of α is *torsion free.* The tautological exact sequence

$$0 \to X \to L \to Y \to 0 \tag{5.3}$$

is thus an exact sequence of free $\mathbf{Z}_\ell$-modules. The quotient of $\mathbf{T} \otimes \mathbf{Z}_\ell$ cut out by Y is the ℓ-adic completion of the q-new quotient of $\mathbf{T}$, while the quotient of $\mathbf{T} \otimes \mathbf{Z}_\ell$ cut out by X is the completion of the analogous q-old quotient of $\mathbf{T}$; this is the same quotient that is defined by the action of $\mathbf{T}$ on $S_k(\Gamma_1(N)) \oplus S_k(\Gamma_1(N))$.

In particular, let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}$ with residue characteristic ℓ. Because $\mathbf{T}$ acts faithfully on L, $\mathfrak{m}$ lies in the support of L. In view of (5.3), it follows that $\mathfrak{m}$ lies either in $\operatorname{Supp} X$ or in $\operatorname{Supp} Y$ and perhaps in both. The statement to be proved may be rephrased as follows: if $\mathfrak{m} \in \operatorname{Supp} X$ and if $\eta \in \mathfrak{m}$, then $\mathfrak{m} \in \operatorname{Supp} Y$. Note that the two hypotheses on $\mathfrak{m}$ may be combined into the statement that $\mathfrak{m}$ belongs to the support of $X/\eta X$. Indeed, if $\mathfrak{m} \in \operatorname{Supp} X$, then $X/\mathfrak{m}X$ is nonzero, by Nakayama's lemma. If furthermore we have $\eta \in \mathfrak{m}$, then $X/\mathfrak{m}X$ is a quotient of $X/\eta X$. Hence $\mathfrak{m}$ belongs to the support of $X/\eta X$, since $\mathfrak{m}$ belongs to the support of $X/\mathfrak{m}X$. Conversely, if $\mathfrak{m}$ lies in the support of $X/\eta X$, it is clear that η is an element of $\mathfrak{m}$ and that $\mathfrak{m} \in \operatorname{Supp} X$. Thus one must prove

$$\text{If } \mathfrak{m} \in \operatorname{Supp}(X/\eta X), \text{ then } \mathfrak{m} \in \operatorname{Supp} Y. \tag{5.4}$$

Diamond proves (5.4) by constructing a $\mathbf{T}$-equivariant surjection $Y \to X/\eta X$. To construct this map, we first dualize (5.3), thereby obtaining an exact sequence

$$0 \to \operatorname{Hom}(Y, \mathbf{Z}_\ell) \to \operatorname{Hom}(L, \mathbf{Z}_\ell) \to \operatorname{Hom}(X, \mathbf{Z}_\ell) \to 0.$$

As Diamond recalls, perfect pairings constructed by Hida [**19,** Theorem 3.2] identify $\operatorname{Hom}(L, \mathbf{Z}_\ell)$ with L and $\operatorname{Hom}(X, \mathbf{Z}_\ell)$ with X. Using these pairings, we obtain an exact sequence

$$0 \to \operatorname{Hom}(Y, \mathbf{Z}_\ell) \to L \xrightarrow{\beta} X \to 0. \tag{5.5}$$

The map β is "almost" $\mathbf{T}$-equivariant: it intertwines the maps labeled T_n on L and on X whenever $(n, q) = 1$, but intertwines the q^{th} Hecke operator U of L with the endomorphism

$$\left(\begin{array}{cc} 0 & q \\ -\langle q\rangle q^{k-2} & T_q \end{array}\right)$$

of X (cf. [**19,** Proposition 3.3]).

Consider the automorphism

$$\omega = \begin{pmatrix} -q^{k-2}\langle q\rangle & T_q \\ 0 & -\langle q\rangle \end{pmatrix}$$

of X. A computation shows that the composite $\omega \circ \beta \colon L \to X$ is **T**-equivariant. Moreover, $(\omega \circ \beta) \circ \alpha$ is the endomorphism η of X. It follows that $\omega \circ \beta$ induces a surjection $L/\alpha X \to X/\eta X$. Since Y is, by definition, $L/\alpha X$, we obtain the desired map.

6. Character groups and component groups

In this section, we recall some material from **[41]** concerning the component groups and character groups associated with bad reductions of ordinary modular curves and Shimura curves. In addition, we explore in further detail a relation that was first found by Jordan and Livné [23] and then deepened by [**41,** Theorem 4.3].

For each $N \geq 1$, we let $X_0(N)$ be the classical modular curve (over $\mathbf{Q}$) which classifies elliptic curves that are endowed with cyclic subgroups of order N. The space of holomorphic differentials on $X_0(N)_{/\mathbf{C}}$ may be identified with $S_2(\Gamma_0(N))$ in a canonical fashion. The curve $X_0(N)$ comes equipped with a family of correspondences T_n ($n \geq 1$); the correspondence T_n induces on $S_2(\Gamma_0(N))$ the endomorphism T_n. The Jacobian $\mathrm{Pic}^0 X_0(N)$ will be denoted $J_0(N)$, as usual; we write again simply T_n for the endomorphism of $J_0(N)$ that is induced by the correspondence T_n of $X_0(N)$.

Consider first a prime number q and a positive integer N prime to q. We are interested in the modular curves $X_0(qN)$ and $X_0(N)$ and in their Jacobians $J_0(qN)$ and $J_0(N)$. The two standard degeneracy coverings

$$X_0(qN) \rightrightarrows X_0(N)$$

induce, by pull-back, a map $\delta \colon J_0(N) \times J_0(N) \to J_0(qN)$. The kernel of δ is finite; it is the "antidiagonal" image in $J_0(N) \times J_0(N)$ of the kernel of the pull-back of the natural covering $X_1(N) \to X_0(N)$ [**38**]. (The curve $X_1(N)$ classifies elliptic curves that are furnished with a point of order N.)

For each $n \geq 1$, a Hecke operator labeled T_n acts diagonally on the product $J_0(N) \times J_0(N)$, and an identically-named operator acts on the target $J_0(qN)$ of δ. When n is prime to q, δ is compatible with the two operators T_n. However, when $n = q$, the situation is more subtle. Reserve the symbol T_q for the q^{th} Hecke operator on $J_0(N)$ and let U denote the q^{th} Hecke operator on $J_0(qN)$. Then $U \circ \delta = \delta \circ U$, where the latter operator U denotes the matrix

$$\begin{pmatrix} T_q & q \\ -1 & 0 \end{pmatrix}$$

of endomorphisms of $J_0(N)$. (We view this matrix as an endomorphism of $J_0(N) \times J_0(N)$.) Let $\mathbf{T} = \mathbf{T}_{qN}$ be the subring of $\mathrm{End}\, J_0(qN)$ generated by the T_n for $n \geq 1$, i.e., by the T_n for n prime to q and the operator

$U = T_q$. We endow $J_0(N) \times J_0(N)$ with the operation of $\mathbf{T}$ such that the elements U and T_n of $\mathbf{T}$ (with n prime to q) act as the operators of $J_0(N) \times J_0(N)$ with the same names. This is the unique operation for which δ is $\mathbf{T}$-equivariant.

We note in passing that the endomorphisms T_n and U and the homomorphism δ are defined over $\mathbf{Q}$. Also, in analogy with the situation of §5, we let $\eta = U^2 - 1$ in $\mathbf{T}$.

(6.1) Theorem. *There is a unique homomorphism of abelian varieties*

$$\sigma\colon J_0(qN) \to J_0(N) \times J_0(N)$$

such that $\sigma \circ \delta = \eta$. This homomorphism is $\mathbf{T}$-equivariant.

To construct σ, we consider the map $\delta'\colon J_0(qN) \to J_0(N) \times J_0(N)$ which is induced by the degeneracy coverings $X_0(qN) \rightrightarrows X_0(N)$ using Albanese (i.e., covariant) functoriality of the Jacobian. The definition of T_q as a correspondence on $X_0(N)$ shows that

$$\delta' \circ \delta = \begin{pmatrix} q+1 & T_q \\ T_q & q+1 \end{pmatrix}.$$

Define

$$\sigma = \begin{pmatrix} -1 & T_q \\ 0 & -1 \end{pmatrix} \circ \delta'.$$

A computation shows that $\sigma \circ \delta$ coincides with η as a matrix of endomorphisms of $J_0(N)$.

For the unicity and the $\mathbf{T}$-equivariance, we observe that

$$\mathrm{Hom}(Q,\, J_0(N)\times J_0(N)) = 0,$$

where $Q = J_0(qN)/\delta(J_0(N)\times J_0(N))$. This fact may be seen arithmetically, by noting that Q has purely toric reduction at the prime q **[11]**, whereas $J_0(N)$ has good reduction at q. It proves the unicity, since the difference between two operators σ that satisfy $\sigma \circ \delta = \eta$ is an operator that vanishes on the image of δ. Similarly, let $T \in \mathbf{T}$, and consider $\sigma \circ T - T \circ \sigma$, which is a priori a homomorphism $h\colon J_0(qN) \to J_0(N)\times J_0(N)$. We have $h \circ \delta = \eta T - T\eta = 0$, so that h factors through Q; as a consequence, $h = 0$. □

Now specialize to the case where $N = pM$, where p is a prime number that does not divide M. The level qN becomes the product pqM in which the two prime numbers p and q play symmetrical roles. The space $\mathscr{S}$ of weight-two cusp forms on $\Gamma_0(pqM)$ is then analogous to the space denoted $\mathscr{S}$ in §5, in the special case $N = pM$, $k = 2$. The main difference is that we have replaced $\Gamma_1(N) \cap \Gamma_0(q)$ by $\Gamma_0(qN)$.

Let $J_0(pqM)_{/\mathbf{F}_p}$ be the fiber at p of the Néron model of $J_0(pqM)$. This group variety is an extension of a finite "component group" $\mathbf{\Theta}_p$ by a connected group $J_0(pqM)^0_{/\mathbf{F}_p}$, which is in turn an extension of the product of two copies of $J_0(qM)$ by a torus $\mathcal{T}$ over $\mathbf{F}_p$. Let L_p be the character group of this torus, and let L_q be the character group of the analogous torus for $J_0(pqM)_{/\mathbf{F}_q}$. Similarly, let X_p be the analogue of L_p for $(J_0(pM)\times J_0(pM))_{/\mathbf{F}_p}$ and let X_q be the analogue of X_p with p replaced by q. The groups X_p and X_q are naturally direct sums of two copies of the character groups coming from $J_0(pM)_{/\mathbf{F}_p}$ and $J_0(qM)_{/\mathbf{F}_q}$, respectively; thus, we may represent endomorphisms of X_p, say, as two-by-two matrices of endomorphisms of the character group associated to $J_0(pM)_{/\mathbf{F}_p}$. The endomorphisms T_n (and U) of $J_0(pqM)$, $J_0(pM)\times J_0(pM)$, and $J_0(qM)\times J_0(qM)$ induce maps on the four character groups L_p, X_p, L_q, and L_p. In what the author hopes is an acceptable abuse of notation, these maps will be denoted simply T_n and U. In particular, the endomorphism U of X_p is the matrix of endomorphisms

$$\begin{pmatrix} T_q & -1 \\ q & 0 \end{pmatrix}.$$

Continuing the abuse of notation, we will write simply δ for the map $L_p \to X_p$ induced by the degeneracy map $\delta\colon J_0(Mp)\times J_0(Mp) \to J_0(Mpq)$. The surjectivity of $\delta\colon L_p \to X_p$ was established in [**41**, Theorem 3.15]. (In [**41**], the author wrote "X" for the character group associated to a *single* copy of $J_0(pM)$ or $J_0(qM)$.) Theorem 4.1 of [**41**] identifies the kernel of δ with an analogue of L_q in which $X_0(pqM)$ is replaced by an appropriate Shimura curve. Namely, let C be the Shimura curve made with a quaternion algebra of discriminant pq and $\Gamma_0(M)$-type level structure (cf. [**41**, §4]). Let J be the Jacobian of C, and let Y_p and Y_q be the analogues of L_p and L_q for J. The groups $J_{/\mathbf{F}_p}$ and $J_{/\mathbf{F}_q}$ are extensions of their component groups Ψ_p and Ψ_q by the tori whose character groups are Y_p and Y_q. (As was proved by Čerednik [**8**] and Drinfeld [**14**], all components of $C_{/\mathbf{F}_p}$ and $C_{/\mathbf{F}_q}$ have genus zero.) In other words, the reductions of J at p and q are "semiabelian," with trivial abelian variety parts. Then Theorem 4.1 of [**41**] provides an exact sequence

(6.2) $$0 \to Y_q \xrightarrow{\iota} L_p \xrightarrow{\delta} X_p \to 0$$

and an analogue

(6.3) $$0 \to Y_p \to L_q \to X_q \to 0$$

in which the roles of p and q have been permuted. These sequences are compatible with the Hecke operators labeled T_n on J and on $J_0(pqM)$; note, incidentally, that the operators T_p and T_q on J are in fact *involutions.* (See [**41**, Theorem 4.1] and also the results proved in [**40**].) The sequences (6.2) and (6.3) seem to play a role analogous to that played by (5.3) and its $\mathbf{Z}_\ell$-dual in the discussion of §5.

One deduces from (6.2) or (6.3) that the Hecke ring $\mathbf{Z}[\dots, T_n, \dots]$ in $\operatorname{End} J$ associated to J is naturally a quotient of the ring $\mathbf{T}$ associated to $J_0(pqM)$. Thus, we may speak of the action of $\mathbf{T}$ on J.

Let us introduce the notation $'$ to denote a dual abelian variety; thus $J_0(pqM)'$, for example, will be the abelian variety dual to $J_0(pqM)$. We use T'_n and U' to denote the endomorphisms of the dual objects that are induced functorially by the operators T_n and U. Thus, for instance, there is a natural action of the ring $\mathbf{T}$ on J', in which the element of $\mathbf{T}$ labeled T_n acts on J' as the endomorphism T'_n. We let Y'_q, L'_p, and so on, represent the character groups of the tori arising from the dual abelian varieties. The operator T_n (say) in $\mathbf{T}$ then induces an endomorphism of L'_q which will be denoted T'_n; this endomorphism is induced by the endomorphism T'_n of $J_0(pqM)'$. Similar remarks apply to the component groups which we have introduced: Θ_p and Θ_q in the case of $J_0(pqM)$, and Ψ_p and Ψ_q in the case of J.

Next, we discuss monodromy pairings. To fix ideas, we will first consider L_p. The pairing associated with L_p is a bilinear map $\langle\ ,\ \rangle : L_p \times L'_p \to \mathbf{Z}$, which induces an injection $L'_p \hookrightarrow \mathrm{Hom}(L_p, \mathbf{Z})$ with finite cokernel. According to [18, §11], the cokernel of this injection is canonically isomorphic to Θ_p. The exact sequence

$$0 \to L'_p \to \mathrm{Hom}(L_p, \mathbf{Z}) \to \Theta_p \to 0$$

is $\mathbf{T}$-equivariant when the elements of $\mathbf{T}$ act in the natural way, i.e., when $T_n \in \mathbf{T}$ acts as T_n on Θ_p, as T'_n on L'_p, and as $\mathrm{Hom}(T_n, \mathbf{Z})$ on $\mathrm{Hom}(L_p, \mathbf{Z})$. In a similar way, we have monodromy pairings

$$Y_q \times Y'_q \to \mathbf{Z}, \qquad X_p \times X'_p \to \mathbf{Z}.$$

One potential source of confusion is that the abelian varieties under discussion are Jacobians of modular curves or else products of two Jacobians. It follows that they are naturally self-dual. After using the autoduality of the Jacobian to identify J' and J, for example, we find that the lattices Y_q and Y'_q are *equal* or at least canonically isomorphic. In point of fact, it is often fruitful to identify both Y_q and Y'_q with $H_1(\Gamma, \mathbf{Z})$, where Γ is the "dual graph" attached to the mod q reduction of C. The group $H_1(\Gamma, \mathbf{Z})$ thereby acquires two actions of $\mathrm{End}\, J$; these differ by the Rosati involution of $\mathrm{End}\, J$ that arises from the theta divisor on C. Similarly, the group X_p has two natural actions of $\mathbf{T}$, which differ by the Rosati involution

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix}$$

on $\mathrm{End}\, J_0(pM)\times J_0(pM)$, where $'$ is the Rosati involution on $\mathrm{End}\, J_0(pM)$. For example, the operator U, which acts on X_p as

$$\begin{pmatrix} T_q & -1 \\ q & 0 \end{pmatrix}$$

in its standard action, becomes

$$\begin{pmatrix} T'_q & q \\ -1 & 0 \end{pmatrix}$$

in the other action. (It might be worth pointing out that the endomorphisms T_q and T'_q of $J_0(N)$ are equal.) As a mnemonic device, we will attempt whenever possible to use the $'$ notation when the Hecke operators in $\mathbf{T}$ act as T'_n and U'.

These examples show that the natural autoduality $J_0(N) \approx J_0(N)'$ does not necessarily intertwine the actions of $\mathbf{T}$ on these two abelian varieties. Nonetheless, it is a fact that there are natural $\mathbf{T}$-equivariant isomorphisms $J \approx J'$ and $J_0(N) \approx J_0(N)'$ for each N. In the case of $J_0(N)$, the isomorphism is provided by the involution $w = w_N$ on $J_0(N)$ which is induced by the Atkin-Lehner involution $(E, C) \mapsto (E/C, E[N]/C)$ on the set of elliptic curves with cyclic subgroups of order N (see **[52,** Proposition 3.54]). In the case of J, there is an analogous involution w on the set of "fake elliptic curves" which are classified by C. These $\mathbf{T}$-equivariant isomorphisms are occasionally useful.

Once we identify L_p with L'_p, the monodromy pairings associated to L_p and Y_q become maps $L_p \times L_p \to \mathbf{Z}$ and $Y_q \times Y_q \to \mathbf{Z}$, respectively. Both pairings are symmetric; they are given by explicit formulas involving the dual graphs of the mod p reduction of $X_0(pqM)$ and the mod q reduction of C. According to **[41,** Theorem 4.1], the restriction to $Y_q \times Y_q$ of the monodromy pairing $L_p \times L_p \to \mathbf{Z}$ agrees with the monodromy pairing on Y_q. Here, one uses ι to embed Y_q in L_p.

(6.4) PROPOSITION. *The embedding* $\iota\colon Y_q \hookrightarrow L_p$ *remains* $\mathbf{T}$-*equivariant when it is regarded as a map* $Y'_q \to L'_p$.

We regard L'_p and Y'_q as the same physical groups as L_p and Y_q, only with different actions of $\mathbf{T}$. Fix $T \in \mathbf{T}$, and use the symbols T_Y and T_L to denote the endomorphisms of Y_q and of L_p induced by T. Write T'_Y and T'_L for the endomorphisms of these groups which T induces in its "dual" actions. The endomorphism $T_{L'}$ (say) is the adjoint of T_L with respect to the monodromy pairing $\langle\, ,\, \rangle_L$ on L_p. Proposition (6.4) states the formula $T'_L \circ \iota = \iota \circ T'_Y$.

Consider the homomorphism $\kappa\colon J_0(Mp)' \times J_0(Mp)' \to J_0(Mpq)'$ which is dual to the map σ of (6.1). It induces a $\mathbf{T}$-equivariant map $\kappa\colon L'_p \to X'_p$; this map is adjoint to the map $X_p \to L_p$ induced by σ. The construction of σ shows that κ differs from δ by an automorphism of X_p. Therefore, the kernel of κ coincides with the kernel of δ, which is the image of ι. It follows that T'_L preserves this image, since we may write $\kappa \circ T'_L \circ \iota = T'_X \circ \kappa \circ \iota = 0$. Hence, the map $\iota \circ T'_Y - T'_L \circ \iota$ maps Y_q to $\iota(Y_q)$. Consequently, to show that $\iota \circ T'_Y = T'_L \circ \iota$, it suffices to verify the equality of $\langle T'_Y y, z\rangle_Y$ and $\langle T'_L \circ \iota y, \iota z\rangle_L$ for all $y, z \in Y_q$. Here, we make use of the fact that the monodromy pairing $\langle\, ,\, \rangle_Y$ on Y_q is the restriction to Y_q of the pairing $\langle\, ,\, \rangle_L$ on L_p.

To verify the desired equation, we note the series of equalities

$$\langle T_L' \iota y, \iota z\rangle_L = \langle \iota y, T_L \iota z\rangle_L = \langle \iota y, \iota T_Y z\rangle_L = \langle y, T_Y z\rangle_Y = \langle T_Y' y, z\rangle_Y.$$

These follow from the adjunction relations and the equivariance of ι with respect to the standard actions of $\mathbf{T}$, plus the compatibility between the two monodromy pairings. □

(6.5) Corollary. *We have an exact sequence of* $\mathbf{T}$-*modules*

$$0 \to Y_q' \xrightarrow{\iota} L_p' \xrightarrow{\kappa} X_p' \to 0.$$

Theorem 4.3 of **[41]** asserts that there is an exact sequence of $\mathbf{T}$-modules

$$0 \to \mathscr{K} \to X_p/\eta X_p \to \Psi_q \to \mathscr{C} \to 0.$$

Here, $\mathscr{K}$ and $\mathscr{C}$ are each closely related to Θ_p and to the component group of the group variety $(J_0(pM)\times J_0(pM))_{/\mathbf{F}_p}$; in particular, they are Eisenstein $\mathbf{T}$-modules in the sense that they are annihilated by $T_r-(r+1)$ for almost all prime numbers r; cf. **[41,** Theorem 3.12] or **[39]**. (This allows us to neglect $\mathscr{C}$ and $\mathscr{K}$ in practice.) The author now believes that he was not sufficiently attentive to the distinction between the two actions of $\mathbf{T}$ and that the exact sequence of **[41,** Theorem 4.3] should read

$$0 \to \mathscr{K}' \to X_p/\eta X_p \to \Psi_q' \to \mathscr{C}' \to 0, \tag{6.6}$$

where $\mathscr{C}'$ and $\mathscr{K}'$ are variants of $\mathscr{K}$ and $\mathscr{C}$, which are again Eisenstein. Note, however, that Ψ_q and Ψ_q' are *isomorphic* as $\mathbf{T}$-modules, in view of the Atkin-Lehner automorphism $w: J \approx J'$ which intertwines the two actions of $\mathbf{T}$. Hence **[41,** Theorem 4.3] can be used as stated.

To derive a version of (6.6), we consider the monodromy pairings on L_p and X_p as injections

$$X_p \hookrightarrow \mathrm{Hom}(X_p', \mathbf{Z}) \quad \text{and} \quad L_p \hookrightarrow \mathrm{Hom}(L_p', \mathbf{Z}),$$

respectively. Their cokernels are the component groups Φ_p' and Θ_p' associated with the mod p reductions of $J_0(pM)\times J_0(pM)$ and $J_0(pqM)$. Using ι, we regard Y_q as a submodule of L_p. We consider the $\mathbf{T}$-equivariant maps $\sigma: X_p \to L_p$ and $\kappa: L_p' \to X_p'$ induced by the homomorphism

$$\sigma: J_0(pqM) \to J_0(pM)\times J_0(pM)$$

and use the abbreviation κ^* to refer to the map $\mathrm{Hom}(\kappa, \mathbf{Z})$ which is the $\mathbf{Z}$-linear dual of κ. We have a commutative square

$$\begin{array}{ccc} L_p & \to & \mathrm{Hom}(L_p', \mathbf{Z}) \\ \uparrow \sigma & & \uparrow \kappa^* \\ X_p & \to & \mathrm{Hom}(X_p', \mathbf{Z}) \end{array}$$

in which the horizontal maps are the injections coming from the two monodromy pairings. On dividing L_p by its submodule Y_q, we obtain a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & L_p/Y_q & \to & \mathrm{Hom}(L'_p, \mathbf{Z})/Y_q & \to & \Theta'_p & \to & 0 \\ & & \uparrow \sigma & & \uparrow \kappa^* & & \uparrow & & \\ 0 & \to & X_p & \to & \mathrm{Hom}(X'_p, \mathbf{Z}) & \to & \Phi'_p & \to & 0 \end{array}$$

in which the rows are exact. The map $\delta\colon L_p \to X_p$ identifies L_p/Y_q with X_p. Also, we have on X_p the identity $\delta \circ \sigma = \eta$, in view of the equation $\sigma \circ \delta = \eta$ involving homomorphisms of abelian varieties. Hence the diagram may be rewritten

$$\begin{array}{ccccccccc} 0 & \to & X_p & \to & \mathrm{Hom}(L'_p, \mathbf{Z})/Y_q & \to & \Theta'_p & \to & 0 \\ & & \uparrow \eta & & \uparrow \kappa^* & & \uparrow & & \\ 0 & \to & X_p & \to & \mathrm{Hom}(X'_p, \mathbf{Z}) & \to & \Phi'_p & \to & 0. \end{array}$$

In view of (6.5), the cokernel of κ^* is $\mathrm{Hom}(Y'_q, \mathbf{Z})$. Therefore, the cokernel of the middle vertical map is the finite group $\mathrm{Hom}(Y'_q, \mathbf{Z})/Y_q = \Psi'_q$. By a dimension count, this shows, in particular, that the middle vertical map is injective. From the snake lemma, we deduce a version of (6.6) in which the groups $\mathscr{K}'$ and $\mathscr{C}'$ are the kernel and cokernel of the right-hand vertical map.

(6.7) PROPOSITION. *The* $\mathbf{T}$*-modules* Ψ_q *and* Ψ'_q *are mutually* $\mathbf{Q}/\mathbf{Z}$*-dual.*

PROOF. The indicated duality between the two component groups may be read directly from the recipe for making Ψ_q and Ψ'_q from the monodromy pairing on Y_q. For more details, see [**18**, §11.4], which treats the more general case of an abelian variety with semistable reduction, and also [**18**, §11.3], which discusses the prime-to-p parts of component groups associated to abelian varieties over a local field of residue characteristic p that do not necessarily have semistable reduction. □

As O. Gabber and R. Livné explained to the author, Grothendieck conjectured in [**18**, §1.2.1] that there is a perfect $\mathbf{Q}/\mathbf{Z}$-duality between the component groups associated to two dual abelian varieties over a local field. In [**18**], Grothendieck did not treat the p-primary parts of component groups associated to arbitrary abelian varieties over local fields of residue characteristic p. This gap was filled, in the case of a perfect residue field, by L. Bégueri as Theorem 8.3.3 of [**3**]. Another proof, for an abelian variety over the fraction field of a discrete valuation ring with finite residue field, is given in [**33**].

In the statement of the following corollary, we introduce Mazur's notation "$[\mathfrak{m}]$" for a kernel: if $\mathfrak{m}$ is an ideal of $\mathbf{T}$ and M is a $\mathbf{T}$-module, then $M[\mathfrak{m}]$ denotes the submodule of M consisting of those element of M that are annihilated by all elements of $\mathfrak{m}$.

(6.8) COROLLARY. *Let* $\mathfrak{m} \subset \mathbf{T}$ *be a maximal ideal. Then* $\mathfrak{m}$ *lies in the support of* Ψ_q *if and only if* $\mathfrak{m}$ *lies in the support of* Ψ'_q*. Moreover, the*

$\mathbf{T}/\mathfrak{m}$-dimensions of $\Psi_q/\mathfrak{m}\Psi_q$ and $\Psi'[\mathfrak{m}]$ are equal; similarly, the dimensions of $\Psi'_q/\mathfrak{m}\Psi'_q$ and $\Psi_q[\mathfrak{m}]$ coincide.

PROOF. The two equalities of dimension result directly from (6.7). Suppose now that $\mathfrak{m}$ lies in the support of Ψ_q. By Nakayama's Lemma, $\Psi_q/\mathfrak{m}\Psi_q$ is nonzero. According to (6.7), however, $\Psi_q/\mathfrak{m}\Psi_q$ and $\Psi'_q[\mathfrak{m}]$ are duals of each other. Thus $\Psi'_q[\mathfrak{m}]$ is nonzero. This nonvanishing implies immediately that $\mathfrak{m}$ lies in the support of Ψ'_q. Similarly, the support of Ψ'_q is contained in the support of Ψ_q. □

(6.9) THEOREM. *Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}$ for which the associated representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is irreducible. Then $\mathfrak{m}$ lies in the support of Ψ_q if and only if $\mathfrak{m}$ contains η and lies in the support of X_p. Moreover, in this case we have the equalities of $\mathbf{T}/\mathfrak{m}$-dimensions*

$$\dim X_p/\mathfrak{m}X_p = \dim \Psi'_q/m\Psi'_q = \dim \Psi_q[\mathfrak{m}].$$

PROOF. Assume that the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ associated with $\mathfrak{m}$ is irreducible. By **[41,** 5.2c], this hypothesis implies that $\mathfrak{m}$ does not lie in the support of either of the modules $\mathscr{K}'$ or $\mathscr{C}'$ which appear in (6.6). Accordingly, $X_p/\eta X_p$ and Ψ'_q are isomorphic locally at $\mathfrak{m}$. Thus $\mathfrak{m}$ lies in the support of Ψ'_q if and only if it lies in the support of $X_p/\eta X_p$. By (6.8), the supports of Ψ_q and Ψ'_q coincide. Also, the support of $X_p/\eta X_p$ clearly consists of those $\mathfrak{m}$ in the support of X_p that contain η. Finally, whenever $\mathfrak{m}$ is prime to the flanking modules in (6.6) and $\eta \in \mathfrak{m}$, we have an isomorphism between $X_p/\mathfrak{m}X_p$ and $\Psi'_q/m\Psi'_q$. This observation, together with (6.8), gives the assertion concerning dimensions. □

7. Proof of Theorem 1.5

Let ρ, ℓ, M, and p be as in the statement of Theorem 1.5. In other words, we assume that $\ell \geq 3$ is a prime, that $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ is an irreducible representation arising from an eigenform f of weight two on $\Gamma_0(Mp)$, and that p is prime to ℓM. Suppose that ρ is unramified at p. We shall prove that ρ arises from a weight-two eigenform on $\Gamma_0(M)$.

The first step in the proof is the introduction of an "auxiliary prime" into the level (cf. the "cuspidal case" discussion in §4 and **[41,** §7]). Choose a prime number q, prime to $p\ell M$, such that $\rho(\mathrm{Frob}_q)$ is conjugate to each matrix $\rho(c)$, with $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ a complex conjugation. The characteristic polynomial of $\rho(\mathrm{Frob}_q)$ then coincides with that of the $\rho(c)$; since ρ is an odd representation, $\rho(c)$ is conjugate to $\left(\begin{smallmatrix}-1 & 0\\ 0 & 1\end{smallmatrix}\right)$ and thus has characteristic polynomial $(T-1)(T+1)$. Since the determinant of ρ is the mod ℓ cyclotomic character χ, we have $q \equiv -1 \pmod{\ell}$. (Because ℓ is odd, this congruence gives $q \not\equiv +1 \bmod \ell$.) Thus $(T-1)(T+1) = (T-a)(T-qa)$, where a can be taken to be either $+1$ or -1, so that condition II of §5 is satisfied.

An antecedent of Theorem 5.1 implies that ρ arises from a weight-two newform on $\Gamma_0(pqM)$ whose level is divisible by q. (See **[38]**, **[42]**, and **[41,** §7].) Curiously, we will not use this fact in what follows, although it will appear implicitly. Instead, we consider the Hecke algebra $\mathbf{T} = \mathbf{T}_{pqM}$ which appeared in §6 and mimic the construction of the maximal ideal $\mathfrak{m}$ of §5. Namely, let U again be the q^{th} Hecke operator in $\mathbf{T}$, and let $\eta = U^2 - 1$. The construction of §5 enables one to find a maximal ideal $\mathfrak{m}$ of $\mathbf{T}$ such that $\mathfrak{m}$ contains η, together with an embedding $\varphi\colon \mathbf{T}/\mathfrak{m} \hookrightarrow \mathbf{F}$, such that T_r is mapped to $\operatorname{trace}(\rho(\mathrm{Frob}_r))$ for almost all prime numbers r. Fix such a maximal ideal in what follows. (For the purposes of the proof of Theorem 1.5, $\mathbf{T}$ could be defined, more sparsely, as the subring of $\operatorname{End} S_2(\Gamma_0(pqM))$ which is generated by the T_n with n prime to pqM, along with the operator $U = T_q$. Since no "multiplicity-one" theorems appear in the following argument, it is possible, and perhaps desirable, to work with this more economical Hecke algebra.)

Our next step is to replace ρ by a model $\rho_{\mathfrak{m}}\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}/\mathfrak{m})$ of ρ over the finite field $\mathbf{T}/\mathfrak{m}$. Note that ρ is an irreducible finite-dimensional representation of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ over $\mathbf{F}$ and that the characteristic polynomials of $\rho(g)$ lie in $\mathbf{T}/\mathfrak{m}$ for each $g \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. It is well known that this necessary condition for descending ρ to $\mathbf{T}/\mathfrak{m}$ is also sufficient.

To descend ρ directly, we write K for $\mathbf{T}/\mathfrak{m}$ and observe that ρ descends, in any case, to some finite extension L of K. (Only a finite number of matrices are required to describe ρ.) Let W be an $L[\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module that affords a model of ρ over L, and let V be the vector space W, viewed as a module over $K[\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$. (Thus V is the "restriction of scalars" of W from L to K.) Let X be a minimal (nonzero) $K[\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-submodule of W. The commutant $E := \operatorname{End}_{K[\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]} X$ of X is a division algebra by Schur's lemma; since E has finite cardinality, it is a finite commutative field. It follows that the $L[\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module $X \otimes_K L$ has commutant $E \otimes_K L$. Because this commutant is commutative, $X \otimes_K L$ must be multiplicity free.

On the other hand, $X \otimes_K L$ is a submodule of $V \otimes_K L$. This latter module is a direct sum $\bigoplus_\gamma {}^\gamma W$, where γ runs through $\operatorname{Gal}(L/K)$ and where ${}^\gamma W$ represents the twist of W by γ, i.e., the tensor product $W \otimes_L L$ where L is viewed as a base extension of itself via $\gamma\colon L \to L$. The hypothesis on characteristic polynomials ensures that the various ${}^\gamma W$ are isomorphic to W (Brauer-Nesbitt theorem). Hence $V \otimes_K L$ is isomorphic to a sum of copies of W, and it follows in particular that $X \otimes_K L$ is a sum of copies of W. Since $X \otimes_K L$ is multiplicity free, there can be only one summand. This means that $X \otimes_K L$ is isomorphic to W, so that X affords a model of ρ over K.

Now that $\rho_{\mathfrak{m}}$ has been constructed, we will abuse notation by writing simply ρ for this representation.

We will prove Theorem 1.5 by an indirect reasoning: we will assume that ρ is *not* modular of level qM and obtain a contradiction from this assumption.

This reasoning will show that ρ is modular of level qM, but then Mazur's Principle (4.7) will enable us to deduce the desired conclusion that ρ is modular of level M.

We continue to work with the series of abelian varieties which appeared in §6: $J_0(pqM)$, $J_0(pM)^2 = J_0(pM)\times J_0(pM)$, $J_0(qM)^2 = J_0(qM)\times J_0(qM)$, and the Jacobian $J = \mathrm{Pic}^0 C$ of the Shimura curve C. We attach "multiplicities" λ and μ to $J_0(pqM)$ and J in the following manner. Consider the kernel $J[\mathfrak{m}]$ of $\mathfrak{m}$ on $J(\overline{\mathbf{Q}})$. The main theorem of **[4]** implies that this $\mathbf{T}/\mathfrak{m}[\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module is isomorphic to a direct sum $V\oplus\cdots\oplus V$, where V is a two-dimensional $\mathbf{T}/\mathfrak{m}$-vector space with a $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-action equivalent to ρ. We define μ to be the number of factors and refer to it as the multiplicity of ρ in J. The multiplicity λ is defined analogously, by considering $J_0(pqM)[\mathfrak{m}]$ in place of $J[\mathfrak{m}]$.

Assume from now on that ρ is not modular of level qM. Using this assumption, we shall prove that both multiplicities λ and μ are zero. This will give the desired contradiction, since λ is positive. Indeed, $\mathbf{T}$, by definition, operates faithfully on $J_0(pqM)$, and it is well known that this implies the inequality $\lambda > 0$.

To prove that λ and μ are both zero, we shall establish a series of relations among the dimensions of the $\mathbf{T}/\mathfrak{m}$-vector spaces obtained as the mod $\mathfrak{m}$ reductions of the character groups that appear in (6.2) and (6.3). The hypothesis that ρ is not modular of level qM implies that $\mathfrak{m}$ generates the unit ideal in the "p-old" quotient of $\mathbf{T}$. This implies, for example, that $\mathfrak{m}$ does not lie in the support of the $\mathbf{T}$-module X_q, since $\mathbf{T}$ operates on X_q through its p-old quotient.

(7.1) PROPOSITION. *There are exact sequences of* $\mathbf{T}$*-modules*

$$Y_p/\mathfrak{m}Y_p \approx L_q/\mathfrak{m}L_q, \tag{7.2}$$

$$\cdots \to Y_q/\mathfrak{m}Y_q \to L_p/\mathfrak{m}L_p \to X_p/\mathfrak{m}X_p \to 0. \tag{7.3}$$

PROOF. The isomorphism (7.2) and the exact sequence (7.3) are obtained from (6.2) and (6.3) by first localizing at $\mathfrak{m}$ and then reducing mod $\mathfrak{m}$. The localization at $\mathfrak{m}$ of X_q vanishes, as remarked above. □

Proposition 7.1 implies the following relations among $\mathbf{T}/\mathfrak{m}$-dimensions:

$$\dim Y_p/\mathfrak{m}Y_p = \dim L_q/\mathfrak{m}L_q; \tag{7.4}$$

$$\dim L_p/\mathfrak{m}L_p \le \dim Y_q/\mathfrak{m}Y_q + \dim X_p/\mathfrak{m}X_p. \tag{7.5}$$

(7.6) PROPOSITION. *We have*

$$\dim Y_p/\mathfrak{m}Y_p = 2\mu, \qquad \dim L_p/\mathfrak{m}L_p = 2\lambda.$$

As in **[43]**, these equalities are obtained by considering actions of a decomposition group D_p for p in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. By the main theorem of **[4]**, the kernels of $\mathfrak{m}$ on $J_0(pqM)$ and J are direct sums of copies of ρ, first as representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and then (by restriction) as representations of D_p. In particular, they are unramified at p.

By a well-known theorem of Serre-Tate [**51**], the kernel of $\mathfrak{m}$ in the mod p reduction of J may be identified with the group of I_p-invariants in $J[\mathfrak{m}]$, where I_p is the inertia subgroup of D_p. Since $J[\mathfrak{m}]$ is unramified at p, the group of I_p-invariants is all of $J[\mathfrak{m}]$, and therefore $J_{/\mathbf{F}_p}(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ is of dimension 2μ.

Now, as we recalled above, $J_{/\mathbf{F}_p}$ is an extension of its "group of components" Ψ_p by its toric part. Also, $\mathfrak{m}$ does not lie in the support of X_q. By (6.9), or more precisely the variant of (6.9) obtained by interchanging p and q, the $\mathbf{T}$-module Ψ_p is prime to $\mathfrak{m}$. In view of this fact, we may deduce that the toric part of $J_{\overline{\mathbf{F}}_p}(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ has dimension 2μ. This gives $\dim Y_p/\mathfrak{m}Y_p = 2\mu$, which was the first of the two formulas to be proved.

The formula involving L_p is proved by an analogous argument in which we consider the mod p reduction of $J_0(pqM)$. Let A be the fiber at p of the Néron model of $J_0(pqM)$, and let B be the analogue of A for $J_0(qM) \times J_0(qM)$. Thus B is an abelian variety over $\mathbf{F}_p$. As in the discussion involving J, we know that $A(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ is of dimension 2λ. In contrast to the situation for J, however, the group of components of A is "Eisenstein" ([**41**, Theorem 3.12] or [**39**]) and therefore prime to $\mathfrak{m}$ [**41**, Theorem 5.2c]. Thus, if A^0 is the connected component of 0 in A, $A^0(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ is of dimension 2λ.

By results of Deligne-Rapoport [**11**] and Raynaud, there is an exact sequence

$$0 \to T \to A^0 \to B \to 0,$$

where T is the "toric part" of A^0, i.e., that torus whose character group is L_p (cf. [**41**, p. 446]). Since ρ is not modular of level qM, we have $B(\overline{\mathbf{F}}_p)[\mathfrak{m}] = 0$. (Compare [**41**, Theorem 3.11], and the discussion in §8.) Therefore, $T(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ has $\mathbf{T}/\mathfrak{m}$-dimension 2λ. Consequently, $L_p/\mathfrak{m}L_p$ is of dimension 2λ. □

(7.7) Proposition. *We have*

$$\dim Y_q/\mathfrak{m}Y_q \le \mu, \qquad \dim L_q/\mathfrak{m}L_q \le \lambda.$$

This proposition is proved, analogously, by considering the action of a decomposition group D_q in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for the prime q. Pick such a group, together with a Frobenius element Frob_q in it. It will be useful to view D_q as $\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)$.

By the choice of q, ρ is unramified at q, and $\rho(\mathrm{Frob}_q)$ has the distinct eigenvalues $+1$ and -1. Therefore, on $J[\mathfrak{m}]$, say, Frob_q has the eigenvalues $+1$ and -1, each with multiplicity μ, in view of the fact that $J[\mathfrak{m}]$ is a direct sum of μ copies of ρ.

By **[18,** §5.1] (or by the results of **[51]**), the $\mathbf{T}[D_q]$-module $J[\mathfrak{m}]$ contains the submodule $\mathrm{Hom}(Y_q/\mathfrak{m}Y_q, \mu_\ell)$. (Here, μ_ℓ denotes the group of ℓ^{th} roots of unity of $\overline{\mathbf{Q}}_q$; the author expects that there will be no confusion with the multiplicity μ.)

We claim that the Frobenius element Frob_q acts on Y_q as the qth Hecke operator $U = T_q$. This claim implies that Frob_q acts on $\mathrm{Hom}(Y_q/\mathfrak{m}Y_q, \mu_\ell)$ as the product qT_q and, therefore, in particular, as an element of the field $\mathbf{T}/\mathfrak{m}$. This scalar will be ± 1, since U is an involution and since q is congruent to $-1 \pmod{\mathfrak{m}}$. It follows from this that the dimension of $Y_q/\mathfrak{m}Y_q$ is bounded from above by the multiplicity μ, since each eigenvalue ± 1 occurs at most μ times on $J[\mathfrak{m}]$. In other words, we obtain the desired inequality $\dim Y_q/\mathfrak{m}Y_q \le \mu$.

To prove the claim, we view Y_q as the first integral homology group of the dual graph associated with the reduced Shimura curve $C_{/\mathbf{F}_q}$. This graph may be described explicitly in terms of certain abelian varieties in characteristic q with "quaternionic multiplication" by an order $\mathcal{O}$ in a rational quaternion algebra **[40,** §5]. The abelian varieties A that occur in the description of the graph are all "exceptional" in the language of **[40,** §4]. This means that the kernel of the Frobenius morphism $A \to A^{(q)}$ coincides with the kernel of multiplication by $\mathfrak{q}$ on A, where $\mathfrak{q}$ is the unique two-sided maximal ideal of $\mathcal{O}$ whose residue field has q^2 elements. Since the operator T_q has the modular description $A \mapsto A/A[\mathfrak{q}]$, we may deduce that the actions of T_q and of the Frobenius automorphism coincide on the graph.

The inequality $\dim L_q/\mathfrak{m}L_q \le \lambda$ is obtained similarly. □

(7.8) PROPOSITION. *We have* $\dim X_p/\mathfrak{m}X_p \le \mu$.

Let J' again denote the abelian variety dual to J, and let Y'_q again be the analogue of Y_q for J'. Because J has multiplicative reduction at q, one has an exact sequence of $\mathbf{T}[\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)]$-modules

$$0 \to \mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell) \to J[\ell] \to Y'_q/\ell Y'_q \to 0. \tag{7.9}$$

Indeed, Grothendieck has constructed a "canonical filtration" **[18,** 11.6.5] on the $\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)$-module $A[\ell]$, whenever A is an abelian variety with semistable reduction over $\mathbf{Q}_q$. This filtration has the form

$$0 \subset A[\ell]^{\mathrm{t}} \subset A[\ell]^{\mathrm{f}} \subset A[\ell],$$

where $A[\ell]^{\mathrm{t}}$, the "toric part" of $A[\ell]$, may be written $\mathrm{Hom}(Y/\ell Y, \mu_\ell)$; here, Y is the character group of the toric part of the special fiber of the Néron model of A. Analogously (and dually), the quotient $A[\ell]/A[\ell]^{\mathrm{f}}$ of Galois modules may be identified with $Y'/\ell Y'$, where Y' is the analogue of Y for the abelian variety dual to A **[18,** 11.6.6, **18,** 11.6.7]. In case A has purely multiplicative reduction, Y and Y' are free of rank $\dim A$ and the middle terms $A[\ell]^{\mathrm{t}}$ and $A[\ell]^{\mathrm{f}}$ of the filtration coincide. One therefore obtains an exact sequence like (7.9) whenever the reduction of A is multiplicative. (For a discussion of abelian varieties with *split* multiplicative reduction, the reader may consult the expository account in Chapter III of **[36]**. The exact sequence (7.9) appears as Lemma 3.3.1 in **[36]**.)

Localization of (7.9) at $\mathfrak{m}$ yields an exact sequence of $\mathbf{T}[\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)]$-modules

$$(7.10) \qquad 0 \to \mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)_{\mathfrak{m}} \to J[\ell]_{\mathfrak{m}} \to (Y'_q/\ell Y'_q)_{\mathfrak{m}} \to 0.$$

This sequence *splits* as a sequence of $\mathbf{T}$-modules. Indeed, consider again a Frobenius element Frob_q of $\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)$. This operator commutes with the action of $\mathbf{T}$ and operates on each of $\mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)_{\mathfrak{m}}$ and $(Y'_q/\ell Y'_q)_{\mathfrak{m}}$ as an element of $\mathbf{T}$. Indeed, Frob_q operates on Y_q as the involution $U = T_q$; therefore, it operates as qU on $\mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)_{\mathfrak{m}}$ and as U on $(Y'_q/\ell Y'_q)_{\mathfrak{m}}$. These elements of $\mathbf{T}$ are incongruent mod $\mathfrak{m}$, since q is -1 (mod ℓ). This forces the desired splitting.

Since (7.10) splits, it remains exact after we take "kernels of $\mathfrak{m}$." In other words, we have a sequence of $(\mathbf{T}/\mathfrak{m})[\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)]$-modules

$$0 \to \mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)[\mathfrak{m}] \to J[\mathfrak{m}] \to (Y'_q/\ell Y'_q)[\mathfrak{m}] \to 0.$$

Since Frob_q operates on $\mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)[\mathfrak{m}]$ and $(Y'_q/\ell Y'_q)[\mathfrak{m}]$ as homotheties, the dimensions of each of these modules must be μ. To see this, we observe that the eigenvalues of Frob_q on the two-dimensional representation V are $+1$ and -1. Therefore, the characteristic polynomial arising from the action of Frob_q on $J[\mathfrak{m}]$ is $(T-1)^\mu(T+1)^\mu$. Suppose that $\mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)[\mathfrak{m}]$ has dimension d_1 and that Frob_q operates on $\mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)[\mathfrak{m}]$ as the scalar ξ_1. Define d_2 and ξ_2 similarly, using $(Y'_q/\ell Y'_q)[\mathfrak{m}]$ in place of $\mathrm{Hom}(Y_q/\ell Y_q, \mu_\ell)[\mathfrak{m}]$. Then we clearly have

$$(T-1)^\mu(T+1)^\mu = (T-\xi_1)^{d_1}(T-\xi_2)^{d_2}.$$

Since -1 and $+1$ are distinct, we have $d_1 = d_2 = \mu$.

Let Ψ_q again be the group of components associated with the mod q reduction of J. As was recalled in §6, Ψ_q may be constructed from the monodromy pairing on Y_q. Indeed, if we view this pairing as an inclusion $Y'_q \hookrightarrow Y^*_q$, where $Y^*_q = \mathrm{Hom}(Y_q, \mathbf{Z})$, then Ψ_q is naturally the cokernel of this inclusion. By considering the "multiplication by ℓ" maps in the exact sequence $0 \to Y'_q \to Y^*_q \to \Psi_q \to 0$, and using the Snake Lemma, we obtain an inclusion $\Psi_q[\ell] \subseteq Y'_q/\ell Y'_q$. In particular, we have $\Psi_q[\mathfrak{m}] \subseteq (Y'_q/\ell Y'_q)[\mathfrak{m}]$. Thus, μ is bounded from below by the dimension of $\Psi_q[\mathfrak{m}]$. By (6.9), this dimension coincides with that of $X_p/\mathfrak{m}X_p$, since η lies in $\mathfrak{m}$. Thus we have the desired inequality $\dim X_p/\mathfrak{m}X_p \le \mu$. □

Using (7.4), Proposition 7.6, and Proposition 7.7, we obtain $2\mu \le \lambda$, which implies $\mu \le \lambda$, since μ is a natural number. From (7.5) and Propositions 7.6, 7.7, and 7.8, we get $2\lambda \le 2\mu$. Clearly, these inequalities imply

$\lambda = \mu = 0$. As indicated above, this contradiction completes the proof of Theorem 1.5.

8. Mazur's Principle

In this section, ℓ is an odd prime and $\mathbf{F}$ is again a fixed algebraic closure of $\mathbf{F}_\ell$. Consider an odd irreducible representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ and a prime number q different from ℓ. (The case $q = \ell$ may be included if one introduces the concept of "finiteness" of a representation; see [**49**, p. 189] and [**41**, Lemma 6.2].)

The following result is a variant of (4.7).

(8.1) THEOREM. *Suppose that ρ arises from the space of weight-two cusp forms on $\Gamma_1(N) \cap \Gamma_0(q)$, where N is prime to q. Assume that ρ is unramified at q and that $q \not\equiv 1 \bmod \ell$. Then ρ arises from the space of weight-two cusp forms on $\Gamma_1(N)$.*

The space $\mathscr{S}$ of weight-two cusp forms on $\Gamma_1(N) \cap \Gamma_0(q)$ contains (as its "q-old" subspace) a direct sum $\mathscr{S}_0$ of two copies of $S_2(\Gamma_1(N))$. This subspace is stable under the action of the Hecke operators T_n (for $n \geq 1$) on $\mathscr{S}$. Let $\mathbf{T}$ be the subring of $\mathrm{End}\,\mathscr{S}$ generated by the T_n, and let $\mathbf{T}_0$ be the q-old quotient of $\mathbf{T}$, i.e., the image of $\mathbf{T}$ in $\mathrm{End}\,\mathscr{S}_0$. These algebras contain the diamond bracket operators $\langle d \rangle$ for $d \in (\mathbf{Z}/N\mathbf{Z})^*$. The assumption that ρ arises from $\mathscr{S}$ implies that there is a homomorphism $\varphi\colon \mathbf{T} \to \mathbf{F}$ satisfying

$$\varphi(T_r) = \mathrm{trace}(\rho(\mathrm{Frob}_r)), \qquad \varphi(\langle r \rangle r) = \det(\rho(\mathrm{Frob}_r))$$

for all prime numbers r prime to $\ell q N$. We view φ as an embedding $k \hookrightarrow \mathbf{F}$, where $k := \mathbf{T}/\mathfrak{m}$, and observe that there exists a model for ρ over the field k. This model is a two-dimensional k-vector space V furnished with an action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. The $\mathbf{F}$-vector space $V \otimes_k \mathbf{F}$ affords the representation ρ. As in §5, to prove that ρ arises from $\mathscr{S}_0$, it suffices to show that $\mathfrak{m}$ arises by pull-back from the quotient $\mathbf{T}_0$ of $\mathbf{T}$. We will sketch a proof that this is so; a more detailed proof, but with $\Gamma_1(N)$ replaced by $\Gamma_0(N)$, is given in [**41,** §6].

In the discussion below, we view V as a $\mathbf{T}$-module in the obvious way: $\mathbf{T}$ acts through its quotient k.

Consider the modular curve C attached to the group $\Gamma_1(N) \cap \Gamma_0(q)$. The operators T_n and $\langle d \rangle$ act on C as correspondences. The correspondences T_n and $\langle d \rangle$ of C induce a faithful action of $\mathbf{T}$ on the Jacobian $J := \mathrm{Pic}^0 C$. The endomorphisms of J induced by the T_n and $\langle d \rangle$ will be denoted, as usual, by the same symbols T_n and $\langle d \rangle$. The kernel $J[\mathfrak{m}]$ of $\mathfrak{m}$ on $J(\overline{\mathbf{Q}})$, which is nonzero, is a direct sum of copies of V (cf. [**4**]). In particular, there exist $\mathbf{T}[\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-equivariant maps $V \overset{\tilde{\imath}}{\hookrightarrow} J(\overline{\mathbf{Q}})$. Because V is unramified at q, and in view of the result of Serre-Tate used above, each choice $\tilde{\imath}$

determines a map of $\mathbf{T}[\mathrm{Gal}(\overline{\mathbf{Q}}_q/\mathbf{Q}_q)]$-modules

$$V \stackrel{\imath}{\hookrightarrow} \underline{J}(\overline{\mathbf{F}}_q).$$

Here, $\underline{J}$ represents the fiber over $\mathbf{F}_q$ of the Néron model of J. Fix one $\tilde{\imath}$, and use the resulting map $\imath$ to view V as a subgroup of $\underline{J}(\overline{\mathbf{F}}_q)$.

By results of Deligne-Rapoport **[11]** and Raynaud, $\underline{J}$ is semiabelian. In other words, $\underline{J}$ is an extension of its component group Φ by a connected group $\underline{J}^0$, which in turn is an extension of an abelian variety A by a torus T. More precisely, A is a product of two copies of $J_1(N)_{/\mathbf{F}_q}$, while the character group X of T may be interpreted as the group of degree-zero divisors on $X_1(N)_{/\mathbf{F}_q}$ which are concentrated at the set of supersingular points of $X_1(N)_{/\overline{\mathbf{F}}_q}$. The ring $\mathbf{T}$ acts naturally on Φ, T, and A by functoriality. To show that $\mathfrak{m}$ arises by pull-back from $\mathbf{T}_0$, it suffices to verify the following three points.

The ring $\mathbf{T}$ acts on A through its quotient $\mathbf{T}_0$.

The $\mathbf{T}$-module Φ is prime to $\mathfrak{m}$ in the sense that $\mathfrak{m}$ does not lie in the support of Φ.

The submodule V of $\underline{J}(\overline{\mathbf{Q}}_q)$ cannot be contained in the torus T.

For the first point, the standard degeneracy maps $\alpha, \beta \colon C \rightrightarrows X_1(N)$ induce (in characteristic zero) a map on Picard varieties $\delta \colon J_1(N) \times J_1(N) \to J$ which is known to be injective (cf. **[13]** or **[38]**). It is clear that $\mathbf{T}$ acts on the abelian subvariety $J_1(N) \times J_1(N)$ of J through its old quotient $\mathbf{T}_0$. Indeed, if we dualize δ and then pass to cotangent spaces, we obtain the canonical inclusion $\mathscr{S}_0 \hookrightarrow \mathscr{S}$. The map δ induces in characteristic q a map $\delta_{/\mathbf{F}_q} \colon J_1(N)_{/\mathbf{F}_q} \times J_1(N)_{/\mathbf{F}_q} \to \underline{J}^0$, and the point to be checked is that $\pi \circ \delta_{/\mathbf{F}_q}$ is an isogeny, where π is the structural map $\underline{J}^0 \to A$. If we regard A as $J_1(N)_{/\mathbf{F}_q} \times J_1(N)_{/\mathbf{F}_q}$, then $\pi \circ \delta_{/\mathbf{F}_q}$ becomes a two-by-two matrix of endomorphisms of $J_1(N)_{/\mathbf{F}_q}$. This matrix, as computed by Deligne-Rapoport **[11,** p. 287], has diagonal components equal to the identity endomorphism and off-diagonal components equal to the Verschiebung endomorphism (dual of the Frobenius). This matrix is then visibly an isogeny, since it acts on the cotangent space of $J_1(N)_{/\mathbf{F}_q} \times J_1(N)_{/\mathbf{F}_q}$ as the identity map.

The following circumstance is perhaps worth stressing. In the functorial action of $\mathbf{T}$ on A, the Hecke operators $\langle d \rangle$ and T_n with n prime to q act "as expected." Namely, they are induced by their namesakes on $J_1(N)$, which operate diagonally on $J_1(N) \times J_1(N)$, and then by reduction (mod q) on A. However, the functorial action of T_q on A belongs naturally to characteristic q.

For the second point, we show that Φ is *Eisenstein* in the strong sense that we have on Φ the equations $T_p = 1 + p$ and $\langle p \rangle = 1$ for all p prime to qN. These equations imply immediately that $\mathfrak{m}$ is prime to the support of Φ because ρ is irreducible; cf. **[41,** Theorem 5.2c]. That Φ is Eisenstein was proved in **[39]** for the abelian variety $J_0(qN)$. The result we seek then follows for $N \leq 3$ because the natural map $J_0(qN) \to J$ is an isomorphism.

Assume that $N \geq 4$. Recall that Φ is the cokernel of the map $X \to \mathrm{Hom}(X, \mathbf{Z})$ coming from the monodromy pairing on X. View X as the kernel of the degree map $X^{\sim} \to \mathbf{Z}$, where $X^{\sim}$ is the free abelian group on the set Σ of supersingular points of $X_1(N)_{/\overline{\mathbf{F}}_q}$. The monodromy pairing on X is the restriction to X of a diagonal pairing on $X^{\sim}$. This pairing takes the value $e(\sigma)$ on (σ, σ), where $e(\sigma)$ is essentially the number of automorphisms of any pair (E, P) that represents σ; we understand that E is a supersingular elliptic curve over $\overline{\mathbf{F}}_q$ and $P \in E(\overline{\mathbf{F}}_q)$ is a point of order N. The precise formula for (σ, σ) may be derived from **[11,** p. 286, Théorème 6.9]; we have

$$e(\sigma) = \begin{cases} \frac{1}{2}\#\mathrm{Aut}(E, P) & \text{if } -1 \in \mathrm{Aut}\,\sigma, \\ \#\mathrm{Aut}(E, P) & \text{if } -1 \notin \mathrm{Aut}\,\sigma. \end{cases}$$

The assumption $N \geq 4$ implies that $e(\sigma) = 1$ by **[27,** Corollary 2.7.4]. Hence the monodromy pairing on X is just the restriction to $X \times X$ of the standard diagonal pairing on $X^{\sim}$. It follows that the natural map

$$\xi\colon X^{\sim}/X \longrightarrow \Phi = \mathrm{Hom}(X, \mathbf{Z})/X, \qquad \sum n_\sigma \sigma \longmapsto \left\langle \sum n_\sigma \sigma, \cdot \right\rangle$$

is surjective. Since $X^{\sim}/X$ is mapped isomorphically to $\mathbf{Z}$ by the degree map, Φ is a finite cyclic group.

Fix p prime to Nq, and let T denote the usual pth Hecke operator on $X^{\sim}$: this is the $\mathbf{Z}$-linear map of $X^{\sim}$ that sends (the class of) (E, P) to $\sum(E/\mathscr{C}, P \bmod \mathscr{C})$. Here, the sum ranges over the $p+1$ subgroups $\mathscr{C}$ of $E[p]$ whose order is p. Let T' denote the composite $T \circ \langle p \rangle^{-1}$, where $\langle p \rangle\colon (E, p) \mapsto (E, pP)$. One checks the formula $\langle Tx, y \rangle = \langle x, T'y \rangle$, where $\langle\, , \rangle$ is the Euclidean pairing on $X^{\sim}$. Also, it is clear that both T and T' preserve the subgroup X of $X^{\sim}$ and induce the map "multiplication by $p+1$" on the quotient $X^{\sim}/X$.

The operator T_p on Φ is the map on $\mathrm{Hom}(X, \mathbf{Z})/X$ induced by the operators T on X and $\mathrm{Hom}(T', \mathbf{Z})$ on $\mathrm{Hom}(X, \mathbf{Z})$. The formula relating T, T', and $\langle\, , \rangle$ shows that ξ is T_p-equivariant if T_p acts naturally on $X^{\sim}/X$ by multiplication by $p+1$. Hence $T_p = p+1$ on Φ. In a similar manner, we find that the diamond-bracket operators are trivial on Φ.

For the last point, we must determine the action of Frob_q on $T(\overline{\mathbf{F}}_q)$. Let w be the automorphism of C that corresponds to the map

$$(E, P, \mathscr{C}) \mapsto (E/\mathscr{C}, P \bmod \mathscr{C}, E[q]/\mathscr{C})$$

on triples consisting of an elliptic curve E, a point P on E of order M, and a cyclic subgroup $\mathscr{C}$ of E that has order q. A short computation shows that we have $w^2 = \langle q \rangle$. One checks as well that the correspondence $T_q + w_q$ of C may be written $\beta' \circ \alpha$ where α and β are again the degeneracy maps

$C \rightrightarrows X_1(N)$. It follows from this that $T_q + w_q$ induces an endomorphism of $\underline{J}$ that kills the torus T, since this endomorphism factors through a map $\underline{J} \to J_1(N)$. On the other hand, w coincides with the Frobenius map on Σ but permutes the two components of $C_{/\mathbf{F}_q}$. Therefore, $\mathrm{Frob}_q = -w = T_q$ on X, since X is canonically the first integral homology group of the dual graph associated with $C_{/\mathbf{F}_q}$. Thus $\mathrm{Frob}_q = qT_q$ on $T(\overline{\mathbf{F}}_q)$. Therefore, if $V \hookrightarrow T(\overline{\mathbf{F}}_q)$, then Frob_q acts on V as a scalar, i.e., as an element of k. Since the square of this scalar is $q^2\langle q\rangle$, the determinant of the action of Frob_q on V is $q^2\langle q\rangle$.

On the other hand, it is clear from the Čebotarev Density Theory that the determinant of V is the character $\chi\varepsilon\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to k^*$, where ε is the Dirichlet character $\langle\ \rangle$, i.e., where $\varepsilon(\mathrm{Frob}_r) = \langle r\rangle$ for every prime number r prime to M. In particular, the determinant of the action of Frob_q on V is $q\langle q\rangle$. Therefore, we cannot have $V \hookrightarrow T(\overline{\mathbf{F}}_q)$ unless $q \equiv 1 \mod \ell$.

References

1. A. Atkin and W. Li, *Twists of newforms and pseudo-eigenvalues of W-operators*, Invent. Math. **48** (1978), 221–243.
2. A. Ash and G. Stevens, *Modular forms in characteristic ℓ and special values of their L-functions*, Duke Math. J. **53** (1986), 849–868.
3. L. Bégueri, *Dualité sur un corps local à corps résiduel algébriquement clos*, Mém. Soc. Math. France, (N. S.) **4** (1980).
4. N. Boston, *Families of Galois representations—increasing the ramification*, Duke Math. J. **66** (1992), 357–367.
5. N. Boston, H.W. Lenstra, Jr., and K.A. Ribet, *Quotients of group rings arising from two-dimensional representations*, C. R. Acad. Sci. Paris Sér. I Math. **312** (1991), 323–328.
6. H. Carayol, *Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. École Norm. Sup. $(4)^e$ **19** (1986), 409–468.
7. ——, *Sur les représentations galoisiennes modulo ℓ attachées aux formes modulaires*, Duke Math. J. **59** (1989), 785–801.
8. I.V. Cerednik, *Uniformization of algebraic curves by discrete arithmetic subgroups of* $\mathrm{PGL}_2(k_w)$ *with compact quotients*, Mat. Sb. **100** (1976), 59–88; English transl. in Math USSR-Sb. **29** (1976).
9. R. Coleman and J. Voloch, *Companion forms and Kodaira-Spencer theory*, Invent. Math. **110** (1992), 263–281.
10. P. Deligne, *Formes modulaires et représentations ℓ-adiques*, Sém. Bourbaki **355** (1968/69), Lecture Notes in Math., vol. 179, Springer-Verlag, Berlin and New York, 1971, pp. 139–172.
11. P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973, pp. 143–316.
12. P. Deligne and J.-P. Serre, *Formes modulaires de poids* 1, Ann. Sci. École Norm. Sup. (4^e) **7** (1974), 507–530.
13. F. Diamond, *Congruence primes for cusp forms of weight* $k \geq 2$, Astérisque **196–197** (1991), 205–213.
14. V.G. Drinfeld, *Coverings of p-adic symmetric regions*, Functsional. Anal. i Prilozhen. **10** (1976), 29–40; English transl. in Functional Anal. Appl. **10** (1976).
15. B. Edixhoven, *The weight in Serre's conjectures on modular forms*, Invent. Math. **109** (1992), 563–594.
16. G. Faltings and B. Jordan, *Crystalline cohomology and* $\mathbf{GL}(2, \mathbf{Q})$ (to appear).
17. B.H. Gross, *A tameness criterion for Galois representations associated to modular forms* mod p, Duke Math. J. **61** (1990), 445–517.

18. A. Grothendieck, *Séminaire de Géometrie Algébrique* 7, *Exposé* IX, Lecture Notes in Math., vol. 288, Springer-Verlag, Berlin and New York, 1972, pp. 313–523.
19. H. Hida, *Congruences of cusp forms and special values of their zeta function*, Invent. Math. **63** (1981), 225–261.
20. ——, *Iwasawa modules attached to congruences of cusp forms*, Ann. Sci. École Norm. Sup. (4^e) **19** (1986), 231–273.
21. ——, *Galois representations into* $\mathbf{GL}_2(\mathbf{Z}_p[[X]])$ *attached to ordinary cusp forms*, Invent. Math. **85** (1986), 545–613.
22. K. Iwasawa, *Lectures on p-adic L-functions*, Ann. of Math. Stud. **74**, Princeton Univ. Press, Princeton, NJ, 1972.
23. B. Jordan and R. Livné, *On the Néron model of Jacobians of Shimura curves*, Compositio Math. **60** (1986), 227–236.
24. ——, *Conjecture "epsilon" for weight* $k > 2$, Bull. Amer. Math. Soc. (N. S.) **21** (1989), 51–56.
25. N. M. Katz, *Higher congruences between modular forms*, Ann. of Math (2) **101** (1975), 332–367.
26. ——, *A result on modular forms in characteristic* p, Lecture Notes in Math., vol. 601, Springer-Verlag, Berlin and New York, 1977, pp. 53–61.
27. N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Ann. of Math. Stud., vol. 108, Princeton Univ. Press, Princeton, NJ, 1985.
28. R. P. Langlands, *Modular forms and* ℓ*-adic representations*, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973, pp. 361–500.
29. W. Li, *Newforms and functional equations*, Math. Ann. **212** (1975), 285–315.
30. R. Livné, *On the conductors of mod* ℓ *Galois representations coming from modular forms*, J. Number Theory **31** (1989), 133–141.
31. B. Mazur, *Letter to J.-F. Mestre* (16 *August* 1985), unpublished.
32. B. Mazur and K. A. Ribet, *Two-dimensional representations in the arithmetic of modular curves*, Astérisque **196–197** (1991), 215–255.
33. W. G. McCallum, *Duality theorems for Néron models*, Duke Math. J. **53** (1986), 1093–1124.
34. T. Miyake, *Modular forms*, Springer-Verlag, Berlin and New York, 1989.
35. C. Queen, *The existence of p-adic abelian L-functions*, Number Theory and Algebra (H. Zassenhaus, ed.), Academic Press, New York, 1977.
36. K. A. Ribet, *Galois action on division points on abelian varieties with many real multiplications*, Amer. J. Math. **98** (1976), 751–804.
37. ——, *The* ℓ*-adic representations attached to an eigenform with Nebentypus: A survey*, Lecture Notes in Math., vol. 601, Springer-Verlag, Berlin and New York, 1977, pp. 17–52.
38. ——, *Congruence relations between modular forms*, Proc. Internat. Congr. of Math., Warsaw, 1983, pp. 503–514.
39. ——, *On the component groups and the Shimura subgroup of* $J_0(N)$, exposé 6, Sém. Th. Nombres, Université Bordeaux, 1987–1988.
40. ——, *Bimodules and Abelian surfaces*, Adv. Stud. Pure Math., vol. 17, Academic Press, Boston, MA, 1989, pp. 359–407.
41. ——, *On modular representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), 431–476.
42. ——, *Raising the levels of modular representations*, Prog. Math., vol. 81, Birkhäuser, Boston, MA, 1990, pp. 259–271.
43. ——, *Lowering the levels of modular representations without multiplicity one*, Internat. Math. Res. Notices, 1991, pp. 15–19.
44. J.-P. Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), 259–331.
45. ——, *Congruences et formes modulaires* [*d'après H. P. F. Swinnerton-Dyer*], Sém. Bourbaki **416** (1971/72), Lecture Notes in Math., vol. 317, Springer-Verlag, Berlin and New York, 1973, pp. 319–338.
46. ——, *Formes modulaires et fonctions zêta p-adiques*, Lecture Notes in Math., vol. 350, Springer-Verlag, Berlin and New York, 1973, pp. 191–268.
47. ——, *Divisibilité de certaines fonctions arithmétiques*, Enseign. Math. (2) **22** (1976), 227–260.

48. ——, *Lettre à J.-F. Mestre* (13 *août* 1985), Contemp. Math., vol. 67, Amer. Math. Soc., Providence, RI, 1987, pp. 263–268.
49. ——, *Sur les représentations modulaires de degré 2 de* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. **54** (1987), 179–230.
50. ——, *Letter to K. Ribet* (15 *April* 1987), unpublished.
51. J.-P. Serre and J. Tate, *Good reduction of abelian varieties*, Ann. of Math. (2) **88** (1968), 492–517.
52. G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton Univ. Press, Princeton, NJ, 1971.
53. H. P. F. Swinnerton-Dyer, *On* ℓ*-adic representations and congruences for modular forms*, Lecture Notes in Math., vol. 350, Springer-Verlag, Berlin and New York, 1973, pp. 1–55.
54. John T. Tate, *Number theoretic background*, Proc. Sympos. Pure Math., vol. 33, Amer. Math. Soc., Providence, RI, 1979, pp. 3–26.

University of California, Berkeley
E-mail address: `ribet@math.berkeley.edu`

Recent Titles in This Series

(Continued from the front of this publication)

31 **Joseph L. Doob, editor,** Probability (University of Illinois at Urbana-Champaign, Urbana, March 1976)

30 **R. O. Wells, Jr., editor,** Several complex variables (Williams College, Williamstown, Massachusetts, July/August 1975)

29 **Robin Hartshorne, editor,** Algebraic geometry – Arcata 1974 (Humboldt State University, Arcata, California, July/August 1974)

28 **Felix E. Browder, editor,** Mathematical developments arising from Hilbert problems (Northern Illinois University, Dekalb, May 1974)

27 **S. S. Chern and R. Osserman, editors,** Differential geometry (Stanford University, Stanford, California, July/August 1973)

26 **Calvin C. Moore, editor,** Harmonic analysis on homogeneous spaces (Williams College, Williamstown, Massachusetts, July/August 1972)

25 **Leon Henkin, John Addison, C. C. Chang, William Craig, Dana Scott, and Robert Vaught, editors,** Proceedings of the Tarski symposium (University of California, Berkeley, June 1971)

24 **Harold G. Diamond, editor,** Analytic number theory (St. Louis University, St. Louis, Missouri, March 1972)

23 **D. C. Spencer, editor,** Partial differential equations (University of California, Berkeley, August 1971)

22 **Arunas Liulevicius, editor,** Algebraic topology (University of Wisconsin, Madison, June/July 1970)

21 **Irving Reiner, editor,** Representation theory of finite groups and related topics (University of Wisconsin, Madison, April 1970)

20 **Donald J. Lewis, editor,** 1969 Number theory institute (State University of New York at Stony Brook, Stony Brook, July 1969)

19 **Theodore S. Motzkin, editor,** Combinatorics (University of California, Los Angeles, March 1968)

18 **Felix Browder, editor,** Nonlinear operators and nonlinear equations of evolution in Banach spaces (Chicago, April 1968)

17 **Alex Heller, editor,** Applications of categorical algebra (New York City, April 1968)

16 **Shiing-Shen Chern and Stephen Smale, editors,** Global analysis, Part III (University of California, Berkeley, July 1968)

15 **Shiing-Shen Chern and Stephen Smale, editors,** Global analysis, Part II (University of California, Berkeley, July 1968)

14 **Shiing-Shen Chern and Stephen Smale, editors,** Global analysis, Part I (University of California, Berkeley, July 1968)

13 **Dana S. Scott (Part 1) and Thomas J. Jech (Part 2), editors,** Axiomatic set theory (University of California, Los Angeles, July/August 1967)

12 **William J. LeVeque and Ernst G. Straus, editors,** Number theory (Houston, Texas, January 1967)

11 **S. S. Chern, L. Ehrenpreis, J. Korevaar, W. H. J. Fuchs, and L. A. Rubel, editors,** Entire functions and related parts of analysis (University of California, San Diego, July 1966)

10 **Alberto P. Calderón, editor,** Singular integrals (University of Chicago, April 1966)

9 **Armand Borel and George D. Mostow, editors,** Algebraic groups and discontinuous subgroups (University of Colorado, Boulder, July 1965)

8 **Albert Leon Whiteman, editor,** Theory of numbers (California Institute of Technology, Pasadena, November 1963

(See the AMS catalog for earlier titles)

ISBN 0-8218-1637-3

9 780821 816370