

OVERCONVERGENT MODULAR FORMS

AND THEIR EXPLICIT ARITHMETIC

JAN VONK

ABSTRACT. In these notes, we aim to give a friendly introduction to the theory of overconvergent modular forms and some examples of recent arithmetic applications. The emphasis is on explicit examples and computations.

INTRODUCTION

The theory of p -adic modular forms has its origins in the work of Serre and Katz in the 1970's, and has seen a spectacular amount of development and applications in number theory since then. In this note, we aim to provide its context and sketch the rudiments of the theory, adopting an approach where we favour explicit examples and computations over proofs. This is done with the hope that the uninitiated reader may build up some intuition and working knowledge as a stepping stone to the literature on the subject, which can be somewhat daunting to outsiders, but for which there is no substitute if one wants to become a serious user. We have included references to many of the original texts.

We should warn the reader that by its very design, this article is doomed to be incomplete, and several crucial developments are not discussed in this text. One notable example is that we have omitted a discussion of the theory of overconvergent modular symbols, which often provides an alternative framework that is in its own way highly suited for explicit computation. The author wishes to apologise for this omission, and many others, with the added clarification that this is merely a reflection of his own lack of experience with this approach. Likewise, many recent exciting developments in the area, such as the burgeoning topic of higher Hida theory, will unfortunately not be discussed in any detail here. Finally, it is important to note that a great many excellent (expository) sources on the theory of overconvergent modular forms already exist, which include for instance the beautiful treatments of Emerton [Eme11] and Calegari [Cal13].

This article is a reworked version of the notes for a mini-course taught by the author at a summer school on Iwasawa theory, held in Bordeaux 19–22 June 2019. We are grateful to the organisers of this conference for the invitation to give these lectures, and to Mark Goresky for the invitation to publish them in the Bulletin of the AMS. The author is supported by NSF Grant No. DMS-1638352. All computations were performed using the SAGE [S⁺20] and MAGMA [BCP97] computer algebra systems.

CONTENTS

| | |
|---|---|
| Introduction | 1 |
| 1. Congruences between modular forms | 2 |
| 2. Eisenstein series and Kummer congruences | 6 |

| | |
|--|----|
| 3. Overconvergent modular forms | 11 |
| 4. Explicit computations and arithmetic applications | 24 |
| References | 37 |

1. CONGRUENCES BETWEEN MODULAR FORMS

We start by recalling some basic definitions, and motivate the theme of these notes by discussing some classical congruences for the Ramanujan Δ -function, a weight 12 modular form of level 1. These illustrate different types phenomena, and we highlight the features we wish to explore in these notes.

1.1. Modular forms. Suppose $\Gamma \subseteq \mathrm{SL}_2(\mathbf{Z})$ is a finite index subgroup, then $M_k(\Gamma)$ denotes the space of *modular forms of weight $k \in \mathbf{Z}$* , that is the space of holomorphic functions f on the Poincaré upper half plane $\mathcal{H} := \{z \in \mathbf{C} \mid \mathrm{Im}(z) > 0\}$ which satisfy the transformation law

$$(1) \quad f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z), \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$$

and are holomorphic at the cusps of Γ . The subspace of cuspforms consists of those functions which vanish at all the cusps, and is denoted by $S_k(\Gamma)$. In these notes, Γ will usually be given by the congruence subgroup

$$(2) \quad \Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : c \equiv 0 \pmod{N} \right\}.$$

Any modular form $f \in M_k(\Gamma_0(N))$ is invariant under translation, and admits a Fourier expansion

$$(3) \quad f(q) = a_0 + a_1 q + a_2 q^2 + \dots, \quad q = e^{2\pi i z}.$$

This will be referred to as its *q-expansion*, and the $a_i \in \mathbf{C}$ are called its *Fourier coefficients*. We refer to a_0 as the *constant term* of f , and the a_i for $i \geq 1$ as its *higher Fourier coefficients*. When $a_1 = 1$, we say f is *normalised*.

Classical examples of modular forms are given by the *Eisenstein series*, which are constructed as follows. For any even $k \geq 4$, we have the weight k normalised Eisenstein series

$$(4) \quad \begin{aligned} \mathbf{G}_k(z) &= \sum_{(m,n) \in \mathbf{Z}^2 \setminus (0,0)} (mz+n)^{-k} \\ &= \frac{-B_k}{2k} + \sum_{n \geq 1} \sigma_{k-1}(n) q^n, \end{aligned}$$

where B_k is the k -th Bernoulli number (see equation (29)) and where $\sigma_r = \sum_{d|n} d^r$ is the *divisor function*. They define modular forms of weight k for the full modular group $\mathrm{SL}_2(\mathbf{Z})$. We note that for $k = 2$, the series defining \mathbf{G}_k fails to converge absolutely, and indeed we have $M_2(\mathrm{SL}_2(\mathbf{Z})) = 0$. We will see in § 2.5 that the q -expansion above still has meaning for $k = 2$ as a p -adic modular form.

The dimension of M_k and S_k may typically be calculated using Riemann–Roch, and the theory of modular symbols allows one to compute, to any desired q -adic accuracy, a set of q -expansions of a basis for it. For more details on these computations, see the detailed treatment of Stein [Ste07].

1.2. The Hecke algebra. Two central aspects of the theory of modular forms are the action of the Hecke algebra, and associated Galois representations, which we briefly discuss now.

The spaces of modular forms $M_k(\Gamma_0(N))$ and $S_k(\Gamma_0(N))$ are finite-dimensional, and they are equipped with an action of the *Hecke algebra*, generated by operators T_p for any prime p , where it is customary to use the notation U_p whenever $p \mid N$. In terms of q -expansions (3) they are given by the expressions

$$(5) \quad \begin{aligned} T_p f(q) &= \sum_{n \geq 0} a_{np} q^n + p^{k-1} \sum_{n \geq 0} a_n q^{pn} & p \nmid N, \\ U_p f(q) &= \sum_{n \geq 0} a_{np} q^n & p \mid N. \end{aligned}$$

Any modular form that is an eigenvector for all these Hecke operators is called an *eigenform*. The Eisenstein series \mathbf{G}_k defined in (4) is a simple example of an eigenform, which satisfies

$$(6) \quad T_p \mathbf{G}_k = (1 + p^{k-1}) \mathbf{G}_k.$$

Note that the eigenvalue $(1 + p^{k-1})$ is equal to the divisor function $\sigma_{k-1}(p)$, and therefore also the p -th Fourier coefficient of \mathbf{G}_k displayed in (4). One can verify from the expressions (5) that this is always the case; if f is a normalised eigenform, then its eigenvalue for T_p or U_p is equal to its p -th Fourier coefficient. For an introduction to the basic properties of Hecke operators, see Diamond–Shurman [DS05].

Many spectacular results in number theory revolve around the notion of *Galois representations*. In what follows, this will always mean a continuous representation

$$(7) \quad \rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(k)$$

where $G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is the absolute Galois group of \mathbf{Q} , and k is either the field of complex numbers \mathbf{C} (in which case ρ is called an *Artin representation*), or a p -adic field such as \mathbf{Q}_p . Important examples of the latter arise from elliptic curves. Suppose E is an elliptic curve defined over \mathbf{Q} , choose a prime p , and consider the *p-adic Tate module*, obtained from the inverse limit of the torsion points on E of p -power order

$$(8) \quad \mathbf{Q}_p \otimes_{\mathbf{Z}_p} \left(\varprojlim_n E[p^n] \right), \quad \text{where } E[p^n] = \mathrm{Ker}(E \xrightarrow{\times p^n} E),$$

which is a 2-dimensional \mathbf{Q}_p -vector space. This space has a natural action of the Galois group $G_{\mathbf{Q}}$, given by the Galois action on the coordinates of the p^n -torsion points on E , which are algebraic numbers. Many important arithmetic properties of the elliptic curve E can be recovered from this Galois representation. For instance, for any prime $\ell \neq p$ of good reduction for E we have that

$$(9) \quad \mathrm{Tr} \rho(\mathrm{Frob}_{\ell}) = \ell + 1 - |E(\mathbf{F}_{\ell})|.$$

In other words, the trace of the matrix of Frobenius at ℓ is related to the number of points of E over the finite field \mathbf{F}_{ℓ} . It is striking that this representation depends on the choice of a prime p , but the traces of Frobenius at primes $\ell \neq p$ of good reduction for E are integers, and are independent of the choice of p .

Suppose f is an eigenform of level N and weight $k = 2$, where N is the conductor of E . We say that f is *attached to* E if the traces of Frobenius elements (9) are equal to the Fourier coefficients a_{ℓ} of f . In other words, this means that for all but finitely many ℓ we have

$$(10) \quad a_{\ell} = \ell + 1 - |E(\mathbf{F}_{\ell})|.$$

Important developments, culminating in the work of Wiles [Wil95], Taylor–Wiles [TW95], and Breuil–Conrad–Diamond–Taylor [BCDT01] show that for any elliptic curve E over \mathbf{Q} , there exists a modular form

that is attached to it in this sense. This has led not just to a proof of Fermat's Last Theorem, but subsequent developments continue to this day to settle long-standing conjectures in number theory.

Remark. We briefly mention that the converse was known much earlier. That is, a construction of Eichler–Shimura attaches an elliptic curve E to any eigenform of weight $k = 2$ with integer Fourier coefficients. Generally, to any normalised cuspidal eigenform f of weight k one may attach a 2-dimensional Galois representation ρ of $G_{\mathbf{Q}}$, which is unramified at all primes ℓ away from a finite set, and satisfies

$$(11) \quad \det(1 - \rho(\text{Frob}_{\ell})T) = 1 - a_{\ell}T + \ell^{k-1}T^2$$

where a_{ℓ} is the ℓ -th Fourier coefficient of f . When $k \geq 2$ this representation is valued in a non-archimedean local field, and it was constructed by Deligne [Del71], though it is no longer attached to an elliptic curve as in the aforementioned construction for $k = 2$ due to Eichler–Shimura. When $k = 1$ it is an Artin representation, constructed by Deligne–Serre [DS74] from the representations in higher weight via congruences.

1.3. Some examples of congruences. The Ramanujan Δ -function is the unique normalised cusp form of weight 12 for the group $\Gamma = \text{SL}_2(\mathbf{Z})$. Its q -expansion is given by the infinite product due to Jacobi:

$$(12) \quad \Delta(q) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = \sum_{n=1}^{\infty} \tau(n)q^n$$

This explicit product allows us to easily establish a number of congruences between the Fourier coefficients of Δ and those of various other modular forms, going back to the early twentieth century. For the reader who would like to inspect these manually, we tabulate its first few Fourier coefficients $\tau(p)$ for p prime:

| | | | | | | | | |
|------|-----------|----------|----------|-----------|-----------|------------|-----------|-----------|
| (13) | p | 2 | 3 | 5 | 7 | 11 | 13 | 17 |
| | $\tau(p)$ | -24 | 252 | 4830 | -16744 | 534612 | -577738 | -6905934 |
| | p | 19 | 23 | 29 | 31 | 37 | 41 | 43 |
| | $\tau(p)$ | 10661420 | 18643272 | 128406630 | -52843168 | -182213314 | 308120442 | -17125708 |

Example 1. We begin with a congruence that appears in the work of Ramanujan [Ram16]. Consider the weight k Eisenstein series \mathbf{G}_k introduced in (4). For $k = 12$, its constant term is equal to $\frac{691}{65520}$, whereas for $k = 6$ the constant term is $\frac{-1}{504}$. Since the space $M_{12}(\text{SL}_2(\mathbf{Z}))$ is two-dimensional, spanned by \mathbf{G}_{12} and Δ , the form \mathbf{G}_6^2 must be a linear combination of the two. Computing the first two terms of all three q -expansions, we find that

$$(14) \quad \frac{691}{65520} \cdot 504^2 \cdot \mathbf{G}_6(q)^2 = \mathbf{G}_{12}(q) - \frac{756}{65} \Delta(q)$$

and since all three modular forms involved have 691-integral q -expansions, we obtain

$$(15) \quad \Delta(q) \equiv \mathbf{G}_{12}(q) \pmod{691}.$$

In particular, we see that for any prime p , we get the celebrated *Ramanujan congruences*

$$(16) \quad \tau(p) \equiv 1 + p^{11} \pmod{691}.$$

For a beautiful and very detailed expository discussion of this example in the broader context of ideal class groups of cyclotomic fields and Galois representations, see Mazur [Maz11].

Example 2. This example is due to Wilton [Wil30], and establishes a congruence modulo 23 between Δ and a certain form of weight 1. We have the following congruences¹ for Δ :

$$(17) \quad \begin{aligned} q \prod_{n=1}^{\infty} (1 - q^n)^{24} &\equiv q^{1/24} \prod_{n=1}^{\infty} (1 - q^n) \cdot q^{23/24} \prod_{n=1}^{\infty} (1 - q^{23n}) \pmod{23} \\ &\equiv \frac{1}{2} \sum_{u,v \in \mathbf{Z}} \left(q^{u^2+uv+6v^2} - q^{2u^2+uv+3v^2} \right) \pmod{23} \end{aligned}$$

where the first is a consequence of the fact that for any prime p , the binomial coefficient $\binom{p}{i}$ is divisible by p for all $0 < i < p$, and the second follows from a calculation using the *Euler identity*

$$(18) \quad \prod_{n=1}^{\infty} (1 - q^n) = \sum_{n \in \mathbf{Z}} (-1)^n q^{\frac{3n^2+n}{2}}.$$

It is a classical result, see for instance Hecke [Hec26], that right hand side of (17) is a modular form of weight one. It is in fact a Hecke eigenform, with an associated Artin representation that we can identify easily: The quadratic field $\mathbf{Q}(\sqrt{-23})$ has class number 3, and its Hilbert class field H is obtained by adjoining a root of the cubic polynomial

$$(19) \quad f(x) = x^3 - x - 1,$$

which has discriminant -23 . The natural quotient gives us

$$(20) \quad G_{\mathbf{Q}} \longrightarrow \text{Gal}(H/\mathbf{Q}) \simeq S_3 \longrightarrow \text{GL}_2(\mathbf{C})$$

from the unique 2-dimensional irreducible representation of S_3 . This is the 2-dimensional Artin representation attached to the above weight one form. In particular, this means that the congruence class of $\tau(p)$ modulo 23 may be worked out from the splitting behaviour of the prime p in the extension H/\mathbf{Q} . The reader may enjoy verifying in general, or simply checking on a few small values of ℓ in the table (13), that this boils down to the statement² that for any prime $p \neq 23$ we have

$$(21) \quad \begin{aligned} \tau(p) &\equiv 0 \pmod{23} && \text{if } (-23/p) = -1, \\ \tau(p) &\equiv 2 \pmod{23} && \text{if } (-23/p) = 1 \text{ and } p = u^2 + 23v^2, \\ \tau(p) &\equiv -1 \pmod{23} && \text{if } (-23/p) = 1 \text{ and } p \neq u^2 + 23v^2. \end{aligned}$$

Example 3. As in the previous example, an elementary divisibility of binomial coefficients allows us to obtain from the infinite product expansion the following congruence for $\Delta(q)$

$$(22) \quad \Delta(q) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} \equiv q \prod_{n=1}^{\infty} (1 - q^n)^2 (1 - q^{11n})^2 \pmod{11}$$

The right hand side is a weight 2 normalised cusp form of level $\Gamma_0(11)$. It is associated to the elliptic curve

$$(23) \quad E : y^2 + y = x^3 - x^2 - 10x - 20,$$

so that we obtain in particular the following congruences for $p \neq 11$:

$$(24) \quad \tau(p) \equiv p + 1 - |E(\mathbf{F}_p)| \pmod{11}.$$

The reader may enjoy verifying this for a few small primes, using the table (13) and the equation (23). Unfortunately, the law governing the association $p \mapsto p + 1 - |E(\mathbf{F}_p)|$ cannot be made explicit in the same elementary terms as we did in the previous two examples in (16) and (21). The reason for this was explained

¹The reader familiar with the Dedekind η -function – which is a modular form of weight $1/2$ for some character χ_{24} of the metaplectic double cover of $\text{SL}_2(\mathbf{Z})$ – will recognise the form on the right hand side as $\eta(q)\eta(q^{23})$.

²We use the notation (a/p) for the *Legendre symbol* for $p \nmid a$, which equals 1 if a is a square modulo p , and -1 otherwise.

by Shimura [Shi66], since this law is governed by the traces of the 11-adic representation attached to the elliptic curve (23), and Shimura showed that its mod 11 reduction has image $\mathrm{GL}_2(\mathbf{F}_{11})$. Since this group is not solvable, the law is equivalent to the splitting behaviour of primes p in a non-solvable extension of \mathbf{Q} , which is in contrast with Example 2, where the relevant group was S_3 .

1.4. The context of this article. The three examples of congruences (15), (17), (22) are of very different flavours, and illustrate different but related phenomena that arise in the p -adic theory of modular forms:

- The first is a congruence between a *cusp form* and an *Eisenstein series*, of the *same weight*. Such congruences are central in Iwasawa theory, and related to the notion of the *Eisenstein ideal*, see Mazur [Maz77]. We will not discuss this theme, but mention that this is a fascinating topic that remains today an active area of research, see for instance [Mer96, CE05, Lec18, WWE20]. A beautiful introduction to the ideas in this area can be found in Mazur [Maz11].
- The second and third are both congruences between two *cusp forms*, of *different weights*. This resonates with the framework of p -adic families of modular forms, as developed by [Hid86b, Hid86a, Col97b, CM98] and many others, and it is these types of congruences that form the focus of this document. We note that both these examples are of a very different nature. Example 2 exhibits a congruence between a modular form of weight 1 and one of higher weight³, and results in an elementary description of the congruence class of $\tau(p)$ modulo 23. Example 3 on the other hand exhibits a congruence with a form associated to an elliptic curve. There is no similarly elementary characterisation of the congruence class of the $\tau(p)$ modulo 11.

In these notes, we will focus primarily on the theme of congruences between modular forms of different weights, and p -adic families. Traditionally, the theory was built around the prototypical example of the *Eisenstein family*, as in Coleman [Col97b], until more recent advances due to Pilloni [Pil13] and Andreatta–Iovita–Stevens [AIS14] on the geometric interpolation of line bundles, which allows us to develop the theory abstractly, without relying on the Eisenstein family. From a practical and computational point of view, this family remains of primordial importance, so the next section will quickly review it, motivated by the strategy of Serre to show the existence of the Kubota–Leopoldt p -adic L-function.

2. EISENSTEIN SERIES AND KUMMER CONGRUENCES

We begin with a brief discussion of the Kummer congruences, and introduce Serre’s important idea of inferring the p -adic variation of the constant term of a modular form, from that of its higher Fourier coefficients. This idea appeared in Serre [Ser73], and goes back to observations of Hecke [Hec24] and Siegel–Klingen [Kli62, Sie68]. It will make several appearances throughout these notes.

2.1. The Kummer congruences. Recall that the Riemann zeta function $\zeta(s)$ may be analytically continued to the entire complex plane, except for a simple pole with residue 1 at the point $s = 1$. It satisfies the functional equation

$$(25) \quad \pi^{-s/2} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \pi^{-(1-s)/2} \Gamma\left(\frac{1-s}{2}\right) \zeta(1-s).$$

³The existence of such congruences is an important ingredient in the aforementioned work of Deligne–Serre [DS74] on the existence of Artin representations attached to modular forms of weight 1.

Of special importance are its values at negative odd integers (or equivalently, by the functional equation, at positive even integers), which were computed first by Euler in 1734, and read on 5 December 1735 in the St. Petersburg Academy of Sciences. The starting point for Euler was the easily verified identity

$$(26) \quad \sin(\pi z) = \pi z \prod_{n \geq 1} \left(1 - \frac{z^2}{n^2}\right)$$

By taking the logarithmic derivative, we obtain the following identities

$$(27) \quad \pi z \cot(\pi z) = 1 - 2 \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{z^{2k}}{n^{2k}}$$

$$(28) \quad = 1 - 2 \sum_{k=1}^{\infty} \zeta(2k) z^{2k}$$

On the other hand, the Bernoulli numbers are defined via the generating series

$$(29) \quad \frac{t}{e^t - 1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}$$

and hence we can formally extract the even part of this series as

$$(30) \quad \frac{1}{2} \left(\frac{t}{e^t - 1} - \frac{-t}{e^{-t} - 1} \right) = \frac{t}{2} \cdot \frac{e^{t/2} + e^{-t/2}}{e^{t/2} - e^{-t/2}}$$

$$(31) \quad = \frac{t}{2} \cdot \coth\left(\frac{t}{2}\right)$$

Bearing in mind that $i \coth(iz) = \cot(z)$, we obtain the identity

$$(32) \quad \cot(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{(-1)^k 2^{2k} B_{2k}}{(2k)!} z^{2k-1}.$$

It now follows formally from (28) and (32) that

$$(33) \quad \zeta(2k) = \frac{(-1)^{k-1} (2\pi)^{2k}}{2(2k)!} B_{2k}$$

and hence by the functional equation

$$(34) \quad \zeta(1 - 2k) = \frac{-B_{2k}}{2k}.$$

The fact that the value of the zeta function at negative odd integers is a *rational number* is remarkable. We will revisit this in the more general setting of L-functions of totally real number fields, when we discuss the ‘explicit formula’ for this rational number obtained by Klingen and Siegel [Kli62, Sie68], following an idea of Hecke [Hec24] using diagonal restrictions of Hilbert Eisenstein series. This is discussed in § 4.5.

The Bernoulli numbers have interesting p -adic properties, notably by two results established in the mid-19th century which are the starting point for our investigations: The Clausen–von Staudt theorem [Cla40, vS40] and the *Kummer congruences* [Kum51]. For convenience, we assume henceforth that $p \neq 2$.

Lemma 2.1. *Suppose k, k' are two positive even integers such that $k \equiv k' \pmod{(p-1)p^n}$, then*

$$(35) \quad \begin{aligned} \text{If } (p-1) \nmid k & : (1 - p^{k-1})B_k/k \equiv (1 - p^{k'-1})B_{k'}/k' \pmod{p^{n+1}} \\ \text{If } (p-1) \mid k & : v_p(B_k/k) = -1 - v_p(k) \end{aligned}$$

The Kummer congruences show in particular that the quantity

$$(36) \quad (1 - p^{k-1}) \frac{B_k}{k} \quad k \in 1 + (p-1) \mathbf{Z}_{\geq 0}$$

exhibits p -adic continuity as a function of k . To the optimistic reader, this may already suggest that there may be an interesting p -adic analytic function $\zeta_p : \mathbf{Z}_p \rightarrow \mathbf{C}_p$ whose special values at arguments $s = 1 - k$ of the above form equal the quantities (36). The problem of finding this function is therefore vaguely akin to attempting to reconstruct the complex Riemann zeta function $\zeta(s)$ just from the knowledge of its special values $-B_k/k$ at negative odd arguments $s = 1 - k$. Kubota–Leopoldt [KL64] show:

Theorem 2.2 (Kubota–Leopoldt). *There is a unique p -adic analytic function $\zeta_p : \mathbf{Z}_p \rightarrow \mathbf{C}_p$ such that*

$$(37) \quad \zeta_p(1 - k) = (1 - p^{k-1}) \frac{-B_k}{k} \quad \text{for all } k \in 1 + (p-1) \mathbf{Z}_{\geq 0}$$

What it means precisely for a p -adically continuous function $\mathbf{Z}_p \rightarrow \mathbf{C}_p$ to be analytic will not be important for now, and we will not define it precisely until § 3.6 when we introduce the *Iwasawa algebra*.

2.2. The p -adic family of Eisenstein series. Serre observed that the congruences of Bernoulli numbers given in (35) can be upgraded to congruences between q -expansions of modular forms. Notice first that we see from the Kummer congruences that the Bernoulli numbers need to be modified by a factor $(1 - p^{k-1})$ in order to interpolate nicely as a function of k . Likewise, we need to adjust the Eisenstein series introduced above, by setting

$$(38) \quad \begin{aligned} \mathbf{G}_k^{(p)} &= (1 - p^{k-1} U_p) \mathbf{G}_k \\ &= (1 - p^{k-1}) \frac{-B_k}{2k} + \sum_{n \geq 1} \left(\sum_{p \nmid d|n} d^{k-1} \right) q^n \end{aligned}$$

which is a modular form for $\Gamma_0(p)$, often referred to as the (*ordinary*) p -stabilisation of \mathbf{G}_k . We define $\mathbf{E}_k^{(p)}$ to be its unique multiple with constant coefficient 1. Observe that elementary congruences for the higher Fourier coefficients yield an upgraded version of the congruences (35). More precisely, when

$$(39) \quad k \equiv k' \pmod{(p-1)p^n}$$

we have that $d^{k-1} \equiv d^{k'-1} \pmod{p^{n+1}}$ for any d not divisible by p . Indeed, for $n = 0$ this is the statement of Fermat’s little theorem, and the general case follows by induction. Therefore we obtain

$$(40) \quad \begin{aligned} \text{If } (p-1) \nmid k : \quad \mathbf{G}_k^{(p)}(q) &\equiv \mathbf{G}_{k'}^{(p)}(q) \pmod{p^{n+1}} \\ \text{If } (p-1) \mid k : \quad \mathbf{E}_k^{(p)}(q) &\equiv \mathbf{E}_{k'}^{(p)}(q) \pmod{p^{n+1}} \end{aligned}$$

The observation of Serre [Ser73], which was inspired by earlier ideas of Hecke [Hec24] and Siegel [Sie68], was that in establishing these congruences of Eisenstein series, there is a striking dichotomy between the congruences between the constant terms (which are the Kummer congruences, and hence somewhat deep) and the higher coefficients (which follow trivially from Fermat’s little theorem, and are hence not deep). His idea was to try and obtain the Kummer congruences, and the construction of the Kubota–Leopoldt zeta function $\zeta_p(s)$, by inheriting congruences of a more elementary nature from the higher coefficients.

This idea, whereby information on the constant coefficient is transferred from the higher coefficients, will appear several times throughout these notes, and is very powerful and useful in a variety of contexts. We will mark the paragraphs where it comes back by a small light bulb in the margin as shown here.



2.3. Serre's theory of p -adic modular forms. We now follow Serre [Ser73] and establish some basic definitions of p -adic modular forms. We follow Serre in restricting to the case of level 1 modular forms defined over \mathbf{Q}_p , but reassure the reader who is nervous about this that these assumptions will eventually be lifted when we adopt the more geometric viewpoint due to Katz in the next lecture.

For any formal power series in the variable q given by

$$(41) \quad f(q) = a_0 + a_1q + a_2q^2 + \dots \in \mathbf{Q}_p[[q]],$$

we define $v_p(f) = \inf_n(v_p(a_n))$, where v_p is the usual p -adic valuation on \mathbf{Q}_p . We define the space of p -adic modular forms to be the collection of $f(q) \in \mathbf{Q}_p[[q]]$ such that there is a sequence $f_i \in M_{k_i}(\mathrm{SL}_2(\mathbf{Z}))$ with rational Fourier coefficients, satisfying

$$(42) \quad v_p(f(q) - f_i(q)) \rightarrow \infty.$$

A p -adic modular form $f(q)$ therefore is obtained as a limit of q -expansions of classical modular forms. The following important proposition of Serre [Ser73, § 1.3 Théorème 1] states that the sequence of their weights k_i must tend to a limit p -adically. Its proof lies significantly deeper than the rest of the contents of [Ser73] which are otherwise largely established by more elementary means.

Proposition 2.3. *Let f, g be two classical modular forms of weights k, ℓ on $\mathrm{SL}_2(\mathbf{Z})$, both nonzero and normalised such that $v_p(f) = 0$. Suppose that we have*

$$(43) \quad v_p(f - g) \geq m$$

for some positive integer m , then it must be true that

$$(44) \quad \begin{array}{ll} k \equiv \ell \pmod{(p-1)p^{m-1}} & \text{if } p \geq 3 \\ k \equiv \ell \pmod{2^{m-2}} & \text{if } p = 2 \end{array}$$

As a consequence of this proposition, one easily checks that every p -adic modular form f has a well defined weight

$$(45) \quad k := \varprojlim_i k_i \in \mathbf{Z}_p \times \mathbf{Z} / (p-1)\mathbf{Z} = \varprojlim_m \mathbf{Z} / (p-1)p^m \mathbf{Z}.$$

For instance, it is not difficult to see that the q -expansion

$$\begin{aligned} \mathbf{E}_4(q)^{-1} &= (1 + 240q + 2160q^2 + 6720q^3 + \dots)^{-1} \\ &= 1 - 240q + 55440q^2 - 12793920q^3 + 2952385680q^4 + \dots \end{aligned}$$

is a 2-adic, 3-adic, and 5-adic modular form of weight -4 .

This is the point where Serre is able to realise the idea of “inheriting” congruences for the constant terms of Eisenstein series, from the much more elementary congruences between their higher coefficients. The following result is a corollary of 2.3, and we leave the proof to the reader.

Theorem 2.4 (Serre). *Suppose we have a sequence of p -adic modular forms of weights k_i :*

$$(46) \quad f_i(q) = a_0^{(i)} + a_1^{(i)}q + a_2^{(i)}q^2 + \dots$$

which satisfy the following two properties:

- The sequences $a_n^{(i)}$ tend uniformly to a limit $a_n \in \mathbf{Q}_p$,
- The weights k_i tend to a limit $k \neq 0$.



Then the constant terms $a_0^{(i)}$ tend to a limit a_0 , and the q -series

$$(47) \quad f(q) = a_0 + a_1q + a_2q^2 + \dots \in \mathbf{Q}_p[[q]]$$

is a p -adic modular form.

Notice that we may use the above theorem to show the existence of a continuous function interpolating the constant terms of the Eisenstein family! Indeed, when a sequence of integers k tends to a limit in \mathbf{Z}_p , we already noticed that the higher Fourier coefficients of $\mathbf{G}_k^{(p)}$ tend uniformly to a limit for elementary reasons. This implies that its constant term, which is

$$(48) \quad \zeta_p(1-k) = (1-p^{k-1})\zeta(1-k)$$

also tends to a limit, and it extends to a continuous function of k in $\mathbf{Z}_p \times \mathbf{Z}/(p-1)\mathbf{Z}$, which is precisely the Kubota–Leopoldt p -adic L-function. The paper of Serre pushes this idea further, and strengthens this significantly by deducing also its analytic properties. The above arguments may be strengthened to give an *effective* version of the claimed convergence, whose rate may be controlled to truly recover the Kummer congruences for Bernoulli numbers from elementary congruences between the higher coefficients.

2.4. Hecke operators and their spectrum. The space of p -adic modular forms is equipped with actions of Hecke operators T_ℓ, U_p, V_p , as was shown by Serre [Ser73, §2]. Suppose

$$(49) \quad f(q) = a_0 + a_1q + a_2q^2 + \dots$$

is a p -adic modular form of weight k , then $T_\ell f$ for $\ell \neq p$ and $U_p f$ are given by the expressions (5), and

$$(50) \quad V_p f(q) = \sum_{n \geq 0} a_n q^{np}.$$

One may wonder what can be said about these operators from the point of view of p -adic spectral theory [Ser62], and what, if any, is the arithmetic significance of the eigenvalues. Despite its great successes on the Kummer congruences, this is a point where the theory of p -adic modular forms starts lacking. Its definition – based solely on q -expansions – lacks the rigidity to avoid capturing a tremendous amount of power series in the space of p -adic modular forms whose arithmetic significance is less apparent.

One way to see this is as follows: Let f be any p -adic modular form, and choose any $\lambda \in p\mathbf{Z}_p$. Then

$$f_\lambda = (1 - \lambda V_p)^{-1} (1 - V_p U_p) f$$

exists as a p -adic modular form, has the same weight as f , and satisfies

$$(51) \quad U_p f_\lambda = \lambda f_\lambda.$$

This shows that the operator U_p has a spectrum that is far from discrete, and contains an overwhelmingly large continuous spectrum. To discern a discrete spectrum, this suggests that one should seek a more rigid framework that excludes these pathological eigenforms. We will find such a framework in the subspace of *overconvergent* modular forms defined by Katz, which we discuss in § 3. As we will see, the action of the operator U_p on overconvergent forms has a very rich and interesting discrete spectrum.

2.5. Eisenstein series of weight two. A celebrated borderline example is given by the Eisenstein series of weight two, and is of great arithmetic importance. Serre shows that, for any prime number p , the formal power series

$$\mathbf{E}_2(q) = 1 + 24 \sum_{n \geq 1} \left(\sum_{d|n} d \right) q^n$$

is a p -adic modular form. Indeed, as is well known, the series $\mathbf{E}_2^{(p)}(q) := (1 - pU_p) \mathbf{E}_2(q)$ is a classical modular form of weight 2 and level $\Gamma_0(p)$. Serre shows that any form of level $\Gamma_0(p)$ is a limit of modular forms of level 1, and therefore defines a p -adic modular form in the above sense. It follows that

$$\begin{aligned} \mathbf{E}_2(q) &= (1 - pU_p)^{-1} \mathbf{E}_2^{(p)} \\ &= \mathbf{E}_2^{(p)} + pU_p \mathbf{E}_2^{(p)} + p^2 U_p^2 \mathbf{E}_2^{(p)} + \dots \end{aligned}$$

is then also a p -adic modular form. Note that this argument is valid for any prime p .

In the next section, we introduce the subspace of overconvergent modular forms according to Katz. It was shown by Coleman–Gouvêa–Jochowitz [CGJ95] that the form \mathbf{E}_2 is never overconvergent. This example is nonetheless of tremendous arithmetic importance, and we content ourselves by mentioning its role in the theory of p -adic heights on ordinary elliptic curves. A beautiful discussion, along with a precise quantification of its failure to overconverge can be found in Mazur–Stein–Tate [MST06].

3. OVERCONVERGENT MODULAR FORMS

We encountered the Kubota–Leopoldt p -adic zeta function, and explored an idea of Serre that uses the p -adic Eisenstein family to construct it. This involved the notion of p -adic modular forms, which therefore served a great purpose, but otherwise seemed somewhat lacking in finer structural properties, as evidenced by the absence of an interesting discrete spectrum of Hecke operators. We now follow Katz in reinterpreting the viewpoint of Serre geometrically, and identifying much smaller – though still infinite-dimensional – subspaces of the space of p -adic modular forms. Though we recall much of what we need, we will assume some familiarity with the algebro-geometric theory modular forms. Excellent expositions can be found for instance in Katz [Kat73], Calegari [Cal13], and Loeffler [Loe14].

3.1. The Hasse invariant. Suppose S is a scheme over \mathbf{F}_p , then there is an *absolute Frobenius* morphism

$$(52) \quad F_{\text{abs}} : S \longrightarrow S$$

given on affine opens by the map on functions $f \mapsto f^p$. If X/S is an S -scheme, we define the scheme $X^{(p)} = X \times_S S$ where the fibre product is taken over S , viewed as an S -scheme via F_{abs} . The *relative Frobenius* morphism $F = F_{X/S}$ is defined by the following commutative diagram, where the square is Cartesian:

$$(53) \quad \begin{array}{ccc} X & \xrightarrow{F_{\text{abs}}} & X \\ \downarrow F_{X/S} & \searrow & \downarrow \\ X^{(p)} & \longrightarrow & X \\ \downarrow & & \downarrow \\ S & \xrightarrow{F_{\text{abs}}} & S \end{array}$$

Notice that the relative Frobenius is an S -linear morphism, whereas the absolute Frobenius is not! Also, the scheme $X^{(p)}$ is hardly a mysterious thing: Suppose X is of finite type over $\mathbf{F}_q / \mathbf{F}_p$, then $X^{(p)}$ is given

by the same equations as X , but where all the coefficients are raised to the p -th power. Note that if $q = p$, then we have $X^{(p)} = X$.

Now suppose that E/S is an elliptic curve, then the relative Frobenius $F = F_{E/S}$ is an isogeny, and hence has a dual isogeny V :

$$(54) \quad \begin{array}{lll} F : E & \longrightarrow & E^{(p)} & \text{“Frobenius”} \\ V : E^{(p)} & \longrightarrow & E & \text{“Verschiebung”} \end{array}$$

Suppose now that $S = \text{Spec}(\overline{\mathbf{F}}_p)$, then we say

$$(55) \quad \begin{cases} E \text{ is ordinary} & \text{if } E[p](\overline{\mathbf{F}}_p) \neq 1 \\ E \text{ is supersingular} & \text{if } E[p](\overline{\mathbf{F}}_p) = 1 \end{cases}$$

In general, we say E/S is ordinary/supersingular if all its geometric fibres are.

Proposition 3.1. *Suppose E/S is an elliptic curve, and S is an \mathbf{F}_p -scheme. Then we have:*

- E/S is ordinary if and only if $V : E^{(p)} \longrightarrow E$ is étale.
- $E/\overline{\mathbf{F}}_p$ is supersingular, only if E is defined over \mathbf{F}_{p^2} .

Proof. We can factor the multiplication by p map as

$$[p] : E \xrightarrow{F} E^{(p)} \xrightarrow{V} E.$$

This implies that V is separable if and only if $\text{Ker}(V)(\overline{\mathbf{F}}_p) \neq 1$ on all geometric fibres. Since the kernel of Frobenius only has the trivial geometric point, this is equivalent to $\text{Ker}([p])(\overline{\mathbf{F}}_p) \neq 1$. This proves the first statement. For the second statement, we have that E/S is supersingular if and only if V is inseparable, which means it must factor through Frobenius:

$$V : E^{(p)} \xrightarrow{F} E^{(p^2)} \longrightarrow E$$

The latter map must be finite of degree 1, and hence an isomorphism. Thus E is defined over \mathbf{F}_{p^2} . \square

Finally, we define the *Hasse invariant* of an elliptic curve E/R where R is a ring of characteristic p . First, choose $\omega \in H^0(E, \Omega_{E/R}^1)$ to be an R -basis, and let η be the R -basis of $H^1(E, \mathcal{O}_E)$ defined via Serre duality. The Hasse invariant $A(E, \omega)$ is the element of R defined by

$$(56) \quad F_{\text{abs}}^*(\eta) = A(E, \omega) \cdot \eta$$

Note that by the previous proposition, $E/\overline{\mathbf{F}}_p$ is ordinary if and only if $A(E, \omega) \neq 0$ for any choice of ω .

3.2. Algebraic modular forms. We will now see how to interpret the Hasse invariant as an algebraic modular form over \mathbf{F}_p of weight $p - 1$. Over the field of complex numbers \mathbf{C} , we are used to thinking of modular forms in terms of their q -expansions, which directly describe them as a function on the upper half plane in the variable q . Over other base fields, one adapts a more algebraic viewpoint, where a modular form becomes a section of a line bundle $\omega^{\otimes k}$ over a moduli space of elliptic curves. The algebraically defined notion of q -expansion, defined by its evaluation on the Tate curve, no longer directly describes the form as a “function” on classes of elliptic curves. This is quite striking for the Hasse invariant, which vanishes at supersingular points, yet has q -expansion given by 1. We now discuss these notions, briefly recalling the notion of algebraic modular forms, referring to Katz [Kat73, Chapter 1] for more details.

A *weakly holomorphic modular form* of weight $k \in \mathbf{Z}$ over a ring A is a rule which assigns to any isomorphism class of pairs $(E/R, \omega)$, where

- E/R is an elliptic curve over an A -algebra R ,
- ω is an R -basis for $H^0(E, \Omega_{E/R}^1)$,

an element $f(E/R, \omega) \in R$ such that the following two properties are satisfied:

$$\begin{aligned} f((E/R, \omega) \otimes_{\phi} R') &= \phi(f(E, \omega)) \quad \text{for all } \phi : R \rightarrow R' \text{ of } A\text{-algebras,} \\ f(E, \lambda\omega) &= \lambda^{-k} f(E, \omega) \quad \text{for all } \lambda \in R^{\times}. \end{aligned}$$

The q -expansion of a weakly holomorphic modular form f is defined as

$$(57) \quad f(q) := f((\text{Tate}(q)_{\mathbf{Z}((q))}, \omega_{\text{can}}) \otimes R) \in R((q)),$$

where $\text{Tate}(q)$ is the Tate elliptic curve over $\mathbf{Z}((q))$ defined by

$$(58) \quad y^2 + xy = x^3 + B_q x + C_q, \quad \omega_{\text{can}} = \frac{dx}{2y + x}$$

with coefficients defined by the explicit q -series in $\mathbf{Z}[[q]]$

$$(59) \quad \begin{aligned} B_q &= \sum_{n \geq 1} -5\sigma_3(n)q^n \\ C_q &= \sum_{n \geq 1} \frac{-5\sigma_3(n) - 7\sigma_5(n)}{12} q^n \end{aligned}$$

We say a weakly holomorphic modular form is an *algebraic* (or *holomorphic*) modular form if its q -expansion, which a priori is an element of $R((q))$, is in fact in $R[[q]]$.

Remark. The Tate curve arose first in the work of Tate [Tat95] on p -adic uniformisation of elliptic curves. Over the complex numbers \mathbf{C} , any elliptic curve E is isomorphic to $\mathbf{C}/\langle 1, \tau \rangle$ where $\langle 1, \tau \rangle$ denotes the lattice spanned by $1, \tau$ in \mathbf{C} . By exponentiation, this quotient can also be described as $\mathbf{C}^{\times}/q^{\mathbf{Z}}$ where as usual $q = \exp(2\pi i\tau)$, and the isomorphism with E involves explicit complex analytic functions which go back at least to Weierstraß. Tate showed that this admits a p -adic analogue; more precisely that for every E whose j -invariant is not p -adically integral, there exists an isomorphism with $\text{Tate}(q)$ for some $|q| < 1$, the latter being isomorphic to $\mathbf{C}_p^{\times}/q^{\mathbf{Z}}$ by explicit p -adic analytic power series.

We now see from the definition of the Hasse invariant in the previous section that it is naturally an algebraic modular form of weight $p - 1$ over \mathbf{F}_p . Indeed, it is a rule that attaches to $(E/R, \omega)$ an element of R whose functoriality is clear by definition, and moreover for any $\lambda \in R^{\times}$ we see that:

$$\begin{aligned} A(E, \lambda\omega) \cdot \lambda^{-1}\eta &= F_{\text{abs}}^*(\lambda^{-1}\eta) \\ &= \lambda^{-p} F_{\text{abs}}^*(\eta) \\ &= \lambda^{1-p} \cdot A(E, \omega) \cdot \lambda^{-1}\eta \end{aligned}$$

We conclude that the Hasse invariant A defines a weakly holomorphic modular form of weight $p - 1$ and level one. It has the following important properties, which we will freely use in what follows:

- The q -expansion of the Hasse invariant was computed in [Kat73, KM85] and is given by $A(q) = 1$. The proof is a beautiful argument using the Cartier operator.
- We already know that for E/k over $k = \overline{\mathbf{F}}_p$, the Hasse invariant vanishes if and only if E is supersingular. In fact, it has *simple* zeroes, in the sense that if R is a local Artinian k -algebra, and E/R is such that $V : E^{(p)} \rightarrow E$ induces the zero map on tangent spaces, then it must be true that there is a supersingular elliptic curve E_0/k such that

$$(60) \quad E_0 \times_k R \simeq E.$$

3.3. Overconvergent modular forms. We now come to the main spaces of interest, which were defined by Katz [Kat73]. They revolve around the properties of the Hasse invariant A over \mathbf{F}_p , but are spaces of forms over a p -adic field, and therefore involve liftings \tilde{A} of the Hasse invariant to characteristic zero.

Remark. We now come to a point where *rigid geometry* is most naturally used. The reader unfamiliar with this framework should not be alarmed, as we take a very pedestrian approach that should be digestible if one is willing to take a few things on faith. Moreover, the next section will show these notions in action in an extended example, and the uninitiated reader may prefer to skip ahead. The foundations of rigid geometry are beautifully summarised in Conrad [Con08].

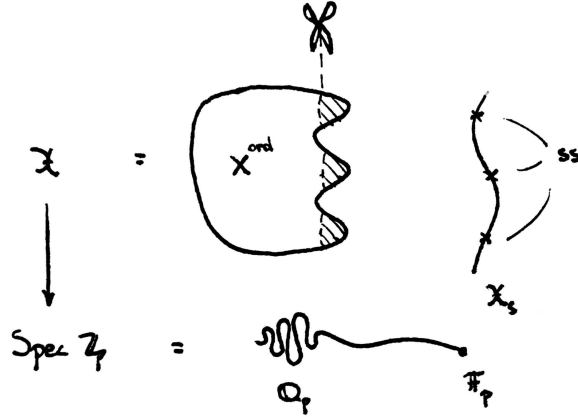
Suppose⁴ $N \geq 5$ and $p \nmid N$ prime. We let $\mathcal{X} / \mathbf{Z}_p$ be the modular curve over \mathbf{Z}_p which classifies generalised elliptic curves with $\Gamma_1(N)$ -level structure, universal curve $\pi : \mathcal{E} \rightarrow \mathcal{X}$, and closed subscheme of cusps \mathcal{I}_C , and denote its generic and special fibres by X and \mathcal{X}_s respectively. Define the line bundle

$$(61) \quad \omega := \pi_* \Omega_{\mathcal{E}/\mathcal{X}}^1(\log \pi^{-1} \mathcal{I}_C).$$

The Hasse invariant is the unique⁵ element of $H^0(\mathcal{X}_s, \omega^{\otimes p-1})$ with q -expansion 1. Let \mathbf{C}_p be the completion of the algebraic closure of \mathbf{Q}_p . Since the relative curve $\mathcal{X} / \mathbf{Z}_p$ is proper, every \mathbf{C}_p -point extends uniquely to an $\mathcal{O}_{\mathbf{C}_p}$ -point, and we obtain a reduction map

$$(62) \quad \text{red} : \mathcal{X}(\mathbf{C}_p) \rightarrow \mathcal{X}_s(\overline{\mathbf{F}}_p).$$

This map provides our first encounter with analytic geometry over the p -adic numbers. The inverse image $\text{red}^{-1}(x)$ of a closed point of the special fibre is isomorphic to a rigid analytic open disk $D = \{x \in \mathbf{C}_p : |x| < 1\}$. We saw previously that the vanishing locus of the Hasse invariant is precisely the supersingular locus of \mathcal{X}_s , which consists of a finite set of closed points. The *ordinary locus* X^{ord} is the affinoid open whose set of \mathbf{C}_p -points correspond to elliptic curves with ordinary reduction, which is therefore the complement of a finite number of rigid analytic open disks, indexed by the supersingular points.



⁴For simplicity, we will choose some auxiliary level structure to rigidify the moduli problem of elliptic curves and work on modular curves. If desired, this can be avoided by working on the moduli stack.

⁵The q -expansion principle states that any modular form with a given q -expansion and weight is uniquely determined.

At this point, we are ready to give a geometric reinterpretation of the spaces of p -adic modular forms introduced by Serre, which were discussed in § 2. The following theorem is due to Katz [Kat73].

Theorem 3.2 (Katz). *The space of p -adic modular forms of weight k is isomorphic as a Hecke module to*

$$(63) \quad H^0(X^{\text{ord}}, \omega^{\otimes k}).$$

In light of this theorem, we see in particular that if \tilde{A} is a *lift of the Hasse invariant*, which is any modular form of weight $p - 1$ over \mathbf{Z}_p whose q -expansion is congruent to 1 modulo p , then it must be invertible as a p -adic modular form. This may be seen explicitly also in Serre's language, since we have

$$(64) \quad v_p \left(\tilde{A}(q)^{p^n-1} - \tilde{A}(q)^{-1} \right) \rightarrow \infty$$

so that the formal q -expansion $\tilde{A}(q)^{-1}$ is a p -adic modular form in the sense of Serre.

Remark. When we do explicit calculations later on, we will see that the fact that the q -expansion of the Hasse invariant is 1 usually allows us to choose an Eisenstein series \mathbf{E}_{p-1} as a lift of the Hasse invariant. Indeed, when $p \geq 5$ this is a modular form of weight $p - 1$ whose q -expansion is congruent to 1 modulo p .

We saw previously that the space of p -adic modular forms is too large to have nice spectral properties, prompting Katz to consider subspaces of sections that extend⁶ to affinoids strictly containing X^{ord} . More precisely, let $0 \leq r \leq 1$, and define $X^{\text{ord}} \subset X_r \subset X^{\text{rig}}$ by

$$(65) \quad X_r(\mathbf{C}_p) := \{x \in X(\mathbf{C}_p) : v_p(\tilde{A}_x) \leq r\},$$

where \tilde{A}_x is a local lift of the Hasse invariant A at x . Note we do not require a global lift of the Hasse invariant to exist, which may fail when $p = 2, 3$. We define the space of r -overconvergent modular forms of integer weight k on $\Gamma_1(N)$ to be

$$(66) \quad M_k^\dagger(r) := H^0(X_r, \omega^{\otimes k}).$$

These spaces come with an action of Hecke operators T_ℓ for $\ell \nmid Np$, and U_ℓ for $\ell \mid N$, defined by restricting the Hecke correspondences on \mathcal{X} . They have the usual effect on q -expansions.

In addition, the operators U_p and V_p defined on p -adic modular forms may be defined geometrically, and preserve the subspace of overconvergent modular forms. More precisely, they are defined for every $r < 1/(p + 1)$ and have the following effect on the rate of overconvergence:

$$(67) \quad \begin{aligned} U_p &: M_k^\dagger(r) \longrightarrow M_k^\dagger(pr) \\ V_p &: M_k^\dagger(pr) \longrightarrow M_k^\dagger(r) \end{aligned}$$

In particular, the operator U_p improves the rate of overconvergence. The reason for the existence of the operators U_p and V_p is the *canonical subgroup section* s of the natural forgetful map of modular curves, which exists for any $r < p/(p + 1)$:

$$(68) \quad \begin{array}{ccc} X(\Gamma_1(N) \cap \Gamma_0(p))^{\text{rig}} & \longrightarrow & X^{\text{rig}} \\ & \searrow \text{dashed } s & \cup \\ & & X_r \end{array}$$

This yields two equivalent ways to view spaces of overconvergent modular forms:

- As sections on affinoid opens of X , with no level at p (the “tame” viewpoint).

⁶This definition may seem obscure to the uninitiated, but already at the time of writing of [Kat73] represented a dominant theme in the school of p -adic analysis surrounding Bernard Dwork and his disciples, and continues to be hugely influential to this day.

- As sections on affinoid opens of the modular curve obtained from X by adding additional $\Gamma_0(p)$ level structure (the “canonical subgroup” viewpoint).

For theoretical questions, the latter is frequently more convenient, for instance it forms the natural setting for questions of analytic continuation and Coleman’s classiciality results discussed in § 3.7. On the other hand, for computational purposes the former often has advantages, since it allows one to compute with auxiliary classical spaces of modular forms, as we will see in § 4.1.

3.4. Interlude: Extended example. Let us explore these abstract definitions in a particular case, to get a feeling for the various objects involved. Consider the case where $p = 2$ and $k = 0$, in level one. In this case, we can be very explicit about the spaces of p -adic and r -overconvergent modular forms, both from the “tame” viewpoint (in level one) or via the canonical subgroup section (on $X_0(2)$). This example is centred around the properties of the Klein j -invariant

$$(69) \quad j(q) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + 864299970q^3 + \dots$$

which is a level one modular form of weight zero with a simple pole at the cusp.

3.4.1. The “tame” viewpoint. Consider the moduli stack X of elliptic curves. Of the 4 values in \mathbf{F}_4 for the j -invariant, only $j = 0$ is supersingular, so that its special fibre at $p = 2$ has a unique supersingular point corresponding to the vanishing locus of j . It follows that the ordinary locus on X is described by $|j^{-1}| \leq 1$, and hence the space of 2-adic modular forms of weight 0 is isomorphic to

$$(70) \quad \mathbf{C}_2\langle j^{-1} \rangle = \{a_0 + a_1j^{-1} + a_2j^{-2} + \dots \mid a_n \rightarrow 0\}$$

For any r , the space $M_0^\dagger(r)$ defines a Banach space contained inside of this Tate algebra, which we can explicitly identify through growth conditions on the coefficients a_n . Precisely, we use the observation that

$$(71) \quad j = \frac{\mathbf{E}_4^3}{\Delta}$$

and $\mathbf{E}_4 = 1 + 240q + \dots$ is the normalised Eisenstein series of weight 4, which is a lift of the fourth power of the Hasse invariant A^4 . In particular, we find that on the supersingular disk (where Δ is invertible, and hence $v_2(\Delta) = 0$, we have that

$$(72) \quad v_2(A) \leq r \quad \Longleftrightarrow \quad v_2(j) \leq 12r.$$

and as a consequence, we get that the subspace of r -overconvergent forms is given by

$$(73) \quad M_0^\dagger(r) = \{a_0 + a_1j^{-1} + a_2j^{-2} + \dots : |a_n|p^{12nr} \rightarrow 0\}$$

Finally, let us compute some Hecke operators, and see whether the obtained results make sense with what is said above. First, note that we can compute very rapidly (most serious computer algebra packages like Magma, PARI/GP, or Sage will already have a function implemented) the q -expansion of j^{-1} . Given any 2-adic modular form of weight 0, we can then compute its j^{-1} -expansion very rapidly by the simple observation that j^{-1} vanishes to order 1 at the cusp infinity, and hence we can inductively subtract powers

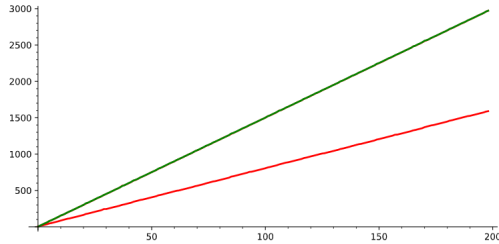
of j^{-1} until we are left with zero. Carrying out this procedure in Magma [BCP97], we obtain that

$$\begin{aligned}
 U_2 j^{-1} &= -744 j^{-1} \\
 &\quad -140914688 j^{-2} \\
 &\quad -16324041375744 j^{-3} \\
 &\quad -1528926232501026816 j^{-4} \\
 &\quad + \dots \\
 T_3 j^{-1} &= 356652 j^{-1} \\
 &\quad -16114360320000 j^{-2} \\
 &\quad +1298216343568384000000/3 j^{-3} \\
 &\quad + \dots \\
 T_5 j^{-1} &= 49336682190 j^{-1} \\
 &\quad -122566701099729715200000 j^{-2} \\
 &\quad +177278377115100363578123747328000000 j^{-3} \\
 &\quad + \dots
 \end{aligned}$$

where we calculated in reality hundreds of terms, which look rather unappetising. Things become very interesting when we look at the 2-adic valuations of the coefficients a_1, a_2, a_3, \dots of $U_2 j^{-1}$ and $T_\ell j^{-1}$ tabulated above, which give us the following sequences:

$$\begin{aligned}
 (74) \quad U_2 j^{-1} &: v_2(a_n) = 3, 12, 20, 28, 35, 46, 52, 60, 67, 76, 86, 94, \dots \\
 T_3 j^{-1} &: v_2(a_n) = 2, 16, 32, 45, 60, 79, 91, 105, 120, 136, 154, 165, \dots \\
 T_5 j^{-1} &: v_2(a_n) = 1, 18, 33, 47, 61, 80, 92, 107, 121, 138, 155, 167, \dots
 \end{aligned}$$

We see very clearly that the latter two sequences grow roughly at the same rate, whereas the first one grows significantly more slowly! In fact, if we plot these three sequences in red, green, and blue respectively, for the first two hundred terms, we obtain the following picture:



They all look like linear functions! The green and blue plots are virtually indistinguishable at this scale, and look roughly like a linear function of slope 15. On the other hand, at this scale the red plot looks roughly like a linear function of slope 8. This is precisely what we expected from the general theory, since j^{-1} is r -overconvergent for any r (indeed, it converges on the entire modular curve X except for a simple pole at the cusp 0!) and its image under the U_2 -operator is therefore only guaranteed to be r -overconvergent for any $r < p/(p+1) = 2/3$. With respect to the identification (73), this shows that the valuation of the coefficients should grow at least like a linear function of slope $8 = (2/3) \cdot 12$.

3.4.2. The “canonical subgroup” viewpoint. Even though we can compute things to our heart’s desire, it is hard to get any more specific information in the tame description (i.e. on $X = X_0(1)$). Following Buzzard–Calegari [BC05], we will now see that we can get a lot of mileage from working on $X_0(2)$ instead, which

we know we can by the theory of the canonical subgroup. Define the *Hauptmodul*

$$(75) \quad h = \Delta(2z)/\Delta(z) = q \prod_{n \geq 1} (1 + q^n)^{24}$$

which is a meromorphic function on $X_0(2)$ with a simple zero at the cusp ∞ , and a pole at the cusp 0. It is related to the j -function by

$$(76) \quad \frac{h}{(1 + 2^8 h)^3} = j^{-1}$$

Using a Newton polygon argument, we see that we can find a canonical section of the forgetful map whenever $v_p(j^{-1}) > -8$ exactly as predicted by the theory of canonical subgroups. Note also that in this case, we see that this section does not extend to any larger region, so the result was optimal! This means that we get an alternative description for (73) of the form

$$(77) \quad M_0^\dagger(r) = \{a_0 + a_1 h + a_2 h^2 + \dots : |a_n| p^{12nr} \rightarrow 0\}$$

The advantage is the following: The Hecke operators are defined as correspondences on $X_0(2)$, and hence we know that $U_2(h)$ and $T_\ell(h)$ are *polynomials* in h ! This is in stark contrast with the tame situation, where we got a rather mysterious set of power series, which we could compute to any accuracy, but never exactly. In contrast, on $X_0(2)$ we can do the computation exactly, and we obtain:

$$\begin{aligned} U_2(h) &= 24h + 2048h^2 \\ T_3(h) &= 300h + 98304h^2 + 16777216/3h^3 \\ T_5(h) &= 18126h + 40239104h^2 + 14696841216h^3 + 1649267441664h^4 + 281474976710656/5h^5 \end{aligned}$$

Together with (77), this can be seen as a complete description of the U_2 -module $M_0^\dagger(r)$. This is what is used by Buzzard–Calegari [BC04] to determine the valuations of all the eigenvalues of U_2 on this space.

3.5. Spectral theory of U_p . We now discuss an important part of the subject, which is the spectral theory of the Hecke operator U_p , acting on the \mathbf{C}_p -Banach space of r -overconvergent forms.

We begin by defining a norm $\|\cdot\|_r$ on $M_k^\dagger(r)$. Pick a point $x \in X_r$, let K be a finite extension of the residue field of x , and let $\text{Spec}(K) \rightarrow \mathcal{X}_{\mathbf{Q}_p}$ be a point whose image corresponds to x . The properness of \mathcal{X} implies that this extends uniquely to a point $\varphi : \text{Spec}(\mathcal{O}_K) \rightarrow \mathcal{X}$. Now let $f \in M_k^\dagger(r)$, then $\varphi^* f = a_f s$ for some section s generating the trivial line bundle $\varphi^* \omega^{\otimes k}$ and some $a_f \in \mathcal{O}_K$. We set

$$(78) \quad |f(x)| := |a_f|,$$

which is independent of the choice of s . The norm

$$(79) \quad \|f\|_r := \sup\{|f(x)| : x \in X_r\}$$

makes $M_k^\dagger(r)$ into a p -adic Banach space. This induces the structure of a p -adic Fréchet space on

$$(80) \quad M_k^\dagger := \varinjlim_{r>0} M_k^\dagger(r)$$

which we call the space of *overconvergent modular forms*. The Banach spaces $M_k^\dagger(r)$ are infinite-dimensional, and there is a priori no meaningful way to talk about the spectrum of an operator, unless we know more.

Suppose we have a continuous bounded operator T on a separable \mathbf{C}_p -Banach space B , then we say that T is *compact* if it is the limit of operators of finite rank. Equivalently, T is compact if and only if the image of the unit ball is relatively compact. There is a well-developed spectral theory for compact operators, see [Dwo62, Ser62, Col97b], which has the following pleasant consequences for compact operators:

- T has a discrete spectrum of non-zero eigenvalues

$$(81) \quad |\lambda_1| \geq |\lambda_2| \geq \dots$$

where $|\lambda_i| \rightarrow 0$ as $i \rightarrow \infty$, whose inverses are the roots of a well-defined characteristic series

$$\begin{aligned} P(t) &= \text{"det}(1 - Tt) \\ &= a_0 + a_1 t + a_2 t^2 + \dots, \quad \text{where } a_i \rightarrow 0 \text{ as } i \rightarrow \infty. \end{aligned}$$

- For every $v \in B$ there are constants c_i and generalised eigenvectors v_i with eigenvalue λ_i such that for any $\varepsilon > 0$ we have (asymptotically in n) that

$$(82) \quad \varepsilon^{-n} \left\| T^n v - T^n \sum_{|\lambda_i| \geq \varepsilon} c_i v_i \right\| \rightarrow 0.$$

The constants c_i are often called the coefficients of the *asymptotic expansion* of v .

We established that the operator U_p exhibits a *contractive* nature, and improves overconvergence as described by (67). This implies that U_p is compact, and hence possesses a well-defined characteristic series. Here is one concrete way to think about this series (and indeed, to compute it in examples!) as explained by Serre [Ser62] and Coleman [Col97b, Theorem A2.1]: Suppose we have an orthonormal basis

$$(83) \quad \{f_1, f_2, f_3 \dots\} \quad \text{for } M_k^\dagger(r),$$

then we obtain an infinite matrix representation of U_p . In the example above, where $p = 2$ and $k = 0$, we already noted that we have an algorithm to compute this matrix *exactly*, or at least any finite submatrix of it. To see what compactness really means in practice, we compute the first 10×10 submatrix with respect to the basis $f_i = (2^8 h)^i$ of the cuspidal subspace, and look at the 2-adic valuations of its entries:

$$(84) \quad v_2(U_2(i, j))_{i,j} = \begin{pmatrix} 3 & 8 \\ 3 & 7 & 11 & 16 \\ & 8 & 12 & 17 & 19 & 24 \\ & 7 & 11 & 15 & 21 & 23 & 27 & 32 \\ & & 11 & 19 & 20 & 25 & 27 & 35 & 35 & \dots \\ & & 11 & 16 & 20 & 24 & 27 & 33 & 35 \\ & & & 17 & 19 & 24 & 29 & 34 & 35 \\ & & & & 15 & 20 & 23 & 27 & 31 & 38 \\ & & & & & 19 & 24 & 27 & 37 & 36 \\ & & & & & & \vdots & & & \ddots \end{pmatrix}$$

Here, we omitted the entries of U_2 that were equal to zero. The compactness of U_p in orthonormalisable situations like this one is *equivalent* to the statement that the column vectors converge uniformly to 0 in the infinite matrix representation. In the above example, that certainly looks plausible, as the entries of the columns seem to have valuation which grows roughly at the same rate. To contrast this with what happens

in general, let us compute with respect to the same basis the first 10×10 submatrix for T_3 :

$$(85) \quad v_2(T_3(i, j))_{i,j} = \begin{pmatrix} 2 & 12 & 16 & & & & & & & \\ 7 & 2 & 11 & 20 & 27 & 32 & & & & \\ 8 & 8 & 2 & 14 & 17 & 28 & 34 & 46 & 48 & \\ & 11 & 8 & 2 & 12 & 19 & 29 & 36 & 43 & \\ & 16 & 9 & 10 & 2 & 12 & 16 & 32 & 34 & \dots \\ & 16 & 15 & 12 & 7 & 2 & 11 & 22 & 28 & \\ & & 18 & 19 & 8 & 8 & 2 & 16 & 18 & \\ & & 23 & 19 & 17 & 12 & 9 & 2 & 13 & \\ & & 24 & 25 & 18 & 17 & 10 & 12 & 2 & \\ & & & & \vdots & & & & & \ddots \end{pmatrix}$$

Notice the stark contrast with the matrix of U_2 . Whereas the general entry of every column seems like it tends to zero (as it should, since T_3 still defines an operator on the Banach space $M_0^\dagger(2/3)$ after all) it does not look like the general column tends uniformly to zero. Most strikingly, the diagonal entries all seem to have valuation 2, suggesting this operator may not have a convergent “trace”.

For the operator U_2 we can also compute an approximation for its characteristic series $P(t)$, using the above matrix. One can easily analyse to which precision the given answer is correct, but we will ignore such issues here. We truncate the matrix for U_2 as above, and obtain a polynomial whose coefficients are 2-adically close to those of $P(t)$. Looking at the Newton polygon, we see that the valuations of the eigenvalues of U_2 on the full space $M_0^\dagger(r)$ for any r are as follows:

$$(86) \quad \mathbf{0}_1, \mathbf{3}_1, \mathbf{7}_1, \mathbf{13}_1, \mathbf{15}_1, \mathbf{17}_1, \dots$$

Here, we denote the valuations of the eigenvalues by bold type, and the multiplicity of that valuation by a subscript. It is striking that these are all integers, since there is no a priori reason that they should be! In this particular example, there is an explicit expression for the general term in this sequence, found by Buzzard–Calegari [BC05]. We give a brief overview of their arguments.

Let h be the Hauptmodul defined in (75). Then a basis for the cuspidal subspace $S_0^\dagger(r) \subset M_0^\dagger(r)$ is given by the powers h, h^2, h^3, \dots where a general element is an infinite sum of these forms, where the coefficients decay in a controlled matter, depending on r . We may try to find an explicit description of the basis for U_2 with respect to this basis. It is easily verified that for $n \geq 2$ we have the recursion

$$U_2(h^n) = (48h + 4096h^2)U_2(h^{n-1}) + hU_2(h^{n-2}).$$

We know from (77) that the powers of $2^6 h$ form an orthonormal basis of $M_0^\dagger(1/2)$, and the above recursion implies that the (i, j) -th entry in the matrix for U_2 with respect to this basis is given by

$$\frac{3j(i+j-1)!2^{2i+2j-1}}{(2i-j)!(2j-i)!}.$$

In spite of the matrix for U_2 being completely explicit, it is still no laughing matter to compute its slopes, and more ideas are required. It was shown by Buzzard–Calegari [BC05, Lemma 4] by a really intriguing direct computation using a hypergeometric summation formula, that there exist matrices A, B with entries in \mathbf{Z}_2 which are both congruent to the identity matrix modulo 2, and such that ADB equals the matrix of U_2 , where D is the diagonal matrix with (i, i) -th entry given by

$$\frac{2^{4i+1}(3i)!^2 i!^2}{3(2i)!^4}.$$

From this, one may deduce that the matrix of U_2 has a characteristic series whose Newton polygon is the same as that for the matrix D , which implies:

Theorem 3.3 (Buzzard–Calegari). *The slope sequence of U_2 on $S_0^\dagger(r)$ for any $r > 0$ is given by*

$$\left\{ 1 + 2v_2 \left(\frac{(3n)!}{n!} \right) \right\}_{n=1, \dots, \infty}$$

The study of *slopes* was very popular in the early 21st century, see for instance [Buz05, BC04, BC05, BP16, BG16] and the references contained therein. A good knowledge of the spectrum, such as the example above, leads to a streamlined way to prove many classical congruences of modular forms, such as that of Lehner [Leh49] for the Fourier coefficients a_n of the j -function, which states that

$$a_n \equiv 0 \pmod{2^{3n+8}} \quad \text{whenever} \quad n \equiv 0 \pmod{2^n}.$$

The appearance of 3 in the exponent is a reflection of the first positive slope being 3 in the theorem of Buzzard–Calegari, and it can be strengthened and refined using the higher slopes. This domain has in recent years shifted its fashions towards the *boundary* of weight space, such as the works [BK05, Roe14, LWX17, AIP18] and many others. This is a fascinating notion that falls outside the narrative we take here, but we mention a spectacular recent application in the proof by Newton–Thorne [NT19] of modularity of $\text{Sym}^n(f)$ when f is a cuspidal eigenform satisfying certain conditions, including all forms of level 1.

3.6. The eigencurve. The above constructions may be extended to incorporate families of modular forms, culminating in the existence of the *eigencurve*. This is a geometric object that provides a powerful picture when thinking about families of overconvergent modular forms. The theory is due mainly to Coleman [Col96, Col97b] and Coleman–Mazur [CM98] and was revisited more recently by Pilloni [Pil13] and Andreatta–Iovita–Stevens [AIS14]. We content ourselves with a very brief discussion in these notes.

Our desire is to explain congruences between modular forms by interpolating between different weights, as in the theory of Serre. The geometric theory of overconvergent forms is restricted to integral weights $k \in \mathbf{Z}$, and to overcome the lack of a sheaf ω^κ for a p -adic weight other than $\kappa \in \mathbf{Z}$, the idea of Coleman was to turn once more to the Eisenstein family, which is defined for any *weight-character*

$$(87) \quad \kappa \in \mathcal{W} := \text{Hom}_{\text{cont}}(\mathbf{Z}_p^\times, \mathbf{C}_p^\times)$$

where we can view a pair (k, χ) consisting of $k \in \mathbf{Z}$ and $\chi : (\mathbf{Z}/p^n \mathbf{Z})^\times \rightarrow \mathbf{C}_p^\times$ as a subset via the embedding defined by the continuous homomorphism

$$(88) \quad (k, \chi) : \mathbf{Z}_p^\times \longrightarrow \mathbf{C}_p^\times, \quad a \longmapsto \chi(a)a^{k-1}.$$

where χ is now thought of as a character of \mathbf{Z}_p^\times by composing with reduction modulo p^n . The subset of weight characters for which κ induces the trivial character on $(\mathbf{Z}/p \mathbf{Z})^\times$ is denoted by \mathcal{W}_0 .

The coefficients of Eisenstein series are naturally functions of (k, χ) , and one can easily show that they extend to functions of \mathcal{W} . The only part that needs clarification is how to view the Kubota–Leopoldt zeta function ζ_p as a function of $\kappa \in \mathcal{W}$. Denote Δ for the torsion subgroup of \mathbf{Z}_p^\times , which is cyclic of order $\phi(q)$, where $q = 4$ if $p = 2$, and $q = p$ otherwise. There is an isomorphism

$$(89) \quad \mathbf{Z}_p^\times \xrightarrow{\sim} \Delta \times (1 + q \mathbf{Z}_p), \quad a \longmapsto (\omega(a), \langle a \rangle).$$

The character ω is called the *Teichmüller character*. Let $\Lambda = \mathbf{Z}_p[[\mathbf{Z}_p^\times]]$ be the Iwasawa algebra, which is the ring of functions on \mathcal{W} , then we have an isomorphism

$$(90) \quad \Lambda \simeq \mathbf{Z}_p[[\Delta]][[T]], \quad 1 + q \longmapsto 1 + T.$$

This way, the Kubota–Leopoldt zeta function ζ_p can be viewed as a function on \mathcal{W} , satisfying

$$(91) \quad \zeta_p((1+q)^{k-1} - 1) = (1 - p^{k-1})\zeta(1-k),$$

giving us the Eisenstein family

$$(92) \quad \begin{aligned} \mathbf{G}_\kappa(q) &= \frac{\zeta_p(\kappa)}{2} + \sum_{n \geq 1} \left(\sum_{p \nmid d|n} \kappa(d)/d \right) q^n & \kappa \notin \mathcal{W}_0 \\ \mathbf{E}_\kappa(q) &= 1 + \frac{2}{\zeta_p(\kappa)} \sum_{n \geq 1} \left(\sum_{p \nmid d|n} \kappa(d)/d \right) q^n & \kappa \in \mathcal{W}_0 \end{aligned}$$

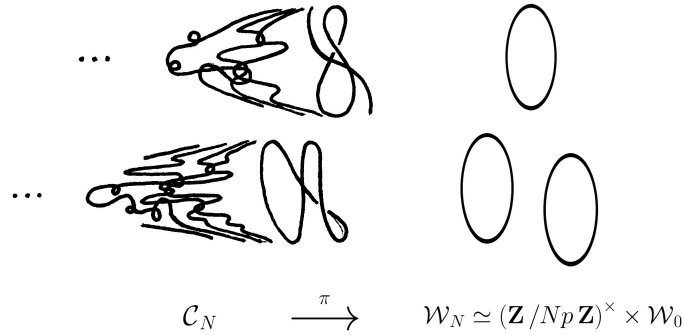
The idea of Coleman was to define an overconvergent modular form of weight κ to be any q -expansion with the property that its quotient by the Eisenstein series of weight κ is an overconvergent modular function⁷. The weights of overconvergent modular forms are naturally parametrised by a geometric object: We define \mathcal{W}_N , the weight space of level N , as a rigid analytic variety, via

$$(93) \quad \mathcal{W}_N = (\mathrm{Spf} \Lambda_N)^{\mathrm{rig}}, \quad \text{where } \Lambda_N = \mathbf{Z}_p[[\mathbf{Z}/N\mathbf{Z}]^\times \times \mathbf{Z}_p^\times].$$

This set of ideas culminated in the construction, due to Coleman–Mazur [CM98], of the *eigencurve* \mathcal{C}_N .

Theorem 3.4 (Coleman–Mazur). *There exists a rigid analytic curve $\mathcal{C}_N \rightarrow \mathcal{W}_N$, whose \mathbf{C}_p -points classify normalised overconvergent eigenforms f which are not in the kernel⁸ of U_p .*

The map $\pi : \mathcal{C}_N \rightarrow \mathcal{W}_N$ simply associates to every overconvergent eigenform f its weight character κ . The geometric properties of \mathcal{C}_N therefore dictate all the possible p -adic variations of modular forms of finite slope in families. Relatively little is known about its geometry. The following picture is a free impression that attempts to depict some of its features. The weight space \mathcal{W}_N decomposes as a finite union of open disks, whereas \mathcal{C}_N contains a particularly well-behaved subspace $\mathcal{C}_N^{\mathrm{ord}}$ that is finite flat over every component of \mathcal{W}_N (to be discussed in § 3.7) and otherwise exhibits a striking contrast between its behaviour close to the boundary, and deeper in the interior (these will be discussed in § 4.2).



We briefly mention here one important property that has been established recently, and is often referred to as the “properness” of the eigencurve. More precisely, it was asked by Coleman–Mazur [CM98] whether

⁷Since then, a more satisfactory – albeit somewhat less immediately suited for explicit computations – definition has been given by Pilloni [Pil13] and Andreatta–Iovita–Pilloni [AIP18], who gave a geometric construction of line bundles ω_κ on the affinoids X_r for some r that depends on κ . He shows that the Eisenstein series of weight κ is a section of his line bundle, therefore giving a completely geometric definition of the space of r -overconvergent forms $M_\kappa^\dagger(r)$ for any weight-character κ , as long as r is sufficiently small.

⁸In this case, we say f is of *finite slope*, where the ‘slope’ refers to the valuation of its U_p -eigenvalue.

the eigencurve can have any ‘holes’, in the sense of a p -adic analytic family of overconvergent eigenforms of finite slope parameterized by a punctured disc, which converges at the puncture to an overconvergent eigenform in the kernel of U_p (such a form is typically said to have *infinite slope*). This is reminiscent of the valuative criterion for properness. It was proved by Buzzard–Calegari [BC06] that no such families exist (and hence the eigencurve is “proper”) when $p = 2$ and $N = 1$, and then by Calegari [Cal08] at integer weights. The general case was established via an intricate, yet elegant, argument by Diaio–Liu [DL16].

3.7. Hida theory. One part of the eigencurve that is fairly well understood is the *ordinary part*, whose discovery by Hida [Hid86b, Hid86a] predates that of the eigencurve by over a decade. An overconvergent form is called *ordinary* if it is a U_p -eigenvector with an eigenvalue that is a p -adic unit, or, said differently, is of slope zero. Hida considered the *ordinary projection* operator

$$(94) \quad e^{\text{ord}} = \lim_{n \rightarrow \infty} U_p^{n!}$$

whose limit exists as an operator on $M_\kappa^\dagger(r)$ for any κ in \mathcal{W}_N . Then Hida showed:

Theorem 3.5 (Hida). *The image of e^{ord} on $M_\kappa^\dagger(r)$ is a finite-dimensional vector space, whose dimension depends only on the connected component of \mathcal{W}_N containing κ .*

This spectacular result shows that even though the slopes of the spectrum of U_p can vary wildly, the dimension of the part of slope 0 is locally constant on \mathcal{W}_N . Note that the connected components of \mathcal{W}_N are indexed by the characters $(\mathbf{Z}/Nq\mathbf{Z})^\times \rightarrow \mathbf{C}_p^\times$, and the dimension of the ordinary subspace is constant over each component. Hida in fact proved the following statement: Suppose

$$(95) \quad \pi^{\text{ord}} : \mathcal{C}_N^{\text{ord}} \rightarrow \mathcal{W}_N$$

is the projection map from the ordinary part of the eigencurve to weight space, then π^{ord} is finite flat. The ordinary part of $\mathcal{C}_N^{\text{ord}}$ is often referred to, though usually only locally, as the *Hida family*.

An extremely powerful tool is the fact that specialisations of Hida families at classical weights $k \geq 2$ are always *classical* modular forms. More generally, the following theorem was proved by Coleman [Col96].

Theorem 3.6 (Coleman). *Suppose that $k \geq 2$ is an integer weight, and $f \in M_k^\dagger$ is a U_p -eigenform of slope strictly less than $k - 1$. Then f is classical, in the sense that it belongs to the finite-dimensional subspace*

$$(96) \quad M_k(\Gamma_0(Np)) \subset M_k^\dagger(\Gamma_0(N)).$$

It is difficult to overstate the importance of this powerful result, which often goes by the name of the *Coleman classicality theorem*. In the literature on overconvergent modular symbols, it is also commonly referred to as *Coleman’s control theorem*. It has far-reaching implications, and is used so frequently in the literature – as well as what follows – that it is often applied without explicit mention.

3.8. Leopoldt’s formula. We end this section with an application of this geometric viewpoint on p -adic modular forms, and prove a classical result on the value at $s = 1$ of p -adic L-functions attached to Dirichlet characters via an incarnation of Serre’s idea to investigate the constant coefficient via the higher Fourier coefficients. In this situation, it allows us to identify the L-value as an explicit combination of units. We follow the treatment in [BCD⁺], which contains several more appearances of Serre’s idea in various guises. A different proof for Leopoldt’s formula for $L_p(1, \chi)$ can be found in [Was97, §5.4].



Suppose that $\chi : (\mathbf{Z}/N\mathbf{Z})^\times \rightarrow \mathbf{C}^\times$ is a primitive even Dirichlet character with conductor $N > 1$ coprime to p , then we have the p -adic Eisenstein family of overconvergent forms

$$(97) \quad \mathbf{E}_k^{(p)}(\chi) = L_p(1-k, \chi) + 2 \sum_{n \geq 1} \sigma_{k, \chi}^{(p)}(n) q^n, \quad \text{where } \sigma_{k, \chi}^{(p)}(n) = \sum_{p \nmid d \mid n} \chi(d) d^{k-1}.$$

This family specialises at $k = 0$ to a rigid analytic function on $X^{\text{ord}} = X_1(N)^{\text{ord}}$, whose value at the cusp ∞ is the value $L_p(1, \chi)$. Now choose a primitive N -th root of unity ζ , then there is a collection of *Siegel units* $g_a \in \mathcal{O}_{Y_1(N)}^\times$ whose q -expansions are given by

$$(98) \quad g_a(q) = q^{1/12} (1 - \zeta^a) \prod_{n \geq 1} (1 - q^n \zeta^a) (1 - q^n \zeta^{-a}), \quad 1 \leq a \leq N-1.$$

Using the operator V_p on p -adic modular forms, we define the rigid analytic function

$$(99) \quad F_\chi^{(p)} = \frac{1}{p\mathfrak{g}(\chi^{-1})} \sum_{a=1}^{N-1} \chi^{-1}(a) \log_p(V_p(g_{pa})g_a^{-1})$$

which is defined on the ordinary locus X^{ord} . Here \mathfrak{g} denotes the standard Gauß sum, obtained by summing $\chi^{-1}(a)\zeta^a$ over a . A direct computation using the expression (98) shows that the higher coefficients of its q -expansion agree with that of $\mathbf{E}_0^{(p)}(\chi)$. Therefore the modular form

$$(100) \quad \mathbf{E}_0^{(p)}(\chi) - F_\chi^{(p)},$$

which is a constant function, must be equal to zero, since it has nebentype χ . We conclude that the constant terms of both series are equal, yielding Leopoldt's formula:

$$(101) \quad L_p(1, \chi) = -\frac{(1 - \chi(p)p^{-1})}{\mathfrak{g}(\chi^{-1})} \sum_{a=1}^{N-1} \chi^{-1}(a) \log_p(1 - \zeta^a).$$

4. EXPLICIT COMPUTATIONS AND ARITHMETIC APPLICATIONS

We now discuss how to compute explicitly with overconvergent modular forms, in more generality than was achieved in the extended example § 3.4, following the approach of Lauder [Lau11]. We then look at a number of different arithmetic applications of this theory, illustrated with explicit examples.

4.1. Computing overconvergent forms. We first explain how to compute explicit bases for the p -adic Banach spaces of r -overconvergent forms, following Katz [Kat73] and Lauder [Lau11]. Note that in the explicit example treated in § 3.4, where $(p, N) = (2, 1)$ and $k = 0$, we were particularly lucky in the sense that the modular curve $X_0(2)$ had genus zero, and the overconvergent regions X_r were isomorphic to a rigid analytic disk, for which we could identify an explicit parameter. This procedure can be repeated for any prime p for which $X_0(p)$ has genus zero (i.e. for $p = 2, 3, 5, 7, 13$), where one can likewise write down a power basis for the space of overconvergent modular forms, for any weight k . See Loeffler [Loe07] for a detailed discussion of this case, as well as many interesting results and computations.

For general values of p , we are faced with a more complicated geometric picture, as the overconvergent regions X_r are isomorphic to the complement of a finite number of disks in \mathbf{P}^1 :



Moreover, in cases where we also have a nontrivial tame level N , the modular curve from which we remove these finitely many disks is no longer isomorphic to \mathbf{P}^1 . Therefore, finding an explicit basis for the set of sections over the overconvergent regions X_r becomes significantly more subtle. In his foundational paper on the subject, Katz [Kat73, Chapter 2] identifies an explicit basis for these spaces, such that any overconvergent form may be written as a unique linear combination of it, referred to as its *Katz expansion*.

Let \mathcal{X} be the modular curve over \mathbf{Z}_p with $\Gamma_1(N)$ -level structure⁹ for $p \nmid N \geq 5$. Let n be the smallest power of p such that the n -th power of the Hasse invariant A^n lifts to a level 1 Eisenstein series E of weight $k_E = n(p-1)$. Throughout this section, we assume $nr \leq 1$. Our notation is summarised by:

| p | 2 | 3 | ≥ 5 |
|-------|-------|-------|-----------|
| E | E_4 | E_6 | E_{p-1} |
| n | 4 | 3 | 1 |
| k_E | 4 | 6 | $p-1$ |

We now describe an explicit basis for the spaces $M_k^\dagger(r)$. Suppose $r = v_p(s)$ for some $s \in \mathbf{C}_p$, then let \mathcal{I}_r be the sheaf of ideals in $\mathrm{Sym}(\omega^{\otimes k_E})$ generated by $E - s^n$, and define the line bundle

$$(102) \quad \mathcal{L} = \mathrm{Spec}_{\mathcal{X}}(\mathrm{Sym}(\omega^{\otimes k_E})/\mathcal{I}_r) \xrightarrow{\pi_{\mathcal{L}}} \mathcal{X}.$$

Assuming that $k \neq 1$, we can apply the base change theorems from [Kat73, Theorem 1.7.1] to show that

$$(103) \quad M_k^\dagger(r) = H^0(\mathcal{L}^{\mathrm{rig}}, \pi_{\mathcal{L}}^* \omega^{\otimes k})$$

$$(104) \quad = H^0(\mathcal{X}, \omega^{\otimes k} \otimes \mathrm{Sym}(\omega^{\otimes k_E})) / H^0(\mathcal{X}, \mathcal{I}_r).$$

Having this concrete description in hand, we now attempt to eliminate the relation $E = s^n$ by investigating the map given by multiplication by E on modular forms as in [Kat73, Lemma 2.6.1] and [Von15, Lemma 1]. More precisely, the injection given by the multiplication by E -map

$$(105) \quad - \times E : H^0(\mathcal{X}, \omega^{\otimes k}) \longrightarrow H^0(\mathcal{X}, \omega^{\otimes k+k_E})$$

splits as a map of \mathbf{Z}_p -modules. This implies that for every $i \geq 0$, we may choose generators $\{a_{i,j}\}_j$ for a complement of the submodule

$$(106) \quad \mathrm{Im}(- \times E) \subseteq H^0(\mathcal{X}, \omega^{\otimes k+ik_E}).$$

This choice is not canonical, but we will fix it once and for all in what follows. As in [Kat73, Proposition 2.6.2], one obtains the following as a consequence of (104) and the splitting of (105).

⁹In practice, there is a lot of flexibility with the setup, and the computations below are usually for $\Gamma_0(N)$ instead of $\Gamma_1(N)$. To justify this, some additional analysis is required to deal with the lack of representability, see [BC05, Appendix].

Theorem 4.1. *The set $\{e_{i,j}\}_{i,j}$ is an orthonormal basis for the p -adic Banach space $M_k^\dagger(r)$, where*

$$(107) \quad e_{i,j} = s^{ni} \frac{a_{i,j}}{E^i}$$

Note that we have avoided the case $k = 1$, which we can still compute with by appropriately twisting by U_p , thereby reducing the computation to one in higher weight for which the results above hold. This technique is often referred to as *Coleman's trick*, see [Col97b, Eqn. (3.3)], and is also frequently useful in other situations. It is based on the observation that multiplication by E^j defines an isomorphism

$$(108) \quad M_k^\dagger(r) \longrightarrow M_{k+jk_E}^\dagger(r),$$

as well as the fact that the U_p -operator is *Frobenius linear* in the sense that

$$(109) \quad U_p(fV_p(E)) = U_p(f)E.$$

It follows from these two simple facts that $P_{k+jk_E}(t)$ equals the characteristic series of $U_p \circ G^j$ on $M_k^\dagger(r)$, where we denote $G = E/V_p E$. This allows us to flexibly change the weights of the spaces of overconvergent forms we are interested in. In particular, we can compute overconvergent forms in weight 1 by reducing the computation to, say, weight p . Likewise, if we would like to compute the operator U_p on $M_k^\dagger(r)$ for some extremely large weight k , we can use Coleman's trick to reduce the computation to a small weight.

Now that we know, by Theorem 4.1, an explicit basis $e_{i,j}$ for the Banach space $M_k^\dagger(r)$, we are in a position to compute approximations of the matrix of U_p on q -expansions. Since we can only compute finitely many of its entries, we need a good estimate on the valuations of its entries, so we know how many elements of the basis we need to compute before we are guaranteed that the end result is correct up to some chosen p -adic precision. To do this, let us first fix some notation for these entries. We write

$$(110) \quad U_p \circ G^j(e_{u,v}) = \sum_{w,z} A_{u,v}^{w,z}(j) e_{w,z},$$

for some $A_{u,v}^{w,z}(j) \in \mathbf{C}_p$. Said differently, the numbers $A_{u,v}^{w,z}(j)$ are the entries of the infinite matrix of $U_p \circ G^j$ with respect to our chosen orthonormal basis for $M_k^\dagger(r)$. The following lemma estimates their p -adic valuations, and is an easy extension of Wan [Wan98, Lemma 3.1], see [Von15].

Lemma 4.2. *We have*

$$(111) \quad v_p(A_{u,v}^{w,z}(j)) \geq wrk_E - 1 - r(n-1).$$

The reader may have wondered why in the above precision estimate, we included the parameter j , corresponding to a twist of the U_p operator by $G^j = (E/V_p E)^j$, rather than simply putting $j = 0$. The reason is that this allows us to easily move between different weights, and perform the computation of U_p in several weights at once. The examples below illustrate this, by computing the U_p -operator in families.

Remark. In what follows, we frequently drop the rate of overconvergence r from the notation. This is justified by the fact that an overconvergent finite slope eigenform must be r -overconvergent for *any* $r < p/(p+1)$. The data below therefore does not depend on r at all, though its computation does. It is clear from (111) that it is helpful in practice to choose r as large as possible, to accelerate convergence.

4.2. The spectral curve. We now have two crucial active ingredients for a working algorithm to compute with spaces of overconvergent modular forms, since we have (a) an explicit basis due to Katz, provided by Theorem 4.1, and (b) a precision estimate for the concomitant entries of the matrix of U_p due to Wan, provided by (111). Lauder [Lau11] combines these two ingredients into an efficient algorithm for computing

U_p on $M_k^\dagger(r)$. We note that the estimate (111) is independent of j , and hence the computation may be performed at several p -adic weights at once. In this example, we compute the resulting 2-variable series $P(\kappa, t)$. The curve in $\mathcal{W}_N \times \mathbf{G}_m$ cut out by this equation is often referred to as the *spectral curve* of U_p , which yields the eigencurve after an additional modification, see [CM98].

Let $f : \mathcal{W}_N \rightarrow \mathbf{C}_p$ be a function in the Iwasawa algebra, and $\{\kappa_0, \kappa_1, \dots, \kappa_n\}$ a finite set of points. Then we denote $f[\kappa_0] = f(\kappa_0)$ and we inductively define the *divided difference* of order n to be

$$f[\kappa_0, \kappa_1, \dots, \kappa_n] := \frac{f[\kappa_1, \dots, \kappa_n] - f[\kappa_0, \dots, \kappa_{n-1}]}{\kappa_n - \kappa_0}.$$

We now define the n -th Newton series to be

$$(112) \quad P_n(\kappa, t) = \sum_{i=0}^n P[\kappa_0, \kappa_1, \dots, \kappa_i](t) \times (\kappa - \kappa_0)(\kappa - \kappa_1) \cdots (\kappa - \kappa_i),$$

where $P[\kappa_0, \dots, \kappa_n](t)$ is the power series in t obtained by taking the corresponding finite differences on the coefficients of $P(\kappa, t)$ of t , which are elements of the Iwasawa algebra by Coleman [Col97a]. The theory of finite differences then shows that upon increasing the number of interpolation points, the n -th Newton series p -adically approaches the series $P(\kappa, t)$. This means that all we need to do to compute an approximation for $P(\kappa, t)$, is to choose our interpolation points carefully and estimate the error term.

We explicitly compute some examples, starting by revisiting the example of Buzzard–Calegari [BC05] familiar from § 3.4, and then venturing into more unfamiliar territory relating to situations that were considered in the literature by Buzzard–Kilford [BK05], Roe [Roe14] and the work on boundary slopes and the spectral halo by Andreatta–Iovita–Pilloni [AIP18] and Bergdall–Pollack [BP16]. We note that an alternative approach using overconvergent modular symbols has been developed in [DHH⁺16], for Hida families. Their algorithms yield explicit q -expansions of Hida families, where the coefficients are elements of Λ , but is only equipped to handle the ordinary part of the spectrum of U_p .

Example 1. We revisit the case of $p = 2$ and tame level $N = 1$, where we computed with the space for $k = 0$ in § 3.4. Using an interpolation as described above, we can compute the two variable power series $P(\kappa, t)$, whose specialisation at $\kappa \in \mathcal{W}$ recovers the characteristic series $P_\kappa(t)$ of U_2 on the space of overconvergent modular forms $M_\kappa^\dagger(r)$. We obtain

$$\begin{aligned} P(\kappa, t) = & 1 + (519736167t + 413685912t^2 + 148708352t^3 + 1065353216t^4) \\ & + \kappa (36306799t + 374998993t^2 + 380696768t^3 + 281739264t^4) \\ & + \kappa^2 (43984100t + 481404364t^2 + 496002384t^3 + 387895296t^4 + 1811939328t^5) \\ & + \kappa^3 (874017364t + 890496879t^2 + 487943741t^3 + 4077568t^4 + 964689920t^5) \\ & + \kappa^4 (392124398t + 264203079t^2 + 839291211t^3 + 908503936t^4 + 817102848t^5) \\ & + O(\kappa^5, 2^{30}), \end{aligned}$$

We actually computed $P(\kappa, t)$ to precision $O(\kappa^{25}, 2^{70})$, which took about 5 minutes, but truncated the result to get output that fits in this document. Let us now investigate various specialisations:

- The computation we did in the previous section is contained in this one, and if we set $\kappa = 5^k - 1 = 0$, which corresponds to weight $k = 0$, we recover the same power series as before, up to the used precision. In particular, we can read off that first few slopes are $\mathbf{0}_1, \mathbf{3}_1, \mathbf{7}_1, \dots$, which agrees with the result of Buzzard–Calegari [BC05] that in weight 0 the n -th slope is equal to

$$1 + 2v_2 \left(\frac{(3n)!}{n!} \right).$$

- As for the other extreme, the main result of Buzzard–Kilford [BK05] states that the slopes on the boundary annulus $1/8 < |\kappa| < 1$ form an arithmetic progression with n -th term $nv_2(\kappa)$, all with multiplicity 1. Indeed, by substituting $\kappa = 2$ we obtain the slope sequence $0, 1, 2, 3, 4, \dots$, while for $\kappa = 4$ we recover $0, 2, 4, 6, 8, \dots$. Our computed power series $P(\kappa, t)$ hence combines the best of both worlds, by describing the spectral curve over the inner regions of \mathcal{W} as well as the outskirts. Notice the striking contrast between the nature of the slope sequence at $k = 0$ and that close to the boundary! A folklore conjecture predicts that the same phenomenon happens in general, and a result of this flavour was obtained by Liu–Wan–Xiao [LWX17].

In the above computation, we focussed on the variation of $P_\kappa(t)$ with the weight κ , but we can interchange the variables κ and t and study instead the powers series in κ appearing as the coefficients of the above series in t . For instance, up to precision $(2^{21}, \kappa^7)$ we obtain

$$\begin{aligned}
P(\kappa, t) \equiv & 1 + t(1739623 + 655215\kappa + 2041060\kappa^2 + 1602132\kappa^3 + 2054126\kappa^4 + 779022\kappa^5 + 1634724\kappa^6) \\
& + t^2(546968 + 1705937\kappa + 1156556\kappa^2 + 1304431\kappa^3 + 2059079\kappa^4 + 1677821\kappa^5 + 644339\kappa^6) \\
& + t^3(1907712 + 1112256\kappa + 1074512\kappa^2 + 1404477\kappa^3 + 430411\kappa^4 + 51909\kappa^5 + 1261732\kappa^6) \\
& + t^4(720896\kappa + 2019328\kappa^2 + 1980416\kappa^3 + 437120\kappa^4 + 1161264\kappa^5 + 1648837\kappa^6) \\
& + t^5(1310720\kappa^4 + 524288\kappa^5 + 1101824\kappa^6) \\
& + O(2^{21}, \kappa^7)
\end{aligned}$$

Investigating the coefficients $a_i(\kappa)$ of $P(\kappa, t)$ for small values, we see that their valuation on $\kappa \in \mathbf{Z}_2$ only seems to depend on $\kappa \pmod{2^6}$. This can be made into a rigorous proof of this fact, by using the uniform estimates in Wan [Wan98] for the Newton polygon in t of $P(\kappa, t)$ recalled above. After possibly redoing the computation to a higher precision, to assure that all the slopes are indeed correct, we recover the following theorem, which may be found in Emerton [Eme98, Theorem 1.1].

Theorem 4.3 (Emerton). *The minimal non-zero slope of U_2 on M_k^\dagger in tame level 1, along with its multiplicity, depends only on $k \pmod{16}$. More precisely, it is given by*

$$\begin{array}{ll}
\mathbf{3}_1 & \text{if } k \equiv 0 \pmod{4}, \\
\mathbf{4}_1 & \text{if } k \equiv 2 \pmod{8}, \\
\mathbf{5}_1 & \text{if } k \equiv 6 \pmod{16}, \\
\mathbf{6}_2 & \text{if } k \equiv 14 \pmod{16}.
\end{array}$$

We note that the calculations of Emerton [Eme98] rely crucially on the explicit uniformisations of 2-adic regions on the genus 0 modular curves $X_0(2^n)$ for small values of n , which are hard to come by in higher levels and primes. Our algorithms do not rely on any specifics of the situation $(p, N) = (2, 1)$, and therefore similar arguments work in more general settings.

Looking further into the above coefficients, let $\lambda(i)$ be the number of roots of $a_i(\kappa)$ in the open unit disk. The following table displays the 2-adic valuations of these roots, along with their multiplicities:

| Coefficient | Valuations | λ |
|-------------------|---|-----------|
| $a_0(\kappa) = 1$ | \emptyset | 0 |
| $a_1(\kappa)$ | \emptyset | 0 |
| $a_2(\kappa)$ | $\mathbf{3}_1$ | 1 |
| $a_3(\kappa)$ | $\mathbf{3}_2, \mathbf{4}_1$ | 3 |
| $a_4(\kappa)$ | $\mathbf{3}_4, \mathbf{4}_1, \mathbf{7}_1$ | 6 |
| $a_5(\kappa)$ | $\mathbf{3}_6, \mathbf{4}_2, \mathbf{5}_1, \mathbf{7}_1$ | 10 |
| $a_6(\kappa)$ | $\mathbf{3}_9, \mathbf{4}_3, \mathbf{5}_2, \mathbf{6}_1$ | 15 |
| $a_7(\kappa)$ | $\mathbf{3}_{12}, \mathbf{4}_5, \mathbf{5}_2, \mathbf{6}_1, \mathbf{8}_1$ | 21 |

By inspecting the 2-adic valuations of the coefficients we computed, we see that this output is provably correct and complete. Note that

$$\lambda(i) = \binom{i}{2},$$

which also follows from the main result of Buzzard–Kilford [BK05]. In Bergdall–Pollack [BP16], precise conjectures are made about the location of the zeroes of a_i .

Example 2. Let us set $(p, N) = (3, 1)$ and compute $P(\kappa, t)$ up to precision $O(3^{90}, \kappa^{60})$. With the same notation as above, we find¹⁰ the following slopes of the zeroes of the coefficients $a_i(\kappa)$:

| Coefficient | Valuations | λ |
|-------------------|--|-----------|
| $a_0(\kappa) = 1$ | \emptyset | 0 |
| $a_1(\kappa)$ | \emptyset | 0 |
| $a_2(\kappa)$ | $\mathbf{1}_2$ | 2 |
| $a_3(\kappa)$ | $\mathbf{1}_5, \mathbf{3}_1$ | 6 |
| $a_4(\kappa)$ | $\mathbf{1}_9, \mathbf{2}_2, \mathbf{3}_1$ | 12 |
| $a_5(\kappa)$ | $\mathbf{1}_{15}, \mathbf{2}_4, \mathbf{3}_1$ | 20 |
| $a_6(\kappa)$ | $\mathbf{1}_{22}, \mathbf{2}_5, \mathbf{3}_2, \mathbf{4}_1$ | 30 |
| $a_7(\kappa)$ | $\mathbf{1}_{30}, \mathbf{2}_8, \mathbf{3}_2, \mathbf{4}_2$ | 42 |
| $a_8(\kappa)$ | $\mathbf{1}_{40}, \mathbf{2}_{11}, \mathbf{3}_2, \mathbf{4}_3$ | 56 |

Again, this output is complete and provably correct. Notice that

$$\lambda(i) = 2 \binom{i}{2},$$

which follows from the main result of Roe [Roe14], who showed that near the boundary, the slopes form an arithmetic progression with an explicit argument that depends on the valuation of κ . Roe tackled this more complicated situation using the same techniques as Buzzard–Kilford [BK05].

Example 3. We now turn to $(p, N) = (2, 3)$ and compute $P(\kappa, t)$ up to precision $O(2^{60}, \kappa^{20})$. This computation took about 90 minutes on a standard laptop. In addition to the notation above, let $\mu(i)$ be the largest power of p that divides $a_i(\kappa)$. The work of Bergdall–Pollack [BP16] uses Koike’s trace formula to prove that $\mu(i) = 0$ whenever $N = 1$. However, in our situation μ appears to be larger for several i :

¹⁰The motivated reader can try to recover this computation, for instance using an explicit basis similar to that used in § 3.4, which is possible since $X_0(3)$ has genus 0. There is a particularly nice basis, described by Loeffler [Loe07], which can be twisted by an Eisenstein series to obtain the computation in all weights.

| Coefficient | Valuations | λ | μ |
|-------------------|--|-----------|-------|
| $a_0(\kappa) = 1$ | \emptyset | 0 | 0 |
| $a_1(\kappa)$ | — | — | — |
| $a_2(\kappa)$ | \emptyset | 0 | 0 |
| $a_3(\kappa)$ | \emptyset | 0 | 1 |
| $a_4(\kappa)$ | $\mathbf{4}_1$ | 1 | 0 |
| $a_5(\kappa)$ | $\mathbf{3}_2$ | 2 | 1 |
| $a_6(\kappa)$ | $\mathbf{3}_2, \mathbf{4}_1$ | 3 | 0 |
| $a_7(\kappa)$ | $\mathbf{3}_2, \mathbf{4}_1, \mathbf{8}_1$ | 4 | 1 |
| $a_8(\kappa)$ | $\mathbf{3}_3, \mathbf{4}_1, \mathbf{5}_1, \mathbf{6}_1$ | 6 | 0 |
| $a_9(\kappa)$ | $\mathbf{3}_4, \mathbf{4}_3, \mathbf{6}_1$ | 8 | 1 |
| $a_{10}(\kappa)$ | $\mathbf{3}_5, \mathbf{4}_3, \mathbf{5}_1, \mathbf{8}_1$ | 10 | 0 |
| $a_{11}(\kappa)$ | $\mathbf{3}_6, \mathbf{4}_4, \mathbf{5}_2$ | 12 | 1 |
| $a_{12}(\kappa)$ | $\mathbf{3}_7, \mathbf{4}_5, \mathbf{5}_2, \mathbf{7}_1$ | 15 | 0 |

Computing $P(\kappa, t)$ up to precision $O(2, \kappa^{30})$ takes about one minute. Extracting the degrees of the t -coefficients, our data suggests the boundary slope sequence

$$\mathbf{0}_2, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_2, \mathbf{3}/\mathbf{2}_2, \mathbf{2}_2, \mathbf{5}/\mathbf{2}_2, \mathbf{3}_2, \mathbf{7}/\mathbf{2}_2, \dots$$

which is indeed in accordance with the Newton polygon of $\lambda + \mu$ computed above, up to the chosen precisions. Notice the similarity with the slope sequence for $(p, N) = (2, 1)$.

Example 4. As above, set $(p, N) = (11, 1)$ and compute $P(\kappa, t)$ up to precision $O(11, \kappa^{60})$, which takes about two minutes. We compute the degrees of the t -coefficients, which suggest the boundary slope sequence:

$$\mathbf{0}_1, \mathbf{1}_1, \mathbf{2}_1, \mathbf{3}_1, \mathbf{4}_2, \mathbf{5}_1, \mathbf{6}_1, \mathbf{7}_1, \mathbf{9}_2, \dots$$

4.3. The Gouvêa–Mazur conjecture. An enormous amount of arithmetic information is encoded in the *slopes* of overconvergent modular forms, which are the valuations of their U_p -eigenvalues. One of the consequences of the theory of Coleman [Col97b] is that for any $\alpha > 0$, there exists a smallest integer N_α with the following property: If k_1 and k_2 are integers such that

$$(113) \quad k_1 \equiv k_2 \pmod{p^{N_\alpha}(p-1)}$$

then the collection of slopes $\leq \alpha$ in weights k_1 and k_2 agree, with multiplicities. Gouvêa and Mazur conjectured in [GM92] that $N_\alpha \leq \lfloor \alpha \rfloor$. However, Wan [Wan98] exhibits an explicit quadratic upper bound for N_α , depending on p and the level.¹¹

The key observation for Wan is that the lower bound (111) is independent of j . After taking determinants, we obtain a lower bound on the coefficients of the characteristic series of U_p in weight $k + jk_E$, again independent of j . Wan then proceeds by proving a very general reciprocity lemma on Newton polygons, which allows him to transform the lower bound for those coefficients into an upper bound for N_α .

Theorem 4.4. *There is an explicitly computable quadratic polynomial $P \in \mathbf{Q}[x]$, depending only on p and the level, such that $N_\alpha \leq P(\alpha)$.*

Since Gouvêa and Mazur conjectured in [GM92] that $N_\alpha \leq \lfloor \alpha \rfloor$, this is still an order of magnitude from what we expect. However, the conjecture of Gouvêa–Mazur is known to be false, and a counterexample

¹¹Strictly speaking, Wan assumes that $p \geq 5$, but his arguments easily extend to $p = 2, 3$ using our basis described above.

was given in [BC04]. It should be noted that the counterexample of Buzzard–Calegari is only a very small violation of the conjecture, and generically it seems that in fact something much stronger than Gouvêa–Mazur is true! Let us illustrate this with two examples.

The case $p = 2$ is prolific soil for finding counterexamples to the Gouvêa–Mazur conjecture. As noted above, the first counterexample was given in [BC04] for $p = 59$ and level 1, and a further one for $p = 79$ in [Lau11]. For $p = 2$, we obtain the following slope sequences in level $\Gamma_0(19)$:

$$\begin{aligned}
 k = -2 : & \quad \mathbf{0}_4, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_3, \mathbf{2}_5, \mathbf{9}/\mathbf{4}_4, \mathbf{4}_3, \mathbf{5}_2, \mathbf{6}_{21}, \mathbf{15}/\mathbf{2}_2, \dots \\
 k = 0 : & \quad \mathbf{0}_4, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_5, \mathbf{3}_{11}, \mathbf{13}/\mathbf{4}_4, \mathbf{7}_{25}, \mathbf{25}/\mathbf{2}_4, \mathbf{13}_{11}, \dots \\
 k = 2 : & \quad \mathbf{0}_4, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_3, \mathbf{3}/\mathbf{2}_2, \mathbf{2}_5, \mathbf{4}_{11}, \mathbf{17}/\mathbf{4}_4, \mathbf{8}_{25}, \mathbf{27}/\mathbf{2}_4, \dots \\
 k = 4 : & \quad \mathbf{0}_4, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_5, \mathbf{5}/\mathbf{2}_2, \mathbf{3}_6, \mathbf{7}/\mathbf{2}_2, \mathbf{4}_3, \mathbf{5}_5, \mathbf{21}/\mathbf{4}_4, \dots \\
 k = 6 : & \quad \mathbf{0}_4, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_3, \mathbf{2}_7, \mathbf{5}/\mathbf{2}_2, \mathbf{4}_3, \mathbf{9}/\mathbf{2}_2, \mathbf{5}_6, \mathbf{11}/\mathbf{2}_2, \dots \\
 k = 8 : & \quad \mathbf{0}_4, \mathbf{1}/\mathbf{2}_2, \mathbf{1}_5, \mathbf{3}_{13}, \mathbf{7}/\mathbf{2}_2, \mathbf{6}_5, \mathbf{13}/\mathbf{2}_2, \mathbf{7}_6, \mathbf{15}/\mathbf{2}_2, \dots
 \end{aligned}$$

Notice the aberration in the dimensions of the slope 1 subspaces, as well as the slope 3 subspaces in weights 0 and 8. Whereas these are all *near misses*, in that the smallest slopes for which discrepancies arise are exactly equal to the valuation of the weight difference, we note a 2-dimensional slope $3/2$ subspace in weight 2, which is completely absent in weight 6, whereas $3/2 < v_2(6 - 2) = 2$. Similarly, the slope $9/4$ subspace in weight -2 does not exist in weight $6 = -2 + 2^3$.

On the other hand, to see how Gouvêa–Mazur is frequently much weaker than the truth, consider the first few slopes of U_3 acting on $M_{278}^\dagger(\Gamma_0(41))$, which we computed using Lauder’s algorithm to be

$$(114) \quad \mathbf{0}_{12}, \mathbf{1}_{14}, \mathbf{3}_{48}, \mathbf{6}_{14}, \mathbf{7}_{22}, \mathbf{8}_6, \mathbf{9}_{22}, \mathbf{10}_{14}, \mathbf{12}_{48}, \mathbf{14}_{14}, \mathbf{16}_{22}, \mathbf{17}_6, \mathbf{18}_{22}, \dots$$

where the subscripts denote multiplicities. Repeating the same computation in weight 8, we find the exact same slope sequence for all the terms we display here, whereas the Gouvêa–Mazur conjecture would only predict the slopes up to 3 to agree. This behaviour seems rather typical in most examples we computed.

4.4. Chow–Heegner points. We now discuss how the computation of spaces of overconvergent forms, using the above algorithms, can be used to construct arithmeto-geometric invariants. We chose to discuss the Heegner-type point construction on elliptic curves, following Darmon–Rotger [DR14].

Let p be a prime and E/\mathbf{Q} an elliptic curve of conductor N , associated to a p -ordinary form $f \in S_2^{\text{new}}(\Gamma_0(N))$. Let g be any other weight 2 newform which is p -ordinary. It can be deduced from the work of Darmon–Rotger [DR14, Theorem 1.3] that there exists a global (rational) point $P_g \in E(\mathbf{Q})$ that satisfies the Gross–Zagier type formula

$$(115) \quad \log(P_g) = 2d_g \cdot \frac{\mathcal{E}_0(g) \mathcal{E}_1(g)}{\mathcal{E}(g, f, g)} \cdot \mathcal{L}_p(\mathbf{g}, \mathbf{f}, \mathbf{g})(2, 2, 2),$$

where the quantities appearing in the formula are

- \log is the formal p -adic logarithm on the elliptic curve E ,
- d_g is an integer described in [DDL15, Remark 3.1.3],
- the \mathcal{E} -factors are quadratic numbers depending only on the p -th coefficients of f and g ,
- $\mathcal{L}_p(\mathbf{g}, \mathbf{f}, \mathbf{g})$ is the Rankin triple product p -adic L-function of the Hida families \mathbf{f}, \mathbf{g} through f, g .

The last item in this list deserves some discussion. We will not define the Rankin triple product p -adic L-function here, as that would lead us too far from the topic of these notes, and the exposition in Darmon–Rotger [DR14] is excellent. We will however explain how one computes the special value appearing in formula (115). As before, we let $e^{\text{ord}} = \lim_n U_p^{n!}$ be Hida’s ordinary projector. Start by computing

$$(116) \quad e^{\text{ord}}(\theta^{-1} f^{[p]} \times g)$$

where $f^{[p]}$ denotes the p -depletion $(1 - V_p U_p)f$ of f . Here, we have used Serre’s differential operator $\theta = qd/dq$, which is an important object in the theory of overconvergent forms, and would surely merit an entire article to do it justice. The inverse of this operator is defined by the p -adic limit

$$(117) \quad \theta^{-1} = \lim_{n \rightarrow \infty} \theta^{p^n - 1}.$$

By Coleman’s criterion, we conclude that the overconvergent form (116) is classical, and hence it can be written as a finite linear combination of Hecke eigenforms of weight 2 and level $\Gamma_0(p)$. The special value $\mathcal{L}_p(\mathbf{g}, \mathbf{f}, \mathbf{g})(2, 2, 2)$ is the coefficient of g in this linear combination.

Example 1. Consider the elliptic curve

$$(118) \quad E : y^2 + xy = x^3 - x^2 - x + 1$$

which has rank 1 and conductor 58. Consider its associated newform f , and let g be the unique newform on $\Gamma_0(58)$ different from f , then

$$(119) \quad \begin{aligned} f(q) &= q - q^2 - 3q^3 + q^4 - 3q^5 + 3q^6 - 2q^7 - q^8 + 6q^9 + 3q^{10} - q^{11} + \dots \\ g(q) &= q + q^2 - q^3 + q^4 + q^5 - q^6 - 2q^7 + q^8 - 2q^9 + q^{10} - 3q^{11} + \dots \end{aligned}$$

Both f and g are 2-ordinary. Letting $P = (0, 1)$ be a generator for $E(\mathbf{Q})$, we compute that

$$(120) \quad \mathcal{L}_2(\mathbf{g}, \mathbf{f}, \mathbf{g})(2, 2, 2) \equiv 3 \log_E(P) \pmod{2^{200}},$$

as predicted by the theory in [DR14].

Let us end this discussion on a more speculative note. In the above it is important that f is ordinary. Whereas it is conceivable that this may be extended to eigenforms of finite slope through the use of Coleman families, it is not clear that even if the Rankin triple product p -adic L-function may be constructed in cases where f is of infinite slope, that it should be related to global points. Nonetheless, the computation of the special value above yields an explicit number even in those situations, and we now compute a few examples where the Tate module of $E_{\mathbf{Q}}$ is wildly ramified at 2 or 3, and f is of infinite slope.

Example 2a. Consider the elliptic curve

$$(121) \quad E : y^2 + y = x^3 + 9x - 10$$

which is of conductor $4617 = 3^5 \cdot 19$ and rank 1. Consider the newforms

$$(122) \quad \begin{aligned} f(q) &= q - 2q^2 + 2q^4 - 2q^5 - 3q^7 + 4q^{10} - 6q^{11} + \dots \\ g(q) &= q - 2q^3 - 2q^4 + 3q^5 - q^7 + q^9 + 3q^{11} + \dots \end{aligned}$$

where f is associated to E , and g is the unique cuspidal newform of weight 2 on $\Gamma_0(19)$. Despite f being of infinite 3-adic slope, we can run the computation and find a numerical value for $\mathcal{L}_2(\mathbf{g}, “f”, \mathbf{g})(2, 2, 2)$. We find that

$$(123) \quad \mathcal{L}_3(\mathbf{g}, “f”, \mathbf{g})(2, 2, 2) \equiv t \cdot \log_E(P) \pmod{3^{200}} \quad \text{where } 2t^2 + 48t + 729 = 0,$$

where $P = (4, 9)$ is a generator of $E(\mathbf{Q})$. The fact that both quantities are related by a quadratic number t of small height suggests that a more general analogue of the theory for ordinary forms in [DR14], and more specifically equation (115), might exist.

Example 2b. Consider the elliptic curve

$$(124) \quad E : y^2 = x^3 + x^2 - 62893x - 6091893$$

which is of rank 1 and conductor $15104 = 2^8 \cdot 59$. Let f be its associated newform, and let g be the newform of level 118 associated to the elliptic curve with Cremona label 118.a1, then

$$(125) \quad \begin{aligned} f(q) &= q - 2q^3 - 3q^7 + q^9 + 3q^{11} - 3q^{13} + \dots \\ g(q) &= q - q^2 - q^3 + q^4 - 3q^5 + q^6 - q^7 - q^8 - 2q^9 + 3q^{10} - 2q^{11} + \dots \end{aligned}$$

Note that g is 2-ordinary. We compute that

$$(126) \quad \mathcal{L}_2(\mathbf{g}, "f", \mathbf{g})(2, 2, 2) \equiv 6 \log_E(P) \pmod{2^{100}},$$

4.5. p -Adic L-functions of real quadratic fields. We end this article with a discussion of a method to compute p -adic L-functions of totally real fields F , following [LV19]. We closely mirror the approach to p -adic L-functions in § 2 developed in Serre [Ser73] and Deligne–Ribet [DR80], which is rooted in an idea that goes back to Hecke [Hec24] and Siegel [Sie68]. It should be noted that an alternative approach towards p -adic L-functions of Barsky and Cassou-Noguès [Bar78, CN79] based on the explicit formula for zeta values of Shintani [Shi76] was recently used to develop an algorithm for their computation by Roblot [Rob15]. Instead, we take here an approach using diagonal restrictions of Eisenstein series and p -adic interpolation, similar to that of Cohen [Coh76] and Cartier–Roy [CR72].

For simplicity, we restrict to the case where $F = \mathbf{Q}(\sqrt{D})$ is a real quadratic field. Let \mathfrak{d} denote its different ideal. Suppose ψ is a character of F , then Hecke [Hec24] proposed to study the values $L(\psi, 1-k)$ by considering the diagonal restriction of a Hilbert Eisenstein series of weight k over F . This was carried out by Klingen–Siegel [Kli62, Sie68] to show the rationality of such special values, and to give explicit closed formulae for some small values of k . For instance, their methods, which we review shortly, yield classical identities such as

$$(127) \quad \zeta_F(-1) = \frac{-1}{60} \sum_{\substack{b < \sqrt{D} \\ b \equiv D \pmod{2}}} \sigma_1\left(\frac{D-b^2}{4}\right).$$

To explain their arguments, we recall the definition of the Eisenstein series attached to a character ψ of modulus \mathfrak{m} . Suppose $k \geq 1$ is such that ψ has sign $(-1)^k$ at both infinite places. Shimura [Shi78] defines the space $M_k(\mathfrak{m}, \psi)$ of Hilbert modular forms of (parallel) weight k , level \mathfrak{m} and character ψ . We content ourselves by mentioning that the data includes a holomorphic function $f : \mathcal{H}^2 \rightarrow \mathbf{C}$ which satisfies

$$(128) \quad (c_1 z_1 + d_1)^{-k} (c_2 z_2 + d_2)^{-k} f\left(\frac{a_1 z_1 + b_1}{c_1 z_1 + d_1}, \frac{a_2 z_2 + b_2}{c_2 z_2 + d_2}\right) = \psi(\mathfrak{a}) f(z_1, z_2),$$

for all matrices

$$(129) \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathcal{O}_F) \quad \text{such that } c \in \mathfrak{m}.$$

Here, and in what follows, we have used the notation x_i to denote the image of $x \in F$ under the i -th embedding $\sigma_i : F \hookrightarrow \mathbf{R}$. The transformation law (128) implies that every form has a q -expansion, indexed

by the totally positive elements \mathfrak{d}_+^{-1} of the inverse different. The case of interest to us is given by the Eisenstein series

$$(130) \quad \mathbf{G}_{k,k}(\psi) \in M_k(\mathfrak{m}, \psi)$$

whose q -expansion is¹² given by

$$(131) \quad L(\psi, 1-k) + 4 \sum_{\nu \in \mathfrak{d}_+^{-1}} \sum_{\mathfrak{a} | (\nu)\mathfrak{d}} \psi(\mathfrak{a}) \mathrm{Nm}(\mathfrak{a})^{k-1} \exp(2\pi i(\nu_1 z_1 + \nu_2 z_2)).$$

The *diagonal restriction* of $\mathbf{G}_{k,k}(\psi)(z_1, z_2)$ is obtained by setting $z_1 = z_2$, and is a modular form of weight $2k$ and level one, which has q -expansion:

$$(132) \quad L(\psi, 1-k) + 4 \sum_{n \geq 1} \left(\sum_{\substack{\nu \in \mathfrak{d}_+^{-1} \\ \mathrm{Tr}(\nu)=n}} \sum_{\mathfrak{a} | (\nu)\mathfrak{d}} \psi(\mathfrak{a}) \mathrm{Nm}(\mathfrak{a})^{k-1} \right) q^n.$$

Setting $k = 2$, note that $M_4(\mathrm{SL}_2(\mathbf{Z})) = \langle \mathbf{E}_4 \rangle$, so the above form must be a multiple of \mathbf{E}_4 . One quickly determines this multiple, and we deduce the equality (127). For general k , it is more difficult to describe the combination of modular forms we obtain, but we can always choose a basis of classical modular forms with rational Fourier coefficients, of which the diagonal restriction must be a rational linear combination due to the rationality¹³ of its higher Fourier coefficients. It then immediately follows that $L(\psi, 1-k)$ must also be rational. We illustrate this with a simple example.

Example 1. Let us consider $F = \mathbf{Q}(\sqrt{5})$. The narrow ray class group attached to the prime (3) is

$$(133) \quad \mathrm{Cl}_{(3)}^+ \simeq \mathbf{Z}/2\mathbf{Z}$$

and the unique quadratic character ψ of conductor (3) is totally odd. We compute the diagonal restriction in (132) for $k = 3$, and find it has q -expansion given by

$$(134) \quad L(\psi, -2) = 1144q - 39696q^2 - 291448q^3 - 1261696q^4 + \dots$$

This is a modular form of weight 6 and level $\Gamma_1(3)$, and the space $M_6(\Gamma_1(3))$ is 3-dimensional, and has a basis of the form

$$(135) \quad \begin{cases} f_1 = & 1 & & - & 504q^3 & & + & \dots, \\ f_2 = & & q & & + & 45q^3 & + & 166q^4 & + & \dots, \\ f_3 = & & & q^2 & + & 6q^3 & + & 27q^4 & + & \dots \end{cases}$$

Using the first three higher Fourier coefficients, we easily determine that the diagonal restriction is equal to

$$\frac{32}{9} f_1 - 1144 f_2 - 39696 f_3,$$

and we deduce that $L(\psi, -2) = 32/9$.

For general weights and characters, our inability to easily describe the first few Fourier coefficients of a rational basis for the space of modular forms that contains the diagonal restriction is what stood in the way of giving a clean explicit formula of the sort of (127). Computationally, we may easily determine such a basis as in the above example, and obtain such a formula in any given case that merits our consideration. This effectively reduces the computation of the constant term to the efficient computation of the higher

¹²In the case where $k = 1$ and $\mathfrak{m} = (1)$, the constant term of (131) must be modified suitably, and we refer the interested reader to the statements contained in Darmon–Dasgupta–Pollack [DDP11, Proposition 2.11] for more details.

¹³Strictly speaking, here we mean rational over the smallest number field containing the values of ψ .

Fourier coefficients of (132), which are of a much more elementary nature. For real quadratic fields, an efficient algorithm was presented in [LV19] in terms of the theory of reduced cycles of indefinite quadratic forms [Gau01, BV07]. The p -adic variation of the constant term is then computed by interpolation, and we therefore find an algorithmic incarnation of the idea of Serre of studying this variation by studying the corresponding variation of the higher Fourier coefficients.

Example 2. Let $F = \mathbf{Q}(\sqrt{3 \cdot 71})$, and ψ the genus character attached to the extension

$$L = \mathbf{Q}(\sqrt{-3}, \sqrt{-71}).$$

The p -adic L-function is naturally an element in the Iwasawa algebra $\mathbf{Z}_p[[\kappa]]$ where for any positive integer $k \equiv 1 \pmod{p-1}$ we have the interpolation property

$$(136) \quad L_p((1+p)^{k-1} - 1, \psi) = L(1-k, \psi) \cdot \prod_{\mathfrak{p} \mid p} (1 - \psi(\mathfrak{p}) \text{Nm}(\mathfrak{p})^{k-1})$$

which, together with the method described above, we use to compute numerically that

$$\begin{aligned} L_7(\kappa, \psi) &= -103777561 \cdot 7 \kappa - 96435328 \kappa^2 - 15935394 \kappa^3 + \dots \pmod{7^{10}, \kappa^4} \\ L_{11}(\kappa, \psi) &= -8645808191 - 10894273842 \kappa + 4315116763 \kappa^2 + \dots \pmod{11^{10}, \kappa^4} \end{aligned}$$

By inspection of the Newton polygon, we see that the 7-adic L-function has precisely two zeroes in the open unit disk. One of the zeroes is a so-called *exceptional zero*, which is caused by the vanishing of the Euler factor at $\kappa = 0$ (corresponding to $k = 1$) in the equality (136). Such a zero is not present in the 11-adic L-function, since 11 splits into two ideals in F , neither of which are in the kernel of ψ . The other zero of $L_7(\kappa, \psi)$ is more interesting, and we compute its approximate value

$$\kappa = 2669714 \cdot 7 \pmod{7^{10}}.$$

The presence of this zero predicts linear growth of the 7-part of the class number of the cyclotomic tower over L relative to F . In fact, such a divisibility is already observed at the bottom layer, since

$$\text{Cl}_L \simeq \mathbf{Z}/7\mathbf{Z}.$$

A celebrated feature of these p -adic L-functions is contained in the Gross–Stark conjecture, which occurs in situations where p is inert in F , so there is an exceptional zero as in the example above. Indeed, in this case it is known by the work of Darmon–Dasgupta–Pollack [DDP11] that

$$L_p(0, \psi) = 0, \quad L'_p(0, \psi) = \log_p(u), \quad u \in \mathcal{O}_H[1/p]^\times$$

where H is the ray class field cut out by ψ . The numerical computation of the quantity $L'_p(0, \psi)$ may be done without first computing the series $L_p(\kappa, \psi)$, using a more direct and efficient approach. To explain it, assume for simplicity that ψ is unramified, and note the following, see [DPV20]:

- The p -adic family $\mathbf{G}_{\kappa, \kappa}(\psi)(z, z)$ obtained from the diagonal restrictions of the Hilbert Eisenstein series specialising to the p -stabilisations of the Eisenstein series (131) attached to ψ vanishes at $\kappa = 0$ (which corresponds to weight $k = 1$), i.e. we have:

$$\mathbf{G}_{0,0}(\psi)(z, z) = 0.$$

In other words, the exceptional zero of the constant term propagates to the higher coefficients.

- Its first order derivative with respect to κ is the q -expansion of an overconvergent modular form of tame level one, whose constant coefficient is equal to $L'_p(0, \psi)$.



Since the higher Fourier coefficients are elementary even after taking this first order derivative, we may proceed with the strategy as above. We compute the higher Fourier coefficients, write the form in terms of a precomputed basis for the space $M_2^\dagger(\mathrm{SL}_2(\mathbf{Z}))$, and obtain a p -adic approximation for the constant term $L'_p(0, \psi)$. This leads to very significant speed-ups when compared to the naive approach that first computes the series $L_p(\kappa, \psi)$. The following example appears in [LV19, § 4.4]:

Example 3. Let $F = \mathbf{Q}(\sqrt{321})$, which has class number 6, then we compute in under 5 seconds that when ψ is the unique unramified quadratic character, we have $L'_7(0, \psi) = \log_7(u)$, where u is a root of

$$7^{16}u^6 - 20976 \cdot 7^8u^5 - 270624 \cdot 7^4u^4 + 526859689u^3 - 270624u^2 - 20976u + 74 = 0,$$

which is a 7-unit in the Hilbert class field of F .

Remark 1. A beautiful alternative method for the computation of the Gross–Stark unit was developed for real quadratic fields by Dasgupta [Das07] and for cubic fields by Slavov [Sla07] based on the Shintani cone refinements of [Das08]. We mention also the unpublished algorithm of Charollois, based on cocycle relations for GL_n as in [CD14, CDG15]. These works are more closely related to the definition of the p -adic L-functions by Barsky and Cassou–Noguès, but yield a suitable refinement of it. Such a refinement may also be obtained for our method above, where the p -adic family of Eisenstein series is replaced by a cuspidal family in the anti-parallel weight direction. This is the subject of the forthcoming paper [DPV21].

Remark 2. We note the striking parallel with the work of Gross–Zagier [GZ86] on singular moduli. They consider the setting of two imaginary quadratic fields $K_1 = \mathbf{Q}(\tau_1)$ and $K_2 = \mathbf{Q}(\tau_2)$, whose biquadratic compositum contains a unique real quadratic subfield F . The real analytic family $\mathbf{G}_s(z_1, z_2)$ over F attached to the associated genus character is the principal actor in the analytic part of the arguments of Gross–Zagier. They consider its diagonal restriction $\mathbf{G}_s(z, z)$, and show that

- When $s = 0$, we have $\mathbf{G}_s(z, z) = 0$,
- The holomorphic projection of the first derivative

$$\left(\frac{\partial}{\partial s} \mathbf{G}_s(z, z) \right) \Big|_{s=0}^{\mathrm{hol}}$$

has Fourier coefficients related to $\log \mathrm{Nm}(j(\tau_1) - j(\tau_2))$.

Since the holomorphic projection is a holomorphic modular form of weight 2 and level 1, it must vanish! This vanishing gives Gross–Zagier their explicit formula for $\mathrm{Nm}(j(\tau_1) - j(\tau_2))$.

In [DPV20], we observe a suitable p -adic analogue of these phenomena. More precisely, we show that

$$(137) \quad \left(\frac{\partial}{\partial \kappa} \mathbf{G}_{\kappa, \kappa}(\psi)(z, z) \right) \Big|_{\kappa=0}^{\mathrm{ord}}$$

is a classical modular form of weight two, whose component along the Eisenstein series $\mathbf{E}_2^{(p)}$ has constant term $L'_p(0, \psi)$. For Gross–Zagier, the higher Fourier coefficients were related to the norms of differences of singular moduli. In our setting, the higher Fourier coefficients of (137) are related to the RM (real multiplication) values of certain *rigid cocycles*, which were introduced in [DV21] as a framework for singular moduli in the case of real quadratic fields, which are conjecturally algebraic.

In this case, the conjecture may be proved using the idea of Serre in the reverse direction, whereby information on a higher Fourier coefficient is inferred from the constant coefficient. Indeed, using the theory of p -adic deformation of Galois representations, it was proved in Darmon–Dasgupta–Pollack [DDP11] that

the constant term is a rational multiple of the logarithm of a unit in $\mathcal{O}_H[1/p]^\times$. The same then follows for the higher Fourier coefficients, giving the algebraic nature of the RM values of rigid cocycles conjectured in [DV21], at least in the special case of the so-called *Dedekind–Rademacher cocycle*.



REFERENCES

- [AIP18] F. Andreatta, A. Iovita, and V. Pilloni. Le halo spectral. *Ann. Sci. ENS.*, 4(51):603–655, 2018. ↑21, 22, 27.
- [AIS14] F. Andreatta, A. Iovita, and G. Stevens. Overconvergent modular sheaves and modular forms for $\mathrm{GL}_{2,F}$. *Israel J. Math.*, 201(1):299–359, 2014. ↑6, 21.
- [Bar78] D. Barsky. Fonctions zêta padiques d’une classe de rayon des corps totalement réels. *Groupe d’études d’analyse ultramétrique*, 1977–1978. ↑33.
- [BC04] K. Buzzard and F. Calegari. A counterexample to the Gouvêa–Mazur conjecture. *C. R. Math. Acad. Sci. Paris*, 338:751–753, 2004. ↑18, 21, 31.
- [BC05] K. Buzzard and F. Calegari. Slopes of overconvergent 2-adic modular forms. *Compositio Math.*, 141(3):591–604, 2005. ↑17, 20, 21, 25, 27.
- [BC06] K. Buzzard and F. Calegari. The 2-adic eigencurve is proper. *Doc. Math. Extra Volume*, pages 211–232, 2006. ↑23.
- [BCD⁺] M. Bertolini, F. Castella, H. Darmon, S. Dasgupta, K. Prasanna, and V. Rotger. p -Adic L-functions and Euler systems: A tale in two trilogies. In *Automorphic forms and Galois representations*, LMS Lecture Note Series 414, pages 52–101. London Math. Soc. ↑23.
- [BCDT01] C. Breuil, B. Conrad, F. Diamond, and R. Taylor. On the modularity of elliptic curves over \mathbf{q} . *J. Amer. Math. Soc.*, 14(4):843–939, 2001. ↑3.
- [BCP97] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system I: The user language. *J. Symb. Comp.*, 24(3–4):235–265, 1997. ↑1, 17.
- [BG16] K. Buzzard and T. Gee. Slopes of modular forms. In *Families of automorphic forms and the trace formula*, Simons Symp., pages 93–109. Springer-Verlag, 2016. ↑21.
- [BK05] K. Buzzard and L. J. P. Kilford. The 2-adic eigencurve at the boundary of weight space. *Compositio Math.*, 141(3):605–619, 2005. ↑21, 27, 28, 29.
- [BP16] J. Bergdall and R. Pollack. Arithmetic properties of Fredholm series for p -adic modular forms. *Proc. London Math. Soc.*, 113(4):419–444, 2016. ↑21, 27, 29.
- [Buz05] K. Buzzard. Questions about slopes of modular forms. *Astérisque*, 298, 2005. ↑21.
- [BV07] J. Buchmann and U. Vollmer. *Binary Quadratic Forms*, volume 20 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2007. ↑35.
- [Cal08] F. Calegari. The Coleman–Mazur eigencurve is proper at integral weights. *Algebra and Number Theory*, 2(2), 2008. ↑23.
- [Cal13] F. Calegari. Congruences between modular forms. In *Arizona Winter School*, 2013. ↑1, 11.
- [CD14] P. Charollois and S. Dasgupta. Integral eisenstein cocycles on GL_n , i : Sczech’s cocycle and p -adic l-functions of totally real fields. *Camb. J. Math.*, 2(1):49–90, 2014. ↑36.
- [CDG15] P. Charollois, S. Dasgupta, and M. Greenberg. Integral eisenstein cocycles on GL_n , ii : Shintani’s method. *Comment. Math. Helv.*, 90(2):435–477, 2015. ↑36.
- [CE05] F. Calegari and M. Emerton. On the ramification of Hecke algebras at Eisenstein primes. *Invent. Math.*, 160:97–144, 2005. ↑6.
- [CGJ95] R. Coleman, F. Gouvêa, and N. Jochowitz. E_2 , θ , and overconvergence. *Int. Math. Res. Not.*, (1):23–41, 1995. ↑11.
- [Cla40] T. Clausen. Theorem. *Astronomische Nachrichten*, 17(22):351–352, 1840. ↑7.
- [CM98] R. Coleman and B. Mazur. The eigencurve. In *Galois representations in arithmetic algebraic geometry*, volume Durham symposium of *LMS Lecture Note Series* 254. CUP, 1998. ↑6, 21, 22, 27.
- [CN79] P. Cassou-Noguès. Valeurs aux entiers négatifs des fonctions zeta et fonctions zeta p -adiques. *Invent. Math.*, 51:29–60, 1979. ↑33.
- [Coh76] H. Cohen. Variations sur un thème de siegel et hecke. *Acta Arith.*, 30(1):63–93, 1976. ↑33.
- [Col96] R. Coleman. Classical and overconvergent modular forms. *Invent. Math.*, 124:215–241, 1996. ↑21, 23.
- [Col97a] R. Coleman. On the coefficients of the characteristic series of the U-operator. *Proc. Natl. Acad. Sci. USA*, 94:11129–11132, 1997. ↑27.
- [Col97b] R. Coleman. p -Adic Banach spaces and families of modular forms. *Invent. Math.*, 127:417–479, 1997. ↑6, 18, 19, 21, 26, 30.
- [Con08] B. Conrad. p -Adic geometry: Lectures from the 2007 Arizona Winter School, volume 45 of *University Lecture Series*, chapter Several approaches to non-archimedean geometry. Amer. Math. Soc., Providence, RI, 2008. ↑14.

- [CR72] P. Cartier and Y. Roy. Certains calculs numériques relatifs à l'interpolation p -adique des séries de dirichlet. In *Proc. Internat. Summer School, Univ. Antwerp*, volume 350 of *Lecture Notes in Math.*, pages 269–349. Springer-Verlag, Berlin, 1972. ↑33.
- [Das07] S. Dasgupta. Computations of elliptic units for real quadratic fields. *Canadian J. of Math.*, 59:553–574, 2007. ↑36.
- [Das08] S. Dasgupta. Shintani zeta functions and Gross–Stark units for totally real fields. *Duke Math. J.*, 143(2):225–279, 2008. ↑36.
- [DDL15] H. Darmon, M. Daub, S. Lichtenstein, and V. Rotger. Algorithms for Chow–Heegner points via iterated integrals. *Math. Comp.*, 2015. ↑31.
- [DDP11] S. Dasgupta, H. Darmon, and R. Pollack. Hilbert modular forms and the Gross–Stark conjecture. *Ann. of Math. (2)*, 174(1):439–484, 2011. ↑34, 35, 36.
- [Del71] P. Deligne. Formes modulaires et représentations l -adiques. *LNM 179*, pages 139–172, 1971. ↑4.
- [DHH⁺16] E. Dummit, M. Hablicsek, R. Harron, L. Jain, R. Pollack, and D. Ross. Explicit computations of hida families via overconvergent modular symbols. *Res. Number Theory*, 2(25):54, 2016. ↑27.
- [DL16] H. Diao and R. Liu. The eigencurve is proper. *Duke Math. J.*, 165(7):1381–1395, 2016. ↑23.
- [DPV20] H. Darmon, A. Pozzi, and J. Vonk. Diagonal restrictions of p -adic Eisenstein families. *Math. Ann.*, 2020. ↑35, 36.
- [DPV21] H. Darmon, A. Pozzi, and J. Vonk. The values of the Dedekind–Rademacher cocycle at real multiplication points. *Preprint*, 2021. ↑36.
- [DR80] P. Deligne and K. Ribet. Values of abelian L -functions at negative integers over totally real fields. *Invent. Math.*, 59:227–286, 1980. ↑33.
- [DR14] H. Darmon and V. Rotger. Diagonal cycles and Euler systems I. *Ann. Sci. ENS*, 47(4):779–832, 2014. ↑31, 32, 33.
- [DS74] P. Deligne and J.-P. Serre. Formes modulaires de poids 1. *Ann. Sci. ENS*, 7(4):507–530, 1974. ↑4, 6.
- [DS05] F. Diamond and J. Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005. ↑3.
- [DV21] H. Darmon and J. Vonk. Singular moduli for real quadratic fields. *Duke Math. J.*, 170(1):23–93, 2021. ↑36, 37.
- [Dwo62] B. Dwork. On the zeta function of a hypersurface. *Publ. Math. IHÉS*, (12):5–68, 1962. ↑18.
- [Eme98] M. Emerton. *2-adic modular forms of minimal slope*. PhD thesis, Harvard University, 1998. ↑28.
- [Eme11] M. Emerton. p -Adic families of modular forms (after hida, coleman, and mazur). *Séminaire Bourbaki*, 339(1013):31–61, 2011. ↑1.
- [Gau01] C. F. Gauss. *Disquisitiones Arithmeticae*. Lipsiae : Apud G. Fleischer, 1801. ↑35.
- [GM92] F. Gouvêa and B. Mazur. Families of modular eigenforms. *Math. Comp.*, 58, 1992. ↑30.
- [GZ86] B. Gross and D. Zagier. Heegner points and derivatives of L -series. *Invent. Math.*, 84(2):225–320, 1986. ↑36.
- [Hec24] E. Hecke. Analytische Funktionen und Algebraische Zahlen. Zweiter Teil. *Abhandlungen aus dem Mathematischen Universitäts*, 3:213–236, 1924. ↑6, 7, 8, 33.
- [Hec26] E. Hecke. Zur Theorie der elliptischen Modulfunktionen. *Math. Ann.*, 97(1):210–242, 1926. ↑5.
- [Hid86a] H. Hida. Galois representations into $GL_2(\mathbb{Z}_p[[x]])$ attached to ordinary cusp forms. *Invent. Math.*, 85:545–613, 1986. ↑6, 23.
- [Hid86b] H. Hida. Iwasawa modules attached to congruences of cusp forms. *Ann. Sci. ENS*, 19:231–273, 1986. ↑6, 23.
- [Kat73] N. Katz. p -Adic properties of modular schemes and modular forms. In P. Deligne and W. Kuyk, editors, *Modular forms in one variable III*, volume 350 of *LNM*, pages 69–190. Springer-Verlag, 1973. ↑11, 12, 13, 14, 15, 24, 25.
- [KL64] T. Kubota and H.-W. Leopoldt. Eine p -adische Theorie der Zetawerte. I. Einführung der p -adischen Dirichletschen L -Funktionen. *J. Reine Angew. Math.*, 214/215:328–339, 1964. ↑8.
- [Kli62] H. Klingen. über die Werte der Dedekindschen Zetafunktion. *Math. Ann.*, 145:265–272, 1962. ↑6, 7, 33.
- [KM85] N. Katz and B. Mazur. *Arithmetic moduli of elliptic curves*, volume 108 of *Annals of Math. Studies*. Princeton University Press, 1985. ↑13.
- [Kum51] E. Kummer. Über eine allgemeine Eigenschaft der rationalen Entwicklungskoeffizienten einer bestimmten Gattung analytischer Functionen. *J. Reine Angew. Math.*, 41:368–372, 1851. ↑7.
- [Lau11] A. Laufer. Computations with classical and p -adic modular forms. *LMS J. Comp. Math.*, 14:214–231, 2011. ↑24, 26, 31.
- [Lec18] E. Lecouturier. On the Galois structure of the class group of certain Kummer extensions. *J. Lond. Math. Soc. (2)*, 98(1):35–58, 2018. ↑6.
- [Leh49] J. Lehner. Further congruence properties of the fourier coefficients of the modular invariant $j(\tau)$. *Amer. J. Math.*, 71:373–386, 1949. ↑21.
- [Loe07] D. Loeffler. Spectral expansions of overconvergent modular functions. *Int. Math. Res. Not.*, 16, 2007. ↑24, 29.
- [Loe14] D. Loeffler. Lecture notes for TCC course “Modular Curves”. https://warwick.ac.uk/fac/sci/maths/people/staff/david_loeffler/teaching/modularcurves/, 2014. ↑11.
- [LV19] A. Laufer and J. Vonk. Computing p -adic L -functions of totally real fields. *Preprint*, 2019. ↑33, 35, 36.

- [LWX17] R. Liu, D. Wan, and L. Xiao. The eigencurve over the boundary of weight space. *Duke Math. J.*, 166(9):1739–1787, 2017. ↑21, 28.
- [Maz77] B. Mazur. Modular curves and the Eisenstein ideal. *IHÉS Publ. Math.*, 47:33–186, 1977. ↑6.
- [Maz11] B. Mazur. How can we construct abelian Galois extensions of basic number fields? *Bull. Amer. Math. Soc.*, 48(2):155–209, 2011. ↑4, 6.
- [Mer96] L. Merel. L’accouplement de Weil entre le sous-groupe de Shimura et le sous-groupe cuspidal de $j_0(p)$. *J. Reine. Angew. Math.*, 477:71–115, 1996. ↑6.
- [MST06] B. Mazur, W. Stein, and J. Tate. Computation of p -adic heights and log convergence. *Documenta Math.*, pages 577–614, 2006. ↑11.
- [NT19] J. Newton and J. Thorne. Symmetric power functoriality for holomorphic modular forms. *ArXiv preprint*, 2019. ↑21.
- [Pil13] V. Pilloni. Formes modulaires surconvergentes. *Ann. Inst. Fourier*, 63(1):219–239, 2013. ↑6, 21, 22.
- [Ram16] S. Ramanujan. On certain arithmetical functions. *Trans. Cambridge Phil. Soc.*, 22:159–184, 1916. ↑4.
- [Rob15] X. Roblot. Computing p -adic L-functions of totally real number fields. *Math. Comp.*, 84(292):831–874, 2015. ↑33.
- [Roe14] D. Roe. The 3-adic eigencurve at the boundary of weight space. *Int. J. Number Theory*, 10(7):1791–1806, 2014. ↑21, 27, 29.
- [S⁺20] W.A. Stein et al. *Sage Mathematics Software (Version 9.0)*. The Sage Development Team, 2020. <http://www.sagemath.org>. ↑1.
- [Ser62] J.-P. Serre. Endomorphismes complètement continus des espaces Banach p -adiques. *Publ. Math. IHÉS*, 12:69–85, 1962. ↑10, 18, 19.
- [Ser73] J.-P. Serre. Formes modulaires et fonctions zêta p -adiques. In P. Deligne and W. Kuyk, editors, *Modular forms in one variable II*, volume 350 of *LNM*, pages 191–268. Springer-Verlag, 1973. ↑6, 8, 9, 10, 33.
- [Shi66] G. Shimura. A reciprocity law in non-solvable extensions. *J. Reine Angew. Math.*, 221(10.68):209–220, 1966. ↑6.
- [Shi76] T. Shintani. An evaluation of zeta-functions of totally real algebraic fields at non-positive integers. *J. Fac. Sci. Univ. Tokyo*, 23:393–417, 1976. ↑33.
- [Shi78] G. Shimura. The special values of the zeta functions associated with Hilbert modular forms. *Duke Math. J.*, 45:637–678, 1978. ↑33.
- [Sie68] C.-L. Siegel. Berechnung von Zetafunktionen an ganzzahligen Stellen. *Nachr. Akad. Wiss. Göttingen, Math.-Phys. Klasse*, 2:7–38, 1968. ↑6, 7, 8, 33.
- [Sla07] K. Slavov. Gross–stark units for totally real number fields. Master’s thesis, Harvard University, 2007. ↑36.
- [Ste07] W. Stein. *Modular forms, a computational approach*, volume 79 of *Graduate Texts in Mathematics*. Amer. Math. Soc., Providence, RI, 2007. ↑2.
- [Tat95] J. Tate. A review of non-Archimedean elliptic functions. In *Elliptic curves, modular forms, and Fermat’s last theorem*, pages 162–184. Int. Press, 1995. ↑13.
- [TW95] R. Taylor and A. Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141(3):553–572, 1995. ↑3.
- [Von15] J. Vonk. Computing overconvergent forms for small primes. *LMS J. Comp. Math.*, 18(1):250–257, 2015. ↑25, 26.
- [vS40] C. von Staudt. Beweis eines Lehrsatzes, die Bernoullischen Zahlen betreffend. *J. Reine. Angew. Math.*, 21:372–374, 1840. ↑7.
- [Wan98] D. Wan. Dimension variation of classical and p -adic modular forms. *Invent. Math.*, 133:449–463, 1998. ↑26, 28, 30.
- [Was97] L. Washington. *Introduction to Cyclotomic Fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, 2nd edition, 1997. ↑23.
- [Wil30] J. R. Wilton. Congruence properties of ramanujan’s function $\tau(n)$. *Proc. London Math. Soc.*, s2-31(1):1–10, 1930. ↑5.
- [Wil95] A. Wiles. Modular elliptic curves and Fermat’s Last Theorem. *Ann. of Math. (2)*, 141(3):443–551, 1995. ↑3.
- [WWE20] P. Wake and C. Wang-Erickson. The rank of Mazur’s Eisenstein ideal. *Duke Math. J.*, 169(1):31–115, 2020. ↑6.

INSTITUTE FOR ADVANCED STUDY, PRINCETON NJ 08540, UNITED STATES

Email address: vonk@ias.edu