p-ADIC HEIGHTS OF HEEGNER POINTS ON SHIMURA CURVES

by

Daniel Disegni

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

ABSTRACT

p-adic Heights of Heegner points on Shimura curves

Daniel Disegni

Let f be a primitive Hilbert modular form of weight 2 and level N for the totally real field F, and let $p \nmid N$ be an odd rational prime such that f is ordinary at all primes $\wp | p$. When E is a CM extension of F of relative discriminant Δ prime to Np, we give an explicit construction of the p-adic Rankin–Selberg L-function $L_p(f_E, \cdot)$ and prove that when the sign of its functional equation is -1, its central derivative is given by the p-adic height of a Heegner point on the abelian variety A associated to f. This p-adic Gross–Zagier formula generalises the result obtained by Perrin-Riou when $F = \mathbf{Q}$ and N satisfies the so-called Heegner condition. We deduce applications to both the p-adic and the classical Birch and Swinnerton-Dyer conjectures for A.

Table of Contents

Acknowledgments

I am grateful to my advisor Professor Shou-Wu Zhang for suggesting this area of research and for his support and encouragement; and to Amnon Besser, Luis Garcia Martinez, Yifeng Liu, Jeanine Van Order, Hui Xue and Shou-Wu Zhang for useful conversations or correspondence related to this work. During these years I have learned a lot especially from Professors Dorian Goldfeld, Aise Johan de Jong, Eric Urban, Shou-Wu Zhang and Wei Zhang, as well as from many discussions with other faculty members and fellow Ph.D. students; while the staff at the Columbia Department of Mathematics has greatly contributed to the smooth and pleasant running of my doctorate. It is a pleasure to thank them all here.

This thesis is the natural offspring, or, to use a more mathematical metaphor, the *fibre product*, of the works of Perrin-Riou [30] and Zhang [45–47] – the 'base' being, of course, the pioneering work of Gross–Zagier [13]. My debt to their ideas cannot be underestimated and will be obvious to the reader.

Last but not least, I would like to warmly thank all my family and friends, near and far, for all their friendship and support. This work is dedicated to them.

Introduction

In this work we generalise Perrin-Riou's p-adic analogue of the Gross–Zagier formula [30] to totally real fields, in a generality similar to the work of Zhang [45–47]. We describe here the main result and its applications.

The p-adic Rankin–Selberg L-function. — Let f be a primitive Hilbert modular form of parallel weight 2, level N and trivial character for the totally real field F of degree g and discriminant D_F. Let p be a rational prime coprime to $2N$. Fix embeddings ι_∞ and ι_p of the algebraic closure $\overline{\mathbf{Q}}$ of F into $\mathbf{C}$ and $\overline{\mathbf{Q}}_p$ respectively. We assume that f is *ordinary* at p, that is, that for each prime $\wp$ of $\mathcal{O}_F$ dividing p, the coefficient $a(f,\wp)$ of the $\wp^{\text{th}}$ Hecke polynomial of f

$$P_{\wp,f}(X) = X^2 - a(f,\wp)X + \mathbf{N}\wp$$

is a p-adic unit for the chosen embedding; we let $\alpha_\wp = \alpha_\wp(f)$ be the unit root of $P_{\wp,f}$.

Let $E \subset \overline{\mathbf{Q}}$ be a CM (that is, quadratic and purely imaginary) extension of F of relative discriminant Δ coprime to Np, let

$$\varepsilon = \varepsilon_{E/F} : F_{\mathbf{A}}^\times / F^\times \to \{\pm 1\}$$

be the associated Hecke character and $\mathfrak{N} = N_{E/F}$ be the relative norm. If

$$\mathscr{W} : E_{\mathbf{A}}^\times / E^\times \to \overline{\mathbf{Q}}^\times$$

is a finite order Hecke character[(1)] of conductor $\mathfrak{f}$, the Rankin–Selberg L-function $L(f_E, \mathscr{W}, s)$ is the entire function defined for $\Re(s) > 3/2$ by

$$L(f_E, \mathscr{W}, s) = L^{N\Delta(\mathscr{W})}(\varepsilon\mathscr{W}|_{F_{\mathbf{A}}^\times}, 2s-1)\sum_m \frac{a(f,m)r_{\mathscr{W}}(m)}{\mathbf{N}m^s},$$

where $\Delta(\mathscr{W}) = \Delta\mathfrak{N}(\mathfrak{f})$, $r_{\mathscr{W}}(m) = \sum_{\mathfrak{N}(\mathfrak{a})=m}\mathscr{W}(\mathfrak{a})$ (the sum running over all nonzero ideals of $\mathscr{O}_E$) and

$$L^{N\Delta(\mathscr{W})}(\varepsilon\mathscr{W}|_{\mathscr{O}_F}, s) = \sum_{(m, N\Delta(\mathscr{W}))=1}\varepsilon(m)\mathscr{W}(m)\mathbf{N}m^{-s}.$$

This L-function admits a p-adic analogue (§4). Let E'_∞ be the maximal abelian extension of E unramified outside p, and E_∞ the maximal $\mathbf{Z}_p$-extension of E. Then $\mathscr{G} = \mathrm{Gal}(E_\infty/E)$ is a direct factor of finite, prime to p index in $\mathscr{G}' = \mathrm{Gal}(E'_\infty/E)$. (It has rank $1+\delta+g$ over $\mathbf{Z}_p$, where δ is the Leopoldt defect of F.)

Theorem A. — *There exists a bounded Iwasawa function $L_p(f_E) \in \mathbf{Q}_p[\![\mathscr{G}']\!]$ satisfying the interpolation property*

$$L_p(f_E)(\mathscr{W}) = \frac{\tau(\mathscr{W})\mathbf{N}(\Delta(\mathscr{W}))^{1/2}V_p(f,\mathscr{W})\overline{\mathscr{W}}(\mathfrak{D})}{\alpha_{\mathfrak{N}(\mathfrak{f})}\Omega_f}L(f_E, \overline{\mathscr{W}}, 1),$$

for all finite order characters $\mathscr{W}$ of $\mathscr{G}'$ which are ramified exactly at the places $w|p$. Here both sides are algebraic numbers[(2)], $\overline{\mathscr{W}} = \mathscr{W}^{-1}$ and

$$\Omega_f = (8\pi^2)^g\langle f, f\rangle_N$$

[(1)] We will throughout use the same notation for a Hecke character, the associated ideal character, and the associated Galois character.

[(2)] By a well-known theorem of Shimura [35]. They are compared via ι_p^{-1} and ι_∞^{-1}.

with $\langle\cdot,\cdot\rangle_N$ *the Petersson inner product* (1.1.2)*;* $\tau(\mathscr{W})$ *is a Gauß sum,* $V_p(f,\mathscr{W})$ *is a product of Euler factors and* $\alpha_{\mathfrak{N}(\mathfrak{f})} = \prod_{\wp|p} \alpha_\wp^{v_\wp(\mathfrak{N}(\mathfrak{f}))}$.

The restriction of $L_p(f_E)$ *to* $\mathscr{G}$ *satisfies a functional equation with sign* $(-1)^g\varepsilon(N)$ *relating its values at* $\mathscr{W}$ *to its values at* $\mathscr{W}^{-c}$*, where* c *is the nontrivial automorphism of* E/F *and* $\mathscr{W}^c(\sigma) = \mathscr{W}(c\sigma c)$.

This is essentially a special case of results of Panchishkin [28] and Hida [15]; we reprove it entirely here (see §4) because the precise construction of $L_p(f_E)$ will be crucial for us. It is achieved, using a technique of Hida and Perrin-Riou, via the construction of a convolution Φ of Eisenstein and theta measures on $\mathscr{G}'$ valued in p-adic modular forms, giving an analogue of the kernel of the classical Rankin–Selberg convolution method. The approach we follow is adelic; one novelty introduced here is that the theta measure is constructed via the Weil representation, which seems very natural and would generalise well to higher rank cases.

On the other hand, Manin [25] and Dabrowski [11] have constructed a p-adic L-function $L_p(f,\cdot)$ as an analogue of the standard L-function $L(f,s)$. It is a locally analytic function on the p-adic Lie group of continuous characters $\chi : \mathscr{G}_F' \to \mathbf{C}_p$, where $\mathscr{G}_F'$ is the Galois group of the maximal abelian extension of F unramified outside p; it satisfies the interpolation property

$$L_p(f,\chi) = \frac{\tau(\chi)}{\alpha_P} \prod_{\substack{\wp|p \\ \wp\nmid P}} \left(1 - \frac{1}{\alpha_\wp}\right)^2 \frac{L(f,\overline{\chi},1)}{\Omega_f^+}$$

for all finite order characters χ of conductor P which are trivial at infinity. (Here Ω_f^+ is a suitable period, cf. §9.1.)

The corresponding formula for complex L-functions implies a factorisation

$$L_p(f_E,\chi\circ\mathfrak{N}) = \frac{\Omega_f^+\Omega_{f_\varepsilon}^+}{D_E^{-1/2}\Omega_f} L_p(f,\chi)L_p(f_\varepsilon,\chi),$$

where f_ε is the form with coefficients $a(f_\varepsilon,m) = \varepsilon(m)a(f,m)$ and $D_E = \mathbf{N}(\Delta)$.

Heegner points on Shimura curves and the main theorem. — Suppose that $\varepsilon(N) = (-1)^{g-1}$, where $g = [F:\mathbf{Q}]$. Then for each embedding $\tau: F \to \mathbf{C}$, there is a quaternion algebra $B(\tau)$ over F ramifed exactly at the finite places $v|N$ for which $\varepsilon(v) = -1$ and the infinite places different from τ; it admits an embedding $\rho: E \hookrightarrow B(\tau)$, and we can consider an order R of $B(\tau)$ of discriminant N and containing $\rho(\mathscr{O}_E)$. This data defines a *Shimura curve* X. It is an algebraic curve over F, whose complex points for any embedding $\tau: F \to \mathbf{C}$ are described by

$$X(\mathbf{C}_\tau) = B(\tau)^\times\backslash\mathfrak{H}^\pm \times \widehat{B}(\tau)^\times/\widehat{F}^\times\widehat{R}^\times \cup \{\text{cusps}\}.$$

It plays the role of the modular curve $X_0(N)$ in the works of Gross–Zagier [13] and Perrin-Riou [30] who consider the case $F = \mathbf{Q}$ and $\varepsilon(v) = 1$ for all $v|N$ (it is only in this case that the set of cusps is not empty).

The curve X is connected but not geometrically connected. Let $J(X)$ be its Albanese ($\cong$ Jacobian) variety; it is an abelian variety defined over F, geometrically isomorphic to the product of the Albanese varieties of the geometrically connected components of X. There is a natural map $\iota: X \to J(X)\otimes\mathbf{Q}$ given by $\iota(x) = [x] - [\xi]$, where $[\xi] \in \mathrm{Cl}(X)\otimes\mathbf{Q}$ is a

canonical divisor class constructed in [45] having degree 1 in every geometrically connected component of X; an integer multiple of ι gives a morphism $X \to J(X)$ defined over F.

As in the modular curve case, the curve X admits a finite collection of *Heegner points* defined over the Hilbert class field H of E and permuted simply transitively by $\mathrm{Gal}(H/E)$. They are the points represented by $(\sqrt{-1}, t)$ for $t \in \widehat{E}^\times / E^\times \widehat{F}^\times \widehat{\mathcal{O}}_E^\times$ when we use the complex description above and view $E \subset B$ via ρ. We let y be any such Heegner point, and let $[z]$ denote the class

$$[z] = u^{-1}\iota\left(\mathrm{Tr}_{H/E} y\right) \in J(X)(E) \otimes \mathbf{Q},$$

where $u = [\mathcal{O}_E^\times : \mathcal{O}_F^\times]$.

As a consequence of Jacquet–Langlands theory, the Hecke algebra on Hilbert modular forms of level N acts through its quaternionic quotient on $J(X)$. Let $z_f \in J(X)(E) \otimes \overline{\mathbf{Q}}$ be the f-component of $[z]$.

Heights and the formula. — On any curve X over a number field E, there is a notion (§5.1) of p-adic height $\langle \cdot, \cdot \rangle_\ell$ attached to the auxiliary choices of splittings[(3)] of the Hodge filtrations on $H^1_{\mathrm{dR}}(X/E_w)$ for $w|p$ and of a *p-adic logarithm* $\ell : E_{\mathbf{A}}^\times / E^\times \to \mathbf{Q}_p$. It is a symmetric bilinear pairing on the group of degree zero divisors on X modulo rational equivalence, which we can view as a pairing on $J(X)(E)$. And more generally, for any abelian variety A/E there is defined a p-adic height pairing on $A(E) \times A^\vee(E)$.

Let $\mathscr{W}$ be a Hecke character of E taking values in $1 + p\mathbf{Z}_p \subset \mathbf{Z}_p^\times$. Under the assumption $\varepsilon(N) = (-1)^{g-1}$, the value $L_p(f_E, \mathbf{1})$ is zero by the (complex or

(3) In our case there will be a canonical choice for those splittings at least on the direct factor of $H^1_{\mathrm{dR}}(X/E_w)$ of interest to us.

p-adic) functional equation. Then we can consider the derivative of $L_p(f_E)$ in the $\mathscr{W}$-direction

$$L'_{p,\mathscr{W}}(f_E,\mathbf{1}) = \frac{d}{ds}\bigg|_{s=0} L_p(f_E)(\mathscr{W}^s).$$

Assume that $\Delta_{E/F}$ is totally odd, and that every prime $\wp|p$ is split in E. (These assumptions can be removed *a posteriori* if the left-hand side of the formula below is nonzero – see §8.2.)

Theorem B. — *Suppose that* $\varepsilon_{E/F}(N) = (-1)^{g-1}$. *Then* $L_p(f_E,\mathbf{1}) = 0$ *and*

$$L'_{p,\mathscr{W}}(f_E,\mathbf{1}) = D_F^{-2}\prod_{\wp|p}\left(1-\frac{1}{\alpha_\wp}\right)^2\left(1-\frac{1}{\varepsilon(\wp)\alpha_\wp}\right)^2\langle z_f, z_f\rangle_{\mathscr{W}}$$

where $\langle\cdot,\cdot\rangle_{\mathscr{W}}$ *is the height pairing on* $J(X)(E)$ *associated to the logarithm* $\ell = \frac{d}{ds}\big|_{s=0}\mathscr{W}^s$.

Applications to the conjecture of Birch and Swinnerton-Dyer. — It is conjectured that to the Hilbert modular newform f one can attach a simple abelian variety $A = A_f$ over F, characterised uniquely up to isogeny[(4)] by the equality of L-functions

$$L(A,s) = \prod_{\sigma : M_f \to \mathbf{C}} L(f^\sigma, s)$$

up to Euler factors at places dividing N. Here $M = M_f$ is the field generated by the Fourier coefficients of f; A has dimension $[M:\mathbf{Q}]$ and its endomorphism algebra contains M (we say that A is *of* $GL_2(M)$*-type*; in fact since F is totally real, A is of *strict* GL_2-type, that is, its endomorphism algebra

(4) Thanks to Faltings's isogeny theorem [12].

equals M – see e.g. [43, Lemma 3.3]). The conjecture is known to be true ([45, Theorem B]) when either $[F:\mathbf{Q}]$ is odd or $v(N)$ is odd for some finite place v; in this case A is a quotient ϕ of $J(X)$ for a suitable Shimura curve X of the type described above. Viceversa any abelian variety of GL_2-type (for some field M) over a totally real field F is conjectured to be associated to a Hilbert modular form f as above.

In view of known $\mathrm{Aut}(\mathbf{C}/\mathbf{Q})$-equivariance properties of automorphic L-functions and the above equality, the order of vanishing of $L(A,s)$ at $s=1$ will be an integer multiple $r[M:\mathbf{Q}]$ of the dimension of A. We call r the *M-order of vanishing* of $L(A,s)$ or the *analytic M-rank* of A.

***Conjecture* (Birch and Swinnerton-Dyer).** — *Let A be an abelian variety of $GL_2(M)$-type over a totally real field F of degree g.*

1. *The M-order of vanishing of $L(A,s)$ at $s=1$ is equal to the dimension of $A(F)_{\mathbf{Q}}$ as M-vector space.*
2. *The Tate-Shafarevich group $Ш(A/F)$ is finite, and the leading term of $L(A,s)$ at $s=1$ is given by*

$$\frac{L^*(A,1)}{\Omega_A} = D_F^{-d/2}|Ш(A/F)|R_A\prod_{v\nmid\infty}c_v = \mathrm{BSD}(A),$$

where $d=\dim A=[M:\mathbf{Q}]$, the c_v are the Tamagawa numbers of A at finite places (almost all equal to 1),

$$\Omega_A = \prod_{\tau:F\to\mathbf{R}}\int_{A(\mathbf{R}_\tau)}|\omega_A|_\tau$$

for a Néron differential[5] ω_A, *and*

$$R_A = \frac{\det(\langle x_i, y_j\rangle)}{[A(F):\sum \mathbf{Z}x_i][A^\vee(F):\sum \mathbf{Z}y_j]}$$

is the regulator of the Néron-Tate height paring on $A(F)\times A^\vee(F)$, *defined using any bases* $\{x_i\}$, $\{y_j\}$ *of* $A(F)_{\mathbf{Q}}$ *and* $A^\vee(F)_{\mathbf{Q}}$.

By the automorphic description of $L(A,s)$ and results of Shimura [35], we know that $L(A,s)/\prod_{\sigma:M_f\to\mathbf{C}}\Omega^+_{f^\sigma}$ is an algebraic number. Comparison with the Birch and Swinnerton-Dyer conjecture suggests the following conjecture.

Conjecture. — *We have*

$$\Omega_A \sim \prod_{\sigma:M_f\to\mathbf{C}} \Omega^+_{f^\sigma} \qquad \text{in } \mathbf{C}^\times/\overline{\mathbf{Q}}^\times.$$

The conjecture is known for $F=\mathbf{Q}$ [36] or when A has complex multiplication (over $\overline{\mathbf{Q}}$) [5]; see §9 below for a more precise conjecture and some further evidence and motivation. Assuming the conjecture, we can define a p-adic L-function $L_p(A)$ for A by

$$L_p(A) = \frac{\prod_\sigma \Omega^+_{f^\sigma}}{\Omega_A} \prod_{\sigma:M_f\to\mathbf{C}} L_p(f^\sigma) \prod_{v|N} \frac{L_{v,p}(A)}{\prod_\sigma L_{v,p}(f^\sigma)}$$

for any prime p of good reduction. (Here $L_{v,p}(A)$, $L_{v,p}(f^\sigma)$ interpolate the bad Euler factors.)

[5]When it exists, which is only guaranteed if $F=\mathbf{Q}$. Otherwise, we take for ω_A any generator of $H^0(A,\Omega^d_{A/F})$ and to define Ω_A we divide by the product of the indices $[H^0(\mathscr{A}_v,\Omega^d_{\mathscr{A}_v/\mathscr{O}_{F,v}}):\mathscr{O}_{F,v}\widetilde{\omega_A}]$ of (the extension of) ω_A in the space of top differentials on the local Néron models $\mathscr{A}_v/\mathscr{O}_{F,v}$ of A.

Then, fixing a ramified Hecke character $\nu\colon \mathscr{G}'_F \to 1+p\mathbf{Z}_p \subset \mathbf{Z}_p^\times$ which we omit from the notation, one can formulate a p-adic version of the Birch and Swinnerton-Dyer conjecture similarly as above for $L_p(A,\nu^s)$:[6] the formula reads

$$\prod_{\wp|p}(1-\alpha_\wp^{-1})^{-2}L_p^*(A,\mathbf{1}) = \mathrm{BSD}_p(A)$$

where $\mathrm{BSD}_p(A)$ differs from $\mathrm{BSD}(A)$ only in the regulator term, which is now the regulator of the p-adic height pairing on $A(F)\times A^\vee(F)$ associated to the p-adic logarithm ℓ deduced from ν as in Theorem B.

Similarly, one can formulate a main conjecture of Iwasawa theory for $L_p(A)$.

Then, just as in [30], we can deduce the following arithmetic application of Theorem B.

Theorem C. — *For the abelian varitey $A = A_f$ we have:*

1. *The following are equivalent:*
   (a) *The p-adic L-function $L_p(A,\nu^s)$ has M_f-order of vanishing $r \le 1$ at the central point.*
   (b) *The complex L-function $L(A,s)$ has M_f-order of vanishing $r \le 1$ at the central point and the p-adic height pairing associated to ν is non-vanishing on $A(F)$.*
2. *If either of the above assumptions holds, the rank parts of the classical and the p-adic Birch and Swinnerton-Dyer conjecture are true for A and the Tate-Shafarevich group of A is finite.*

[6] Here $s \in \mathbf{Z}_p$ and the central point is $s = 0$, corresponding to $\nu^0 = \mathbf{1}$.

3. *If moreover the cyclotomic Iwasawa main conjecture is true for A, then the classical and the p-adic Birch and Swinnerton-Dyer formulas for A are true up to a p-adic unit.*

Proof. — In 1., the statement follows trivially from the construction of $L_p(A)$ if $r = 0$; if $r = 1$, both conditions are equivalent to the assertion that for a suitable CM extension E, the Heegner point $z_f = z_{f,E}$ is nontorsion: this is obvious from our main theorem in case 1a; in case 1b, by the work of Zhang [45, 46] (generalising Gross–Zagier [13] and Kolyvagin [23, 24]), the Heegner point

$$P = \sum_\sigma \mathrm{Tr}_{E/F}\phi(z_{f^\sigma,E}) \in A(F)\otimes \mathbf{Q}$$

(with $\phi : J(X) \to A$) generates $A(F)\otimes\mathbf{Q}$ as M_f-vector space, so that the p-adic height pairing on $A(F)$ is non-vanishing if and only if it is nonzero at z_f. Part 2. then follows from 1. and the results of Zhang [45, 46].

Schneider [33] proves an "arithmetic" version of the p-adic Birch and Swinnerton-Dyer formula for (the Iwasawa L-function associated to) A, which under the assumption of 3. can be compared to the analytic p-adic formula as explained in [30] to deduce the p-adic Birch and Swinnerton-Dyer formula up to a p-adic unit. In the analytic rank 0 case the classical Birch and Swinnerton-Dyer formula follows immediately. In the case of analytic rank 1, recall that the main result of [45, 47] is, in our normalisation, the formula

$$\frac{L'(f_E,1)}{\Omega_f} = \frac{1}{D_F^2 D_E^{1/2}}\langle z_f, z_f\rangle = D_E^{-1/2}\mathrm{GZ}(f_E)$$

(where $\langle\cdot,\cdot\rangle$ denotes the Néron–Tate height); whereas we introduce the notation $\mathrm{GZ}_p(f_E)$ to write our formula (for any fixed ramified cyclotomic character $\mathscr{W}$) as

$$L_p'(f_E,\mathbf{1})=\prod_{\wp\mid p}\left(1-\frac{1}{\alpha_\wp}\right)^2\left(1-\frac{1}{\varepsilon(\wp)\alpha_\wp}\right)^2\mathrm{GZ}_p(f_E).$$

Then, after choosing E suitably so that $L(f_\varepsilon,1)\neq 0$ (which can be done by [4], [40]), we can argue as in [30] to compare the p-adic and the complex Birch and Swinnerton-Dyer formulas via the correspoding Gross–Zagier formulas to get the result. Namely, denoting by $L_N(A)$ a suitable ratio of bad Euler factors[(7)], we have

$$\begin{aligned}\frac{L^*(A,1)}{\Omega_A\mathrm{BSD}(A)}&=\frac{\prod_\sigma\Omega^+_{f^\sigma}}{\Omega_A}\frac{1}{\mathrm{BSD}(A)}\prod_\sigma\frac{L'(f_E^\sigma,1)}{\Omega_{f^\sigma}}\frac{\Omega_{f^\sigma}}{\Omega^+_{f^\sigma}\Omega^+_{f_\varepsilon^\sigma}}\frac{\Omega^+_{f_\varepsilon^\sigma}}{L(f_\varepsilon^\sigma,1)}L_N(A)\\&=\frac{\prod_\sigma\Omega^+_{f^\sigma}}{\Omega_A}\frac{\prod_\sigma\mathrm{GZ}(f_E^\sigma)}{\mathrm{BSD}(A)}\prod_\sigma\frac{D_E^{-1/2}\Omega_{f^\sigma}}{\Omega^+_{f^\sigma}\Omega^+_{f_\varepsilon^\sigma}}\frac{\Omega^+_{f_\varepsilon^\sigma}}{L(f_\varepsilon^\sigma,1)}L_N(A)\end{aligned}$$

by the complex Gross–Zagier formula and the factorisation of $L(f_E,s)$. Similarly,

$$\prod_{\wp\mid p}(1-\alpha_\wp^{-1})^{-2}\frac{L_p^*(A,1)}{\mathrm{BSD}_p(A)}=\frac{\prod_\sigma\Omega^+_{f^\sigma}}{\Omega_A}\frac{\prod_\sigma\mathrm{GZ}_p(f_E^\sigma)}{\mathrm{BSD}_p(A)}\prod_\sigma\frac{D_E^{-1/2}\Omega_{f^\sigma}}{\Omega^+_{f^\sigma}\Omega^+_{f_\varepsilon^\sigma}}\frac{\Omega^+_{f_\varepsilon^\sigma}}{L(f_\varepsilon^\sigma,1)}L_N(A)$$

by the p-adic Gross–Zagier formula, the factorisation of $L_p(f_E)$ and the interpolation property of $L_p(f_\varepsilon)$. Since we know that the left-hand side of the

[(7)]Namely, $L_N(A)=\prod_{v\mid N}L_v(A,1)/\prod_\sigma L_v(f^\sigma,1)$.

last formula is a p-adic unit, the result follows from observing the equality

$$\frac{\prod_\sigma \mathrm{GZ}(f_E^\sigma)}{\mathrm{BSD}(A)} = \frac{\prod_\sigma \mathrm{GZ}_p(f_E^\sigma)}{\mathrm{BSD}_p(A)}$$

of rational numbers.[8] □

Discussion of the assumptions. — The conjecture on periods could be dispensed of if one were willing to work with a "wrong" p-adic L-function for A (namely, one without the period ratio appearing in the definition above). Then at least the rank part of the p-adic Birch and Swinnerton-Dyer conjecture makes sense and parts 1 and 2 of the Theorem hold. The nonvanishing of the p-adic height pairing is only known for CM elliptic curves [1]. The Iwasawa main conjecture is known in most cases for elliptic curves over $\mathbf{Q}$ thanks to the work of Rubin, Kato and Skinner–Urban (see [38]). For Hilbert modular forms, one divisibility in the CM case is known by work of Ming-Lun Hsieh [17] (this implies one divisibilty in the above result), and in the non-CM case there is work in progress by Xin Wan. The other divisibility is not known but could be within reach with current methods, cf. [39, remarks on top of p.6].

Plan of the proof. — The proof of the main formula follows the strategy of Perrin-Riou [30]. It is enough (see §8) to study the case where $\mathscr{W}$ is cyclotomic, since both sides of the formula are zero when $\mathscr{W}$ is anticyclotomic.

[8] The rationality of the ratios follows from the fact that the z_{f^σ} essentially belong to $J(X)(F)$ – that is, they belong to the $+1$-eigenspace for the action of $\mathrm{Gal}(E/\mathbf{Q})$ on $J(X)(E)\otimes\overline{\mathbf{Q}}$ – and that in this sense, their images $\phi(z_{f^\sigma})$ form a $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-invariant basis of $A(F)\otimes\overline{\mathbf{Q}}$, orthogonal for the height pairing.

In the first part of this thesis, we construct a measure Φ on $\mathscr{G}$ valued in p-adic modular forms such that $L_p(f_E)(\mathscr{W})$ essentially equals $L_{f_0}(\Phi(\mathscr{W}))$, where L_{f_0} is a p-adic analogue of the functional "Petersson product with f" on p-adic modular forms. This allows us to write

$$L'_{p,\mathscr{W}}(f_E, \mathbf{1}) \doteq L_{f_0}(\Phi'_{\mathscr{W}}),$$

where $\doteq$ denotes equality uo to suitable nonzero factors, and $\Phi'_{\mathscr{W}} = \frac{d}{ds}\Phi(\mathscr{W}^s)|_{s=0}$ is a p-adic Hilbert modular form.

On the other hand, there is a modular form Ψ with Fourier coefficients given by $\langle z, T(m)z\rangle_{\mathscr{W}}$, so that $L_{f_0}(\Psi) \doteq \langle z_f, z_f\rangle_{\mathscr{W}}$. It can be essentially written as a sum $\Psi_{\text{fin}} + \Psi_p$, where Ψ_{fin} encodes the local contributions to the height from places not dividing p and Ψ_p the contribution from places above p. Then we can show by explicit computation that the Fourier coefficients of Φ' are equal to the Fourier coefficients of Ψ_{fin} up to the action of suitable Hecke operators, whereas by making use of the assumption that f is ordinary at p we prove that $L_{f_0}(\Psi_p) = 0$. The desired formula follows.

One difficulty in the approach just outlined is that compared to the case of modular curves there are no cusps available, so that in this case the divisors z and $T(m)z$ have intersecting supports and the decomposition of the height pairing into a sum of local pairings is not available. Our solution to this problem, which is inspired from the work of Zhang [45], is to make use of p-adic Arakelov theory as developed by Besser [3] (see §5.2) and work consistently in a suitable space of Fourier coefficients.

Notation. — Throughout this text we use the following notation and assumptions, unless otherwise noted:

- F is a totally real field of degree g;
- $\mathbf{N}_F$ is the monoid of nonzero ideals of $\mathcal{O}_F$;
- $|\cdot|_v$ is the standard absolute value on F_v;
- $\mathbf{A} = \mathbf{A}_F$ is the adèle ring of F; if $*$ is a place or a set of places or an ideal of F, the component at $*$ (respectively away from $*$) of an adelic object x is denoted x_* (respectively x^*). For example if $\phi = \prod_v \phi_v$ is a Hecke character and δ is an ideal of $\mathcal{O}_F$ we write $\phi_\delta(y) = \prod_{v|\delta} \phi_v(y_v)$, and $|y|_\delta = \prod_{v|\delta} |y|_v$. We also use the notation

$$|m|_v = |\pi_m|_v,\ |m|_\delta = |\pi_m|_\delta, \quad \phi_v(m) = \phi_v(\pi_m),\ \phi_\delta(m) = \phi_\delta(\pi_m)$$

  if m is an ideal of $\mathcal{O}_F$ and ϕ is unramified at δ (here π_m satisfies $\pi_m \mathcal{O}_F = m$).
- "$>$" denotes the partial order on $\mathbf{A}_F$ given by $x > 0$ if and only if x_∞ is totally positive;
- $R_{\mathbf{A}} = R \otimes_F \mathbf{A}$ if R is an F-algebra;
- $\mathbf{N}m$ is the absolute norm of an ideal m in a number field (the index of m in the ring of integers: it is a positive natural number);
- d_F is the inverse different of F;
- π_N, for N an ideal of $\mathcal{O}_F$, is the idèle with components $\pi_v^{v(N)}$ for $v \nmid \infty$ and 1 for $v|\infty$.
- $D_F = \mathbf{N}d_F$ is the discriminant of F.
- $m^\times = \{a \in F_{\mathbf{A}}^\times \,|\, a\mathcal{O}_F = m\}$ if m is any nonzero fractional ideal of F (this notation will be used with $m = d_F^{-1}$).
- E is a quadratic CM (that is, totally imaginary) extension of F;
- $\mathfrak{D} = \mathfrak{D}_{E/F}$ is the relative inverse different of E/F.
- $\mathfrak{N} = N_{E/F}$ is the realtive norm on E or any E-algebra;

– $\Delta = \Delta_{E/F} = \mathfrak{N}(\mathfrak{D})$ is the relative discriminant of E/F and we assume

$$(\Delta_{E/F}, p) = 1;$$

in §§ 2.5, 4.5 and part of §3.2 we further assume that

$$(\Delta, 2) = 1$$

and in §§ 7.2, 8.1, that

$$(\Delta, 2) = 1 \text{ and all primes } \wp \text{ dividing } p \text{ are split in } E.$$

– $D_E = \mathbf{N}(\Delta)$ is the absolute discriminant of E.

– $U_F(N)$ is the subgroup of $\widehat{\mathscr{O}}_F^\times = \prod_v \mathscr{O}_{F,v}^\times \subset F_{\mathbf{A}^\infty}^\times$ consisiting of elements $x \equiv 1 \bmod N\widehat{\mathscr{O}}_F$, if N is any ideal of $\mathscr{O}_F$;
– $\mathbf{e}_v(x) = \exp(-2\pi i\{\mathrm{Tr}_{F_v/\mathbf{Q}_p}(x)\}_p)$ for $v|p < \infty$ and $\{y\}_p$ the p-fractional part of $y \in \mathbf{Q}_p$ is the standard addtive character of F_v, with conductor $d_{F,v}^{-1}$; for $v|\infty$, $\mathbf{e}_v(x) = \exp(2\pi i \mathrm{Tr}_{F_v/\mathbf{R}}(x))$;
– $\mathbf{e}(x) = \prod_v \mathbf{e}_v(x_v)$ is the standard additive character of $\mathbf{A}_F$.

– $\mathbf{1}_Y$ is the characteristic function of the set Y;
– if φ is any logical proposition, we define $\mathbf{1}[\varphi]$ to be 1 when φ is true and 0 when φ is false – e.g. $\mathbf{1}[x \in Y] = \mathbf{1}_Y(x)$.

PART I

p-ADIC L-FUNCTION AND MEASURES

This part is dedicated to the construction of the measure giving the p-adic Rankin–Selberg L-function $L_p(f_E)$ and to the computation of its Fourier coefficients.

1. p-adic modular forms

1.1. Hilbert modular forms. — Let us define compact subgroups of $\mathbf{GL}_2(\mathbf{A}^\infty)$ as follows:

- $K_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{GL}_2(\widehat{\mathcal{O}}_F) \mid c \equiv 0 \bmod N\widehat{\mathcal{O}}_F \right\}$ if N is an ideal of $\mathcal{O}_F$;
- $K_1(N,n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K_0(N) \mid a \equiv 1 \bmod N/n\widehat{\mathcal{O}}_F, d \equiv 1 \bmod n\widehat{\mathcal{O}}_F \right\}$ if $n|N$ are ideals of $\mathcal{O}_F$.

Let k be a positive integer and ψ be a character of $F_{\mathbf{A}}^\times/F^\times$ of conductor dividing N satisfying $\psi_v(-1) = (-1)^k$ for $v|\infty$. A **Hilbert modular form** of parallel weight k, level $K_1(N,n)$ and character ψ is a smooth function

$$f : \mathbf{GL}_2(F)\backslash\mathbf{GL}_2(\mathbf{A}_F) \to \mathbf{C}$$

of *moderate growth*[9] satisfying[10]

$$f\left(\begin{pmatrix} z & \\ & z \end{pmatrix} g \begin{pmatrix} a & b \\ c & d \end{pmatrix} r(\theta)\right) = \psi(z)\psi_{N/n}(a)\psi_n(d)\mathbf{e}_\infty(k\theta)f(g)$$

for each $z \in \mathbf{A}_F^\times$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K_0(N)$ and $\theta = (\theta_v)_{v|\infty} \in F_\infty$, with $r(\theta) = \prod_{v|\infty} r(\theta_v)$ and $r(\theta_v) = \begin{pmatrix} \cos\theta_v & \sin\theta_v \\ -\sin\theta_v & \cos\theta_v \end{pmatrix} \in \mathbf{SO}_2(F_v)$.

We call f holomorphic if the function on $\mathfrak{H}^{\mathrm{Hom}(F,\mathbf{C})} = \{x_\infty + iy_\infty \in F \otimes \mathbf{C} \mid y_\infty > 0\}$

$$x_\infty + iy_\infty \mapsto (\psi^n)^{-1}(y)|y|^{-k/2} f\left(\begin{pmatrix} y & x \\ & 1 \end{pmatrix}\right)$$

is holomorphic; in this case such function determines f.

Petersson inner product. — We define a Haar measure dg on $Z(\mathbf{A}_F)\backslash\mathbf{GL}_2(\mathbf{A}_F)$ (where $Z \cong \mathbf{G}_m$ denotes the center of $\mathbf{GL}_2$) as follows. Recall the Iwasawa decomposition

$$\mathbf{GL}_2(\mathbf{A}_F) = B(\mathbf{A}_F)K_0(1)K_\infty \tag{1.1.1}$$

where $K_\infty = \prod_{v|\infty} \mathbf{SO}_2(F_v)$. Let $dk = \otimes_v dk_v$ be the Haar measure on $K = K_0(1)K_\infty$ with volume 1 on each component. Let $dx = \otimes_v dx_v$ be the Haar measure such that dx_v is the usual Lebesgue measure on $\mathbf{R}$ if $v|\infty$, and $\mathscr{O}_{F,v}$ has volume 1 if $v \nmid \infty$. Finally let $d^\times x = \otimes_v d^\times x_v$ on $F_{\mathbf{A}}^\times$ be the product

---

[9] That is, for every g the function $\mathbf{A}^\times \ni y \mapsto f\left(\begin{pmatrix} y & \\ & 1 \end{pmatrix} g\right)$ grows at most polynomially in $|y|$ as $|y| \to \infty$.

[10] Recall the notation $\psi_n = \prod_{v|n} \psi_v$.

of the measures given by $d^\times x_v = |dx_v/x_v|$ if $v\infty$ and by the condition that $\mathcal{O}_{F,v}^\times$ has volume 1 if $v|\infty$. Then we can use the Iwasawa decomposition $g = \begin{pmatrix} z & \\ & z \end{pmatrix} \begin{pmatrix} y & x \\ & 1 \end{pmatrix} k$ to define

$$\int_{Z(\mathbf{A})\backslash \mathbf{GL}_2(\mathbf{A})} f(g)\,dg = \int_{F_{\mathbf{A}}^\times} \int_{\mathbf{A}} \int_K f\left(\begin{pmatrix} y & x \\ & 1 \end{pmatrix} k\right) dk\,dx\,\frac{d^\times y}{|y|}.$$

The Petersson inner product of two forms f_1, f_2 on $GL_2(F)\backslash GL_2(\mathbf{A})$ such that $f_1 f_2$ is invariant under $Z(\mathbf{A})$ is defined by

$$\langle f_1, f_2\rangle_{\mathrm{Pet}} = \int_{Z(\mathbf{A})\backslash \mathbf{GL}_2(\mathbf{A})} \overline{f_1(g)} f_2(g)\,dg$$

whenever this converges (this is ensured if either f_1 or f_2 is a cuspform as defined below). It will be convenient to introduce a level-specific inner product on forms f, g of level N:

$$\langle f, g\rangle_N = \frac{\langle f, g\rangle_{\mathrm{Pet}}}{\mu(N)} \tag{1.1.2}$$

where $\mu(N)$ is the measure of $K_0(N)$.

1.2. Fourier expansion. — Let f be a (not necessarily holomorphic) Hilbert modular form. We can expand it as

$$f(g) = C_f(g) + \sum_{\alpha\in F^\times} W_f\left(\begin{pmatrix} \alpha & \\ & 1 \end{pmatrix} g\right)$$

where

$$C_f(g) = D_F^{-1/2} \int_{\mathbf{A}/F} f\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} g\right) dx,$$

$$W_f(g) = D_F^{-1/2} \int_{\mathbf{A}/F} f\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} g\right) \mathbf{e}(-x)\, dx$$

are called the *constant term* and the *Whittaker function* of f respectively. The form f is called *cuspidal* if its constant term C_f is identically zero. The functions of y obtained by restricting the constant term and the Whittaker function to the elements $\begin{pmatrix} y & \\ & 1 \end{pmatrix}$ are called the *Whittaker coefficients* of f. When f is holomorphic, they vanish unless $y_\infty > 0$ and otherwise have the simple form

$$C_f\left(\begin{pmatrix} y & x \\ & 1 \end{pmatrix}\right) = \tilde{a}^0(f,y) = \psi^n(y)|y|^{k/2} a(f,0),$$

$$W_f\left(\begin{pmatrix} y & x \\ & 1 \end{pmatrix}\right) = \tilde{a}(f,y)\mathbf{e}_\infty(iy_\infty)\mathbf{e}(x) = \psi^n(y)|y|^{k/2} a(f, y^\infty d_F)\mathbf{e}_\infty(iy_\infty)\mathbf{e}(x)$$

for functions $\tilde{a}^0(f,y)$, $\tilde{a}(f,y)$ of $y \in F_{\mathbf{A}}^{\infty,\times}$ which we call the *Whittaker-Fourier coefficients* of f, and a function $a(f,m)$ of the fractional ideals m of F which vanishes on nonintegral ideals whose values are called the **Fourier coefficients** of f. (We will prefer to study Whittaker-Fourier coefficients rather than Fourier coefficients when $n \neq 1$.)

For any $\mathbf{Z}$-submodule A of $\mathbf{C}$, we denote by $M_k(K_1(N,n),\psi,A)$ the space of holomorphic Hilbert modular forms with Fourier coefficients in A of

weight k, level $K_1(N,n)$, and character ψ; and by $S_k(K_1(N,n),\psi,A)$ its subspace of cuspidal forms. When $n=1$, we use the notation $M_k(K_1(N),\psi,A)$ and $S_k(K_1(N),\psi,A)$ and refer to its elements as modular forms of level N and character ψ. When the character ψ is trivial we denote those spaces simply by $M_k(K_0(N),A)$ and $S_k(K_0(N),A)$, whereas linear combinations of forms of level $K_1(N)$ with different characters form the space $M_k(K_1(N),A)$. The notion of Whittaker-Fourier coefficients extends by linearity to the spaces $M_k(K_1(N),\mathbf{C})$.

We can allow more general coefficients: if A is a $\mathbf{Z}[1/N]$-algebra, we define $S_k(K_0(N),A) = S_k(K_0(N),\mathbf{Z}[1/N])\otimes A$; this is well-defined thanks to the q-expansion principle, see [20].

p-adic modular forms. — Let N, P be coprime ideals of $\mathcal{O}_F$, ψ a character of conductor dividing N. If f is a holomorphic form of weight k, level $K_1(NP)$ and prime-to-P character ψ (that is, f is a linear combination of forms of level NP and character $\psi\psi'$ with ψ' a character of conductor dividing P), we associate to it the formal q-expansion

$$\sum_m a_p(f,m)q^m \in \mathbf{C}[[q]]$$

where the sum runs over integral ideals of F, and

$$a_p(f,y^\infty d_F) = \psi^{-1}(y)|y|^{-k/2}\widetilde{a}(f,y).$$

Let $\mathscr{A}$ be a complete $\mathbf{Z}_p$-submodule of $\mathbf{C}_p$, N an ideal prime to p. The space

$$\mathbf{M}_k(K_1(N),\psi,\mathscr{A})$$

of **p-adic modular forms** of weight k, *tame level* $K_1(N)$ and character ψ (with conductor of ψ dividing N) is the subspace of formal q-series with coefficients in $\mathscr{A}$ which are uniform limits of q-expansions of modular forms in $M_k(K_1(Np^\infty), \mathscr{A} \cap \overline{\mathbf{Q}})$ with character whose prime-to-p part is equal to ψ, the norm being the sup norm on q-expansion coefficients. It is a a closed $\mathbf{Z}_p$-submodule in a *p-adic Banach space* (a $\mathbf{Q}_p$- or $\mathbf{C}_p$-vector space complete with respect to a norm compatible with that of $\mathbf{Q}_p$ or $\mathbf{C}_p$). We shall view $M_k(K_1(Np^r, \psi, \mathscr{A})$ as a subset of $\mathbf{M}_k(K_1(N), \psi, \mathscr{A})$ via the q-expansion map.

Similarly we use the notation $\mathbf{S}_k(K_1(N), \psi, \mathscr{A})$, $\mathbf{S}_k(K_0(N), \mathscr{A})$; when $k = 2$ we simply

$$\mathbf{S}_N(\mathscr{A}) = \mathbf{S}_2(K_0(N), \mathscr{A})$$

or just $\mathbf{S}_N$ if $\mathscr{A} = \mathbf{Q}_p$ or $\mathscr{A} = \mathbf{C}_p$ (as understood from context). An element of $\mathbf{S}_N(\mathscr{A})$ is called *bounded* if its Fourier coefficients lie in a bounded subset of $\mathscr{A}$.

1.3. Operators acting on modular forms. — There is a natural action of $\mathbf{Q}[\mathbf{GL}_2(\mathbf{A}^\infty)]$ on modular forms induced by right translation. Here we describe several interesting operators arising from this action.

Let m be an ideal of $\mathscr{O}_F$, $\pi_m \in F^\times_{\mathbf{A}^\infty}$ a generator of $m\widehat{\mathscr{O}}_F$ which is trivial at places not dividing m.

The operator $[m]\colon M_k(K_1(N), \psi) \to M_k(K_1(Nm), \psi)$ is defined by

$$[m]f(g) = \mathbf{N}(m)^{-k/2} f\left(g \begin{pmatrix} 1 & \\ & \pi_m \end{pmatrix}\right). \tag{1.3.1}$$

It acts on Fourier coefficients by

$$a([m]f,n)=a(f,m^{-1}n).$$

For any double coset decomposition

$$K_1(N)\begin{pmatrix}\pi_m & \\ & 1\end{pmatrix}K_1(N)=\coprod_i \gamma_i K_1(N),$$

the **Hecke operator** $T(m)$ is defined by the following level-preserving action on forms f in $M_k(K_1(N))$:

$$T(m)f(g)=\mathbf{N}(m)^{k/2-1}\sum_i f(g\gamma_i);$$

For m prime to N, its effect on Fourier coefficients of forms with trivial character is described by

$$a(T(m)f,n)=\sum_{d|(m,n)}\mathbf{N}(d)^{k/2-1}a(f,mn/d^2).$$

Let $\mathbf{T}_N$ be the (commutative) subalgebra of $\mathrm{End}\,S_k(K_0(N))$ generated by the $T(m)$ for m prime to N. A form f which is an eigenfunction of all the operators in $\mathbf{T}_N$ is called a Hecke *eigenform*. It is called a *primitive* form if moreover it is a newform (see §1.4 below for the definition) and it is normalised by $a(f,1)=1$.

As usual (cf. [30, Lemme 1.10]) we will need the following lemma to ensure the modularity of certain generating functions.

Lemma 1.3.1. — *Let A be a $\mathbf{Z}[1/N]$-algebra. For each linear form*

$$a:\mathbf{T}_N\to A$$

there is a modular form in $S_k(K_0(N), A)$ whose Fourier coefficients are given by $a(T(m))$ for m prime to N. Such a form is unique if we require it to be a newform of some level dividing N (see §1.4 for the definition).

When m divides N, we can pick as double coset representatives the matrices $\gamma_i = \begin{pmatrix} \pi_m & c_i \\ & 1 \end{pmatrix}$ for $\{c_i\} \subset \widehat{\mathcal{O}}_F$ a set of representatives for $\mathcal{O}_F/m\mathcal{O}_F$. Then the operator $T(m)$ is more commonly denoted $U(m)$ and we will follow this practice. It acts on Fourier coefficients of forms with trivial character by

$$a(U(m)f, n) = \mathbf{N}(m)^{k/2-1} a(f, mn).$$

Ordinary projection. — Denoting $U_\wp = U(\wp)$ for $\wp | p$ a prime of $\mathcal{O}_F$, we further define the **ordinary projection** operators

$$e_\wp = \lim U_\wp^{n!} \colon \mathbf{S}_N(\mathbf{C}_p) \to S_{Np}(\mathbf{C}_p)$$

and $e_p = \prod_{\wp | p} e_\wp$. The image is precisely the direct factor of $S_{Np}(\mathbf{C}_p)$ on which $U_\wp$ (respectively U_p) acts by a p-adic unit (the *ordinary* part). See [15, §3] for more details.

Atkin-Lehner theory. — For any nonzero ideal M of $\mathcal{O}_F$, let $W_M \in \mathbf{GL}_2(\mathbf{A}^\infty)$ be a matrix with components

(1.3.2)

$$W_{M,v} = \begin{pmatrix} & 1 \\ -\pi_v^{v(M)} & \end{pmatrix} \quad \text{if } v | M, \qquad W_{M,v} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \quad \text{if } v \nmid M$$

where π_v is a uniformiser at v. We denote by the same name W_M the operator acting on modular forms of level N and trivial character by

$$W_M f(g) = f(g\, W_M);$$

it is self-adjoint for the Petersson inner product, and when M is prime to N it is proportional to the operator $[M]$ of (1.3.1). On the other hand when M equals N, or more generally M divides N and is coprime to NM^{-1}, the operator W_M is an involution and its action is particularly interesting. In this case, extending the definition to forms of level N and character[(11)] $\psi = \psi_{(M)}\psi_{(NM^{-1})}$ with $\psi_{(C)}$ of conductor dividing C, we have

$$W_M f(g) = \psi_{(M)}^{-1}(\det g) f(g\, W_M). \tag{1.3.3}$$

The effect of this action on newforms is described by Atkin-Lehner theory; we summarise it here (in the case $M = N$), referring to [9] for the details.

Let π be an irreducbile infinite-dimensional automorphic representation of $\mathbf{GL}_2(\mathbf{A}_F)$ of central character ψ. Up to scaling, there is a unique *newform* f in the space of π. It is characterised by either of the equivalent properties: (a) it is fixed by a subgroup $K_1(N)$ with N minimal among the N' for which $\pi^{K_1(N')} \neq 0$; (b) its Mellin transform is (a multiple of) the L-function $L(\pi, s)$ of π. In the case of a holomorphic cuspform, this is equivalent to requiring that it belongs to the space of newforms defined in §1.4 below. There is a functional equation relating the L-function $L(s, \pi)$ of π and the L-function $L(1-s, \widetilde{\pi})$ of the contragredient representation; as $\widetilde{\pi} \cong \psi^{-1} \cdot \overline{\pi}$, it translates

[(11)] Notice that a decomposition of ψ as described is only unique up to class group characters (that is, Hecke characters of level one). We will only be using the operator W_M for M a proper divisor of N in a case in which a decomposition is naturally given.

into the following description of the action of W_N on newforms. Suppose that the eigenform $f \in S_k(K_1(N), \psi)$ is a newform in the representation π it generates, then we have

$$W_N f(g) = (-i)^{[F:\mathbf{Q}]k} \tau(f) f^\rho(g) \tag{1.3.4}$$

where f^ρ is the form with coefficients

$$a(f^\rho, m) = \overline{a(f, m)} \tag{1.3.5}$$

and $\tau(f) = \tau(\pi)$ is an algebraic number of complex absolute value 1; it is the central root number of the functional equation for $L(s, \pi)$.

Trace of a modular form. — The **trace** of a modular form f of level ND and trivial character is the form of level N

$$\mathrm{Tr}_{ND/N}(f)(g) = \sum_{\gamma \in K_0(ND)/K_0(N)} f(g\gamma).$$

It is the adjoint of inclusion of forms of level N for the rescaled Petersson product:

$$\langle f, \mathrm{Tr}_{ND/N} g \rangle_N = \langle f, g \rangle_{ND}$$

if f has level N and g has level D.

Suppose that D is squarefree and prime to N, in which case we can write $\mathrm{Tr}_D = \mathrm{Tr}_{ND/N}$ without risk of ambiguity. A set of coset representatives for $K_0(ND)/K_0(N)$ is given by elements $\gamma_{j,\delta}$ for $\delta | D$, $j \in \mathcal{O}_{F,v}/\delta\mathcal{O}_{F,v}$, having

components

$$\gamma_{j,\delta,v} = \begin{pmatrix} 1 & j \\ & 1 \end{pmatrix} \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} = \frac{1}{\pi_v} \begin{pmatrix} \pi_v & j \\ & 1 \end{pmatrix} \begin{pmatrix} & 1 \\ -\pi_v & \end{pmatrix}$$

at places $v|\delta$, and $\gamma_{j,\delta,v} = 1$ everywhere else. From the second decomposition given just above, if f has weight 2 we obtain

$$a(\mathrm{Tr}_D(f), m) = \sum_{\delta|D} a(U(\delta) f^{(\delta)}, m) = \sum_{\delta|D} a(f^{\delta}, m\delta) \tag{1.3.6}$$

where $f^{(\delta)}(g) = f(g\,W_\delta)$ with W_δ as in (1.3.2).

Notice that if f is a modular form of level ND and character ψ, then $f^{(\delta)}(g) = f(g\,W_\delta)$ is a modular form of level $K_1(ND, \delta)$ of characrers $\psi_1 = \psi^\delta = \prod_{v\nmid\delta} \psi_v$, $\psi_2 = \psi_\delta = \prod_{v|\delta} \psi_v$.

Remark 1.3.2. — If D is prime to p, the various trace operators Tr_{NDp^r/Np^r} extend to a continuous operator $\mathrm{Tr}_{ND/N}$ on p-adic modular forms of tame level ND. Similarly the operators $[m]$, $T(m)$ and W_m for m prime to Np extend to continuous operators on p-adic modular forms of tame level N.

1.4. Fourier coefficients of old forms. — As we will study modular forms through their Fourier coefficients, we give here a criterion for recognising the coefficients of certain old forms.[(12)] Let N, P be coprime ideals of $\mathcal{O}_F$. The space $S_{NP}^{N\text{-old}} \subset S_{NP}$ is the space spanned by forms $f = [d]f'$ for some $1 \neq d|N$ and some cuspform f' of level $N'P$ with $N'|d^{-1}N$. In the case $P = 1$, we define the space of **newforms** of level dividing N to be the orthogonal

[(12)] Cf. [45, §4.4.4].

to the space of old forms for the Petersson inner product. We denote by $\mathbf{S}_N^{\mathrm{old}} \subset \mathbf{S}_N$ the closed subspace generated by the image of $S_{Np^\infty}^{N\text{-old}}$ in $\mathbf{S}_N$.

(As above, the coefficient ring will always be $\mathbf{Q}_p$ or an algebraic extension of it or $\mathbf{C}_p$ as understood from context.)

Let now $\mathscr{S}$ be the space of functions $f : \mathbf{N}_F \to \mathscr{A}$ modulo those for which there is an ideal M prime to p such that $f(n) = 0$ for all n prime to M. A function $f \in \mathscr{S}$ is called *multiplicative* if it satisfies[13] $f(mn) = f(m)f(n)$ for all $(m,n) = 1$. For h a multiplicative function, a function f is called an *h-derivative* if it satisfies $f(mn) = h(m)f(n) + h(n)f(m)$ for all $(m,n) = 1$.

Let σ_1 and r be the multiplicative elements of $\mathscr{S}$ defined by

$$\sigma_1(m) = \sum_{d|n} \mathbf{N}(d), \qquad r(m) = \sum_{d|m} \varepsilon_{E/F}(d)$$

(where E is a totally imaginary quadratic extension of F of discriminant prime to p).[14] We define a subspace $\mathscr{D}_N \subset \mathscr{S}$ to be generated by σ_1, r, σ_1-derivatives, r-derivatives, and Fourier coefficients of forms in $\mathbf{S}_N^{\mathrm{old}}$.

Lemma 1.4.1. — *The q-expansion map $\mathbf{S}_N/\mathbf{S}_N^{\mathrm{old}} \to \mathscr{S}/\mathscr{D}_N$ is injective.*

The proof is similar to that of [45, Proposition 4.5.1].

Remark 1.4.2. — The operators $U_\wp$ for $\wp | p$ extend to opeators on $\mathscr{S}$ via $U_\wp f(m) = f(m\wp)$. The ordinary projection operators $e_\wp$ and $e = e_p$ of course do not extend, since the limit may not exist; however the kernels $\mathrm{Ker}(e), \mathrm{Ker}(e') \subset \mathscr{S}$ are well-defined.

[13] This relation and the following are of course to be understood to hold in $\mathscr{S}$.
[14] We will see below that σ_1 and r are the Fourier coefficients of weight 1 Eisenstein series and theta series.

1.5. The functional L_{f_0}. — Recall from the Introduction that we have fixed a primitive ordinary Hilbert modular form f of level $K_0(N)$, which is ordinary at all primes $\wp$ dividing p. If $\alpha_\wp$ is the unit root of the $\wp^{\text{th}}$ Hecke polynomial of f (cf. the Introduction) and the operator $[\wp]$ is as in (1.3.1), then the *p-stabilisation* of f is

$$f_0 = \prod_{\wp|p} \left(1 - \frac{\mathbf{N}\wp}{\alpha_\wp}[\wp]\right) f,$$

a form of level $K_0(Np)$ satisfying $U_\wp f_0 = \alpha_\wp f_0$ for all $\wp|p$.

We define a functional, first introduced by Hida, which plays the role of projection onto the f-component. Both sides of our main formula will be images of p-adic modular forms under this operator.

Let P be an ideal of $\mathcal{O}_F$ divisible exactly by the primes $\wp|p$. For a form $g \in M_2(K_0(NP))$ with $r \geq 1$, let

$$L_{f_0}(g) = \frac{\langle W_{NP} f_0^\rho, g\rangle}{\langle W_{NP} f_0^\rho, f_0\rangle}.$$

Lemma 1.5.1. — *The above formula defines a bounded linear functional*

$$L_{f_0} : M_2(K_0(Np^\infty), \overline{\mathbf{Q}}) \to \overline{\mathbf{Q}}$$

satisfying the following properties:

1. *On $M_2(K_0(N))$ we have*

$$L_{f_0} = \prod_{\wp|p} \left(1 - \frac{\mathbf{N}\wp}{\alpha_\wp(f)^2}\right)^{-1} \mathbf{1}_f$$

where $\mathbf{1}_f(g) = \langle f, g\rangle\langle f, f\rangle$.

2. *On $M_2(K_0(N\wp^r))$ we have, for each nonnegative $t \leq r-1$,*

$$L_{f_0} \circ U_\wp^t = \alpha_\wp(f)^t L_{f_0}.$$

3. *It admits an extension to p-adic modular forms still denoted*

$$L_{f_0} : \mathbf{M}_N(\mathbf{C}_p) \to \mathbf{C}_p$$

which is continuous and factors through the ordinary projection operator e.

The proof of these facts is similar to the case of elliptic modular forms, see e.g. [29].

We define spaces $\overline{\mathbf{S}}_N = \mathbf{S}_N/(\mathbf{S}_N^{\mathrm{old}} + \mathrm{Ker}(e))$ and $\overline{\mathscr{S}} = \mathscr{S}/(\mathscr{D}_N + \mathrm{Ker}(e))$. The former injects into the latter, and we denote its image by $\overline{\mathscr{S}}_N$. Note in passing that elements of $\overline{\mathscr{S}}$ are invariant under U_p and therefore are determined by their values on prime-to-p ideals (or on ideals divisible by p).

By the pevious lemma, the operator L_{f_0} extends to a bounded operator, still denoted by the same name,

$$L_{f_0} : \overline{\mathscr{S}}_N \to \mathbf{C}_p.$$

It is defined over $\mathbf{Q}_p$ in the sense that it takes $\mathbf{Q}_p$-valued elements of $\overline{\mathscr{S}}_N$ to $\mathbf{Q}_p$.

2. Theta measure

We construct a measure on the Galois group of the maximal abelian extension of E unramified outside p with values in p-adic theta series, and compute its Fourier expansion.

2.1. Weil representation. — We first define the Weil representation. See [6, §4.8] for an introduction, and [40] or [43] for our conventions on the representation for similitude groups.

Local setting. — Let $V=(V,q)$ be a quadratic space over a local field F of characteristic not 2, with a quadratic form q; we choose a nontrivial additive character $\mathbf{e}$ of F. For simplicity we assume V has even dimension. For $u\in F^\times$, we denote by V_u the quadratic space (V,uq). We let $\mathbf{GL}_2(F)\times\mathbf{GO}(V)$ act on the space $\mathscr{S}(V\times F^\times)$ of Schwartz functions as follows (here $\nu\colon \mathbf{GO}(V)\to\mathbf{G}_m$ denotes the similitude character):

- $r(h)\phi(t,u)=\phi(h^{-1}t,\nu(h)u)\quad$ for $h\in\mathbf{GO}(V)$;
- $r(n(x))\phi(t,u)=\mathbf{e}(xuq(t))\phi(t,u)\quad$ for $n(x)=\begin{pmatrix}1 & x\\ & 1\end{pmatrix}\in\mathbf{GL}_2$;
- $r\left(\begin{pmatrix}a & \\ & d\end{pmatrix}\right)\phi(t,u)=\chi_V(a)\left|\frac{a}{d}\right|^{\frac{\dim V}{4}}\phi(at,d^{-1}a^{-1}u)$;
- $r(w)\phi(x,u)=\gamma(V_u)\hat{\phi}(x,u)$ for $w=\begin{pmatrix}& 1\\ -1 & \end{pmatrix}$.

Here χ_V is the quadratic character associated to V, $\gamma(V_u)$ is a certain square root of $\chi(-1)$, and $\hat{\phi}$ denotes the Fourier transform in the first variable

$$\hat{\phi}(x,u)=\int_V\phi(y,u)\mathbf{e}(-u\langle x,y\rangle)dy$$

where $\langle\cdot,\cdot\rangle$ is the bilinear form associated to q and dy is the self-dual Haar measure.

Global setting. — Given a quadratic space (V,q) over a global field F of characteristic not 2 (and a nontrivial additive character $\mathbf{e}\colon F\backslash\mathbf{A}_F\to\mathbf{C}^\times$), the Weil representation is the restricted tensor product r of the associated local

Weil representations, with spherical functions $\phi_v(t,u) = \mathbf{1}_{\mathscr{V}_v \times \mathscr{O}_{F,v}^\times}(x,u)$ for some choice of lattices $\mathscr{V}_v \subset V(F_v)$.

The case of interest to us is the following: F is a totally real number field, $V = (E, \mathfrak{N})$ is given by a quadratic CM extension E/F with the norm form $\mathfrak{N} = N_{E/F}$ and the lattices $\mathscr{O}_{E,v} \subset E_v$, and the additive character $\mathbf{e}$ is the standard one. We denote $G = \mathbf{GL}_2$, $H = \mathbf{GSO}(V)$, two algebraic groups defined over F; we have $H \cong \mathrm{Res}_{E/F}\mathbf{G}_m$. In this case we have

$$\chi_V = \varepsilon_{E/F} = \varepsilon,$$

where $\varepsilon_{E/F}$ is the quadratic character of $F_{\mathbf{A}}^\times$ associated to the extension E/F. The self-dual measure on E_v is the one giving $\mathscr{O}_{E,v}$ volume $|\mathscr{O}_{E,v}/\mathfrak{D}_v|^{-1/2}$ where $\mathfrak{D}_v$ is the relative inverse different Moreover the constant γ can be explicitly described (see [7, §§ 38.6, 30.4, 23.5]): in the case $v|\Delta_{E/F}$, which is the only one we will be using, such description is in terms of a local Gauß sum $\kappa(v)$:[15]

(2.1.1)
$$\gamma(E_v, u\mathfrak{N}) = \varepsilon_v(u)\kappa(v) = \varepsilon_v(u)|\pi_v|^{\frac{1}{2}} \sum_{x \in (\mathscr{O}_{F,v}/\pi_v\mathscr{O}_{F,v})^\times} \varepsilon(x/\pi_v)\mathbf{e}_v(x/\pi_v).$$

2.2. Theta series. — We define the **theta kernel** to be

$$\theta_\phi(g,h) = \sum_{(t,u) \in V \times F^\times} r(g,h)\phi(t,u)$$

which is an automorphic form for the group $\mathbf{GL}_2(F)\backslash\mathbf{GL}_2(\mathbf{A}_F) \times \mathbf{GO}(V)\backslash\mathbf{GO}(V_{\mathbf{A}_F})$.

If $\mathscr{W}$ is an automorphic function for H which is trivial at infinity (which is the same thing as a linear combination of finite order Hecke characters of

[15] Our $\kappa(v)$ is denoted $\kappa(v)^{-1}$ in [45].

E), we define the **theta series**[16]

$$\theta_\phi(\mathscr{W})(g) = \int_{H(F)\backslash H(\mathbf{A}_F)} \mathscr{W}(h^{-1})\theta_\phi(g,h)\,dh$$

which is an automorphic form on G. Here the measure dh is the product of the measure on $H(\mathbf{A}^\infty)$ which gives volume 1 to the compact $U_0 = \widehat{\mathcal{O}}_E^\times$, and any fixed measure[17] on $H(\mathbf{A}_\infty)$.

Let us explain how to explicitly compute the integral in our situation. For each open compact subgroup $U \subset H(\mathbf{A}_F^\infty) = E_{\mathbf{A}^\infty}$, we have exact sequences

$$1 \to \mathcal{O}_{E,U}^\times \backslash U E_\infty^\times \to E^\times \backslash E_{\mathbf{A}}^\times \to E^\times U \backslash E_{\mathbf{A}^\infty}^\times \to 1$$

and

$$1 \to \mu(U)\backslash U E_\infty^1 \to \mathcal{O}_{E,U}^\times \backslash U E_\infty^\times \xrightarrow{\mathfrak{N}_\infty} N(\mathcal{O}_{E,U}^\times)\backslash F_\infty^+ \to 1.$$

The notation used is the following: $\mathcal{O}_{E,U}^\times = E^\times \cap U \supset \mu(U) =$ the subset of roots of unity, $\mathfrak{N}_\infty : E_\infty^\times \to F_\infty^+$ is the norm map at the infinite places and E_∞^1 is its kernel.

We can choose a splitting ι of the first sequence, for example

$$\iota : E^\times U \backslash E_{\mathbf{A}^\infty}^\times \cong E^\times U \backslash (E_{\mathbf{A}}^\times)^{1,\|} \hookrightarrow E^\times \backslash E_{\mathbf{A}}^\times,$$

where $(E_{\mathbf{A}}^\times)^{1,\|}$ denotes the set of idèles of adelic norm 1 with infinity component $h_\infty = (h,\dots,h)$ for some real number $h > 0$ and the isomorphism

[16] The reason for taking $\mathscr{W}(h^{-1})$ rather than $\mathscr{W}(h)$ is that we want $\theta_\phi(\mathscr{W})$ to be the series classically denoted $\Theta(\mathscr{W})$ for a suitable choice of ϕ – this will be clear from the computations below.

[17] There will be no ambiguity since later we will choose ϕ_∞ to be again any fixed Schwartz function, whose integral over $H(\mathbf{A}_\infty)$ with respect to the chosen measure is a specified function $\overline{\phi}_\infty$.

is the unique one which gives the identity once composed with projection onto the finite part.

We begin to expand the series, evaluating the integral as explained above and exploiting the fact that the action of $H(F_\infty) = E_\infty^\times$ on $\phi(t,u)$ factors through the norm. We take U to be small enough so that $\mathscr{W}$ and ϕ are invariant under U, and denote

$$\overline{\phi}_v(t,u) = \int_{H(F_v)} r(h)\phi_v(t,u)\,dh \qquad \text{if } v|\infty$$

and $\overline{\phi} = \prod_{v\nmid\infty}\phi_v \prod_{v|\infty}\overline{\phi}_v$. A specific choice of $\overline{\phi}_v$ will be made shortly: for the moment we just record, and use in the following computation, that we will take $u \mapsto \overline{\phi}_v(t,u)$ to be supported on $\mathbf{R}^+$.

We have

$$\theta_\phi(\mathscr{W})(g) = \int_{E^\times\backslash E_{\mathbf{A}}^\times} \mathscr{W}(h^{-1})\theta_\phi(g,h)\,dh$$

$$= w_U^{-1}\int_U \int_{E_\infty^1}\int_{\mathfrak{N}(\mathscr{O}_{E,U}^\times)\backslash F_\infty^+}\int_{E^\times U\backslash E_{\mathbf{A}^\infty}^\times} \mathscr{W}(\iota(a)^{-1}) \sum_{(t,u)\in E\times F^\times} r(g,\iota(a)h)\phi(t,u)\,da\,dh$$

Here $w_U = |\mu(U)|$ and dh denotes the measure on $U\times E_\infty^1\times F_\infty^+ = U\times H(F_\infty)$. We partially collapse the integral over $\mathfrak{N}(\mathscr{O}_{E,U}^\times)\backslash F_\infty^+$ and the sum over $u\in F^\times$ and use our choice of ϕ_∞ to get

(2.2.1)

$$= w_U^{-1}\mathrm{vol}(U)\int_{E^\times U\backslash E_{\mathbf{A}^\infty}^\times} \mathscr{W}(\iota(a)^{-1}) \sum_{u\in\mathfrak{N}(\mathscr{O}_{E,U}^\times)\backslash F^+}\sum_{t\in E} r(g,\iota(a))\overline{\phi}(t,u)\,da$$

$$= w^{-1}\frac{h}{h_U}\int_{E^\times U\backslash E_{\mathbf{A}^\infty}^\times} \mathscr{W}(\iota(a)^{-1})\nu_U \sum_{u\in\mathfrak{N}(\mathscr{O}_{E,U}^\times)\backslash F^+}\sum_{t\in E} r(g,\iota(a))\overline{\phi}(t,u)\,da$$

Here in the last step we have defined $\nu_U = [\mathfrak{N}(\mathscr{O}_E^\times) : \mathfrak{N}(\mathscr{O}_{E,U}^\times)]$ and computed $\mathrm{vol}(U) = \mathrm{vol}(U_0)(h/h_U)(w_U/w)\nu_U^{-1}$, where $U_0 = \widehat{\mathscr{O}}_E^\times$, $h_U = |E^\times U \backslash E_{A^\infty}^\times|$, $h = h_{U_0}$, $w = w_{U_0}$. Recall that our measure is such that $\mathrm{vol}(U_0) = 1$. The remaining integral is just a finite sum.

The sum over u is actually finite owing to the integrality constraints imposed by ϕ at finite places.[18]

2.3. Theta measure. — We define a measure with values in p-adic modular forms on the group

$$\mathscr{G}' = \mathrm{Gal}(E'_\infty/E) \cong \varprojlim E^\times U_{p^n} \backslash E_{\mathbf{A}^\infty}^\times$$

where E'_∞ is the maximal abelian extension of E unramified outside p, that is, the union of the ray class fields of E of p-power ray $U_{p^n} = \prod_v \{\text{units} \equiv 1 \bmod p^n \mathscr{O}_{E,v}\}$ and the isomorphism is given by class field theory. The topology is the profinite topology.

Recall that a **measure** on a topological space $\mathscr{G}$ with values in a p-adic Banach space $\mathbf{M}$ is a $\overline{\mathbf{Q}}_p$-linear functional

$$\mu : \mathscr{C}(\mathscr{G}, \mathbf{C}_p) \to \mathbf{M}$$

on continuous $\mathbf{C}_p$-valued functions, which is continuous (equivalently, bounded) with respect to the sup norm on $\mathscr{C}(\mathscr{G}, \mathbf{C}_p)$. The linearity property will be called distributional property in what follows. The boundedness property will in each case at hand be verified on the set of p-adic characters of $\mathscr{G}$,

[18] We will see this in more detail shortly. We are also using the definition of $\overline{\phi}_\infty$ in order to freely replace the sum over $u \in F^\times$ with a sum over $u \in F^+$ – in fact a slight variation would be necessary when $\det g_\infty \notin F_\infty^+$, but this is a situation we won't encounter.

which in our cases generates the whole of $\mathscr{C}(\mathscr{G}, \mathbf{C}_p)$ (classically, the continuity of μ goes under the name of *abstract Kummer congruences* for μ).

When $\mathbf{M} = \mathbf{M}_0 \otimes_{\mathbf{Q}_p} \mathbf{C}_p$ for a p-adic Banach space $\mathbf{M}_0$ over $\mathbf{Q}_p$, the measure μ is said to be *defined over* $\mathbf{Q}_p$ if $\mu(\mathscr{W}) \in \mathbf{M}_0 \otimes \mathbf{Q}_p(\mathscr{W})$ whenever the function $\mathscr{W}$ on $\mathscr{G}$ has values in $\mathbf{Q}_p(\mathscr{W}) \subset \overline{\mathbf{Q}}_p \subset \mathbf{C}_p$.

Definition 2.3.1. — The **theta measure** $d\Theta$ on $\mathscr{G}'$ is defined by

$$\Theta(\mathscr{W}) = \int_{\mathscr{G}'} \mathscr{W}(\sigma)\, d\Theta(\sigma) = \theta_\phi(\mathscr{W}),$$

for any function $\mathscr{W} : \mathscr{G}' \to \overline{\mathbf{Q}}$ factoring through a finite quotient of $\mathscr{G}'$, where the function ϕ is chosen as follows:

- for $v \nmid p\infty$, $\phi_v(t,u) = \mathbf{1}_{\mathscr{O}_{E,v}}(t)\mathbf{1}_{d_{F_v}^{-1,\times}}(u)$;
- for $v|p$,
$$\phi_v(t,u) = [\mathscr{O}_{E,v}^\times : U_v']\mathbf{1}_{U_v'}(t)\mathbf{1}_{d_{F_v}^{-1,\times}}(u),$$
where $U_v' \subset \mathscr{O}_{E,v}^\times$ is any small enough compact set – that is, $U_v' \subset U_v$ if $\mathscr{W}$ is invariant under $U = \prod_v U_v$, and the definition does not depend on the choice of U_v. (In practice, we will choose $U_v' = U_v$ if U_v is maximal with respect to the property just mentioned.)
- for $v|\infty$, $\phi_v(t,u)$ is a Schwartz function such that
$$\int_{H(F_v)} r(h)\phi_v(t,u)\, dh = \overline{\phi}_v(t,u) = 1_{\mathbf{R}^+}(u)\exp(-2\pi u N(t)).$$
(See [43, 4.1] for more details on this choice.)

In Corollary 2.4.3 below we will show that this in fact defines a measure on $\mathscr{G}'$ with values in p-adic Hilbert modular forms of weight one, tame level $\Delta_{E/F}$ and character ε.

2.4. Fourier expansion of the theta measure / I. — We compute the Fourier expansion of the theta measure on $\mathscr{G}'$, carrying on the calculation started in §2.2.

In the case where $g = \begin{pmatrix} y & x \\ & 1 \end{pmatrix}$ with $y_\infty > 0$, the sum over (u,t) in (2.2.1) evaluates to

$$\varepsilon(y)|y|^{1/2}\sum_{u,t}\phi^\infty(a^{-1}yt,\mathfrak{N}(a)y^{-1}u)\mathbf{e}_\infty(iy_\infty uN(t))\mathbf{e}(xu\mathfrak{N}(t)). \tag{2.4.1}$$

Then we compute the sum of this expression over the finite quotient $\mathscr{G}'_U$ of $\mathscr{G}'$, with $\mathscr{G}'_U \cong E^\times U\backslash E^\times_{\mathbf{A}^\infty}$.

We assume $\mathscr{W}$ is a character so $\mathscr{W}(a^{-1}) = \overline{W}(a)$ where $\overline{\mathscr{W}} = \mathscr{W}^{-1}$.

First we pre-compute the product of all the constants appearing in the theta series of (2.2.1), including the one from ϕ – we take

$$\phi_v(t,u) = [\mathscr{O}^\times_{E,v} : U_v]\mathbf{1}_{U_v}(t)\mathbf{1}_{\mathscr{O}_F^\times}(u),$$

so:

$$\begin{aligned} w\frac{h}{h_U}\nu_U[\mathscr{O}^\times_{E,v} : U_v] &= w[\mathscr{O}_E^\times\backslash\widehat{\mathscr{O}}^\times_{E,v} : \mathscr{O}^\times_{E,U}\backslash U]^{-1}[\mathfrak{N}(\mathscr{O}_E^\times) : \mathfrak{N}(\mathscr{O}^\times_{E,U})]^{-1}[\widehat{\mathscr{O}}^\times_E : U] \\ &= w[\mu(\mathscr{O}_E) : \mu(\mathscr{O}_{E,U}] = w_U^{-1}. \end{aligned}$$

This computation together with (2.2.1), (2.4.1) gives

$$\begin{aligned} \Theta(\mathscr{W}) = \varepsilon(y)|y|^{\frac{1}{2}}w_U^{-1}\sum_{a\in E^\times U\backslash E^\times_{\mathbf{A}^\infty}}\overline{\mathscr{W}}(a)\sum_{t\in E, u\in\mathfrak{N}(\mathscr{O}^\times_{E,U})\backslash F^+}\phi^{p\infty}(a^{-1}yt,\mathfrak{N}(a)y^{-1}u) \\ \times\mathbf{1}_{\mathscr{O}^\times_{E,U,p}}(a^{-1}yt)\mathbf{1}_{d_{F_p}^{-1,\times}}(\mathfrak{N}(a)y^{-1}u)\,\mathbf{e}_\infty(iy_\infty u\mathfrak{N}(t))\,\mathbf{e}(xu\mathfrak{N}(t)) \end{aligned}$$

$$=\varepsilon(y)\mathscr{W}(y)|y|^{\frac{1}{2}}w_U^{-1}\sum_{a\in E^\times U\backslash E_{\mathbf{A}^\infty}^\times}\overline{\mathscr{W}}(a)\sum_{t\in E,u\in\mathfrak{N}(\mathscr{O}_{E,U}^\times)\backslash F^+}\mathbf{1}_{\widehat{\mathscr{O}_{E,U}}\cap\mathscr{O}_{E,U,p}^\times}(a^{-1}t)$$
$$\times\,\mathbf{1}[\mathfrak{N}(a)yu\mathscr{O}_F=d_F^{-1}]\,\mathbf{e}_\infty(iy_\infty u\mathfrak{N}(t))\,\mathbf{e}(xu\mathfrak{N}(t))$$

where we have made the change of variable $a\to ay$.

Now we make the substitution $u\mathfrak{N}(t)=\xi$ and observe that the contribution to the ξ^{th} term is equal to zero if $(\xi yd_F,p)\neq 1$, and otherwise it equals $\overline{\mathscr{W}}(a)$ times the cardinality of the set

$$R_{a^{-1}}(\xi,y)=\left\{(t,u)\in\mathscr{O}_{E,U}\times F^+\,|\,u\mathfrak{N}(t)=\xi,\mathfrak{N}(t/a)\mathscr{O}_F=\xi yd_F\right\}\Big/\mathfrak{N}\left(\mathscr{O}_{E,U}^\times\right),$$

which admits a surjection $\pi:(t,u)\mapsto a^{-1}t\mathscr{O}_{E,U}$ to the set $\mathbf{r}_{a^{-1}}(\xi yd_F)$ of proper ideals $\mathfrak{b}\subset\mathscr{O}_{E,U}$ in the U-class a^{-1}, whose norm is $\mathfrak{N}(\mathfrak{b})=\xi yd_F$. The fibres of π are in bijection with $\mathscr{O}_{E,U}^\times/\mathfrak{N}(\mathscr{O}_{E,U}^\times)$ which has cardinality w_U. We deduce the following description of the Fourier coefficients of $\Theta(\mathscr{W})$.

Proposition 2.4.1. — *The series $\Theta(\mathscr{W})$ belongs to $S_1(K_1(N),\varepsilon\mathscr{W}|_{F_{\mathbf{A}}^\times})$, where $\Delta(\mathscr{W})=\Delta\mathfrak{N}(\mathfrak{f}(\mathscr{W}))$. Its Fourier coefficients are given by*

$$a(\Theta(\mathscr{W}),m)=\sum_{\substack{\mathfrak{b}\subset\mathscr{O}_{E,U}\\ \mathfrak{N}(\mathfrak{b})=m}}\mathscr{W}(\mathfrak{b})=r_{\mathscr{W}}(m)$$

for $(m,p)=1$ and vanish for $(m,p)\neq 1$.

Remark 2.4.2. — If $\mathscr{W}$ is ramified at all places $w|p$, the Mellin transform of $\Theta(\mathscr{W})$ is thus precisely the L-function

$$L(\mathscr{W},s)=\sum_{(\mathfrak{b},p)=1}\mathscr{W}(\mathfrak{b})\mathbf{N}(\mathfrak{b})^{-s}=\sum_m r_{\mathscr{W}}(m)\mathbf{N}(m)^{-s}.$$

Therefore by Atkin-Lehner theory $\Theta(\mathscr{W})$ is a newform and according to (1.3.4) we have a functional equation

$$W_{\Delta(\mathscr{W})}\Theta(\mathscr{W})(g) = (-i)^{[F:\mathbf{Q}]}\tau(\mathscr{W})\Theta(\overline{\mathscr{W}})$$

where $\tau(\mathscr{W})$ is an algebraic number of absolute value 1 (essentially a Gauß sum).

Corollary 2.4.3. — *The functional Θ of Definition 2.3.1 is a measure on $\mathscr{G}'$ with values in $\mathbf{S}_1(K_1(\Delta), \varepsilon)$, defined over $\mathbf{Q}_p$.*

Proof. — The distributional property is obvious from the construction or can be seen from the q-expansion given above, from which boundedness is also clear – cf. also [16, Theorem 6.2]. □

2.5. Fourier expansion of the theta measure / II. — For later use in computing the trace of the convolution of the theta measure with the Eisenstein measure (defined below), we need to consider the expansion of $\Theta(\mathscr{W})^{(\delta)}(g) = \Theta(\mathscr{W})(g\,W_\delta)$ for $g = \begin{pmatrix} y & x \\ & 1 \end{pmatrix}$; for such a g we have

$$\begin{pmatrix} y & x \\ & 1 \end{pmatrix} W_\delta = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & \pi_\delta \end{pmatrix} w_\delta$$

where π_δ is an idèle with components π_v at $v|\delta$ and 1 everywhere else. Here π_v is a uniformiser chosen to satisfy $\varepsilon(\pi_v) = 1$.

The modular form $\Theta(\mathscr{W})^{\delta}$ can be expanded in the same way as in §2.4, except that for $v|\delta$ we need to replace $\phi_v(t,u)=\mathbf{1}_{\mathscr{O}_{E,v}}(t)\mathbf{1}_{d_F^{-1,\times}}(u)$ by

$$W_\delta\phi_v(t,u)=\begin{pmatrix}1 & \\ & \pi_v\end{pmatrix}\gamma(u)\widehat{\mathbf{1}_{\mathscr{O}_{E,v}}}(t)\mathbf{1}_{d_{F,v}^{-1,\times}}(u)$$
$$=\varepsilon_v(\pi_\delta^{-1}u)\kappa(v)\mathbf{1}_{\mathfrak{D}_v^{-1}}(t)\mathbf{1}_{d_F^{-1,\times}}(\pi_\delta^{-1}u)$$

Here recall that $\mathfrak{D}$ is the relative inverse different of E/F; and that w acts as Fourier transform in t with respect to the quadratic form associated to $u\mathfrak{N}$, with the normalising constant $\gamma(u)=\gamma(E_v,u\mathfrak{N})$ as described in (2.1.1).

The computation of the expansion can then be performed exactly as in §2.4. We omit the details but indicate that the relevant substitution is now $a\to\pi_{\mathfrak{d}}ay$, where $\mathfrak{d}$ is an ideal of $\mathscr{O}_E$ of norm δ and $\pi_{\mathfrak{d}}\in\widehat{\mathscr{O}_E}$ is a generator with components equal to 1 away from $\mathfrak{d}$.

Proposition 2.5.1. — *The series $\Theta(\mathscr{W})^{(\delta)}$ belongs to $S_1(K_1(\Delta(\mathscr{W}),\delta),\varepsilon_{\mathscr{W}}|_{F_{\mathbf{A}}^\times})$. Its Whittaker-Fourier coefficients are given by*

$$\tilde{a}(\Theta(\mathscr{W})^{(\delta)},y)=\varepsilon_{\mathscr{W}}(y)|y|^{1/2}\kappa(\delta)\mathscr{W}(\mathfrak{d})\varepsilon_\delta(y)r_{\mathscr{W}}(yd_F),$$

where $\kappa(\delta)=\prod_{v|\delta}\kappa(v)$.

3. Eisenstein measure

In this section we construct a measure (cf. §2.3) valued in Eisenstein series of weight one, and compute its Fourier expansion.

3.1. Eisenstein series. — Let k be a positive integer, M an ideal of $\mathscr{O}_F$, and $\varphi\colon F_{\mathbf{A}}^\times/F^\times\to\mathbf{C}^\times$ a finite order character of conductor dividing M satisfying

$\varphi_v(-1)=(-1)^k$ for $v|\infty$. Let

$$L^M(s,\varphi)=\sum_{(m,M)=1}\varphi(m)\mathbf{N}(m)^{-s} \tag{3.1.1}$$

where the sum runs over all nonzero ideals of $\mathscr{O}_F$.

Let $B\subset \mathbf{GL}_2$ be the Borel subgroup of upper triangular matrices; recall the notation from §1.1, and the Iwasawa decomposition (1.1.1); the decomposition is not unique but the ideal of $\widehat{\mathscr{O}}_F$ generated by the lower left entry of the $K_0(1)$-component is well-defined.

For $s\in\mathbf{C}$, define a function $H_{k,s}(g,\varphi)$ on $\mathbf{GL}_2(\mathbf{A}_F)$ by

$$H_{k,s}(g=qur(\theta);\varphi)=\begin{cases}\left|\frac{y_1}{y_2}\right|^s\varphi^{-1}(y_2d)\mathbf{e}_\infty(k\theta) & \text{if } u=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in K_0(M)\\ 0 & \text{if } u\in K_0(1)\setminus K_0(M_0).\end{cases}$$

where we have written $g=qur(\theta)$ with $q=\begin{pmatrix}y_1 & x\\ & y_2\end{pmatrix}\in B(\mathbf{A}_F)$, $u\in K_0(1)$, $r(\theta)\in K_\infty$.

We define two **Eisenstein series**

$$E_k^M(g,s;\varphi)=L^M(2s,\varphi)\sum_{\gamma\in B(F)\backslash\mathbf{GL}2(F)}H_{k,s}(\gamma g;\varphi),$$
$$\widetilde{E}_k^M(g,s;\varphi)=W_ME_k^M(g,s;\varphi)=\varphi^{-1}(\det g)E_k^M(g\,W_M,s;\varphi)$$

which are absolutely convergent for $\Re s>1$ and continue analytically for all s to (non-holomorphic) automorphic forms of level M, parallel weight k and characters φ^{-1} for E and φ for $\widetilde{E}$. Here W_M is as in (1.3.3). The superscript M will be omitted from the notation when its value is clear from context.

3.2. Fourier expansion of the Eisenstein measure. — We specialize to the case where k is odd, $M = \Delta P$ with $(\Delta, P) = 1$, $\varphi = \varepsilon\phi$ with $\varepsilon = \varepsilon_{E/F}$ and ϕ a nontrivial character of conductor dividing P, trivial at infinity (in particular we have $\varphi_v(-1) = \varepsilon_v(-1)\phi_v(-1) = -1$ as required). We assume that Δ is squarefree. For $\delta|\Delta$ we compute[(19)] the Whittaker coefficients (cf. §1.2; we suppress φ, M and k from the notation) of $\widetilde{E}^{(\delta)}$;

$$c_s^\delta(\alpha, y) = D_F^{-1/2} \int_{\mathbf{A}_F/F} \widetilde{E}\left(\begin{pmatrix} y & x \\ & 1 \end{pmatrix} W_\delta, s\right) \mathbf{e}(-\alpha x)\, dx$$

for $\alpha \in F$ and δ dividing Δ; since $c_s(\alpha, y) = c_s(1, \alpha y)$ for $\alpha \neq 0$, we can restrict to $\alpha = 0$ or 1.

Proposition 3.2.1. — *In the case just described, the Whittaker coefficients $c_s^\delta(\alpha, y)$ of the Eisenstein series $\widetilde{E}_k^{(\delta)}(g, s; \varphi)$ are given by $c_s^\delta(0, y) = 0$, and*

$$c_s^\delta(1, y) = \frac{\mathbf{N}(\delta)^{s-1/2}}{D_F^{1/2}\mathbf{N}(\Delta P)^s} \varepsilon\phi(y)|y|^{1-s}\kappa(\delta)\phi(\delta)\varepsilon_\delta(y)\phi_\delta(y^\infty d_F)|y\delta d_F|_\delta^{2s-1}\sigma_{k,s,\varepsilon\phi}(y)$$

if $y d_F$ is integral, and $c_s^\delta(1, y) = 0$ otherwise,

where $\kappa(\delta) = \prod_{v|\delta} \kappa(v)$ with $\kappa(v)$ as in (2.1.1) *and*

$$\sigma_{k,s,\varphi}(y) = \prod_{v \nmid \Delta M \infty} \sum_{n=0}^{v(y d_F)} \varphi_v(\pi_v)^n |\pi_v|^{n(2s-1)} \prod_{v|\infty} V_{k,s}(y_v)$$

with

$$V_{k,s}(y) = \int_{\mathbf{R}} \frac{e^{-2\pi i y x}}{(x^2+1)^{s-k/2}(x+i)^k}\, dx.$$

(19) Cf. [45, §§ 3.5, 6.2].

Proof. — We use the Bruhat decomposition

$$\mathbf{GL}_2(F) = B(F) \coprod \coprod B(F) w N(F)$$

with $w = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$ and the unipotent subgroup $N(F) \cong F$ via $N(F) \ni \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \leftarrow\!\shortmid x \in F$, to get

$$\varepsilon \phi^{-1}(y) c_s^{\delta}(\alpha, y) = L(2s, \varphi) D_F^{-1/2} \int_{\mathbf{A}_F/F} H_s \left(\begin{pmatrix} y & x \\ & 1 \end{pmatrix} W_{M/\delta} \right) \mathbf{e}(-\alpha x)\, dx$$
$$+ L(2s, \varphi) D_F^{-1/2} \int_{\mathbf{A}_F} H_s \left(w \begin{pmatrix} y & x \\ & 1 \end{pmatrix} W_{M/\delta} \right) \mathbf{e}(-\alpha x)\, dx.$$

At any place $v | M/\delta$, we have the decomposition

$$\begin{pmatrix} y_v & x_v \\ & 1 \end{pmatrix} W_{M/\delta, v} = \begin{pmatrix} y_v & \pi_v x_v \\ & \pi_v \end{pmatrix} \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$$

so that the first summand is always zero.

For the second integral, we use the identity

$$w \begin{pmatrix} y & x \\ & 1 \end{pmatrix} = \begin{pmatrix} 1 & \\ & y \end{pmatrix} \begin{pmatrix} & -1 \\ 1 & xy^{-1} \end{pmatrix}$$

and the substitution $x \to xy$ to get

$$\int_{\mathbf{A}_F} H_s \left(w \begin{pmatrix} y & x \\ & 1 \end{pmatrix} W_{M/\delta} \right) \mathbf{e}(-\alpha x)\, dx = |y|^{1-s} \prod_v V_s^{M/\delta}(\alpha_v y_v)$$

where for $y \in F_v$

$$V_s^M(y) = \int_{F_v} H_s\left(\begin{pmatrix} & -1 \\ 1 & x \end{pmatrix} W_{M,v}\right) \mathbf{e}(-xy)\,dx. \tag{3.2.1}$$

Archimedean places. — As in [45, Proposition 3.5.2].

Nonarchimedean places $v \nmid M/\delta$. — If v is a finite place, we have $\begin{pmatrix} & -1 \\ 1 & x \end{pmatrix} \in \mathbf{GL}_2(\mathscr{O}_{F,v})$ if $x \in \mathscr{O}_{F,v}$, and otherwise we have the decomposition

$$\begin{pmatrix} & -1 \\ 1 & x \end{pmatrix} = \begin{pmatrix} x^{-1} & -1 \\ & x \end{pmatrix} \begin{pmatrix} 1 & \\ x^{-1} & 1 \end{pmatrix}.$$

Therefore

$$H_{s,v}\left(\begin{pmatrix} & -1 \\ 1 & x \end{pmatrix}\right) = \begin{cases} \overline{\varphi}_v(x)|x|^{-2s} & \text{if } v(x) \le -1; \\ 1 & \text{if } v \nmid M, v(x) \ge 0; \\ 0 & \text{if } v \mid \delta, v(x) \ge 0. \end{cases} \tag{3.2.2}$$

The case $v \nmid M$. — We deduce that

$$\begin{aligned} V_s^{M/\delta}(y) &= \int_{\mathscr{O}_{F,v}} \mathbf{e}(-xy)\,dx + \sum_{n\ge 1} \int_{\mathscr{O}_{F,v}^\times} \overline{\varphi}_v(x\pi_v^{-n})|x\pi_v^{-n}|^{-2s}\mathbf{e}(-xy\pi_v^{-n})\,d(\pi_v^{-n}x) \\ &= \mathbf{1}[y \in d_F^{-1}] + \sum_{n\ge 1} \varphi_v(\pi_v)^n |\pi_v|^{n(2s-1)} \int_{\mathscr{O}_{F,v}^\times} \mathbf{e}(-xy\pi_v^{-n})\,dx. \end{aligned}$$

The integral evaluates to $1-|\pi_v|$ if $v(yd_F) \ge n$, to $-|\pi_v|$ if $v(yd_F) = n-1$, and to zero otherwise. Therefore we have $V_s^M(y) = 0$ unless $v(yd_F) \ge 0$ in

which case

$$V_s^{M/\delta}(y) = 1 + (1-|\pi_v|)\sum_{n=1}^{v(yd_F)} \left(\varphi_v(\pi_v)|\pi_v|^{2s-1}\right)^n - |\pi_v|(\varphi_v(\pi_v)|\pi_v|^{2s-1})^{v(yd_f)+1}$$

$$= \left(1-\varphi_v(\pi_v)|\pi_v|^{2s}\right)\sum_{n=0}^{v(yd_F)} \varphi_v(\pi_v)^n|\pi_v|^{n(2s-1)}$$

$$= L_v(2s,\varphi)^{-1}\sum_{n=0}^{v(yd_F)} \varphi_v(\pi_v)^n|\pi_v|^{n(2s-1)}.$$

The case $v|\delta$. — Again by (3.2.2) we find

$$V_s^{M/\delta}(y) = \sum_{n\geq 1}\int_{\mathcal{O}_{F,v}^\times} \overline{\varphi}_v(x\pi_v^{-n})|x\pi_v^{-n}|^{-(2s-1)}\mathbf{e}(-xy\pi_v^{-n})\,dx.$$

All the integrals vanish except the one with $n = v(yd_F)+1$ which gives

$$\varepsilon_v(y\pi_v^n)\phi_v(y\pi_{d_F,v}\pi_v)|y\pi_{d_F,v}\pi_v|^{2s-1}|\pi_v|^{1/2}\kappa(v);$$

therefore we have[20]

$$V_s^{M/\delta}(y) = \varepsilon_v(y)\phi_v(y\pi_{d_F,v}\pi_v)|y\pi_{d_F,v}\pi_v|^{2s-1}|\pi_v|^{1/2}\kappa(v)$$

if $v(yd_F)\geq 0$ and $V_s(y)=0$ otherwise.

Places $v|M/\delta$. — For $w\begin{pmatrix}1 & x\\ & 1\end{pmatrix}\begin{pmatrix} & 1\\ -\pi_v^{v(M)} & \end{pmatrix} = \begin{pmatrix}\pi_v^{v(M)} & \\ -x\pi_v^{v(M)} & 1\end{pmatrix}$ we have the decompositions

$$\begin{pmatrix}\pi_v^{v(M)} & \\ -x\pi_v^{v(M)} & 1\end{pmatrix} = \begin{pmatrix}-\pi_v^{v(M)} & \\ & 1\end{pmatrix}\begin{pmatrix}-1 & \\ -x\pi_v^{v(M)} & 1\end{pmatrix}$$

[20]Recall that we always choose π_v so that $\varepsilon_v(\pi_v)=1$.

$$= \begin{pmatrix} x^{-1} & -\pi_v^{v(M)} \\ & x\pi_v^{v(M)} \end{pmatrix} \begin{pmatrix} & 1 \\ -1 & x^{-1}\pi_v^{-v(M)} \end{pmatrix};$$

for $v(x) \geq 0$ we use the first one to find

$$H_s\left(\begin{pmatrix} -\pi_v^{v(M)} & \\ x\pi_v^{v(M)} & -1 \end{pmatrix}\right) = |\pi_v^{v(M)}|^s;$$

for $v(x) < 0$ the second decomposition shows that the integrand vanishes. We conclude that

$$V_s^{M/\delta}(y) = \begin{cases} |\pi_v^{v(M)}|^s & \text{if } v(yd_F) \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The final formula follows from these computations. □

We specialize to the case $s = 1/2$, $k = 1$ and consider the rescaled holomorphic Eisenstein series:

$$\mathbf{E}_{\varepsilon\phi}^{\Delta P}(g) = \frac{D_F^{1/2}\mathbf{N}(\Delta P)^{1/2}}{(-2\pi i)^{[F:\mathbf{Q}]}} E_1^{\Delta P}(g, 1/2; \varepsilon\phi),$$

$$\widetilde{\mathbf{E}}_{\varepsilon\phi}^{\Delta P}(g) = \frac{D_F^{1/2}\mathbf{N}(\Delta P)^{1/2}}{(-2\pi i)^{[F:\mathbf{Q}]}} \widetilde{E}_1^{\Delta P}(g, 1/2; \varepsilon\phi).$$

Corollary 3.2.2. — *The Eisenstein series $\mathbf{E}_{\varepsilon\phi}^{\Delta P}$, $\widetilde{\mathbf{E}}_{\varepsilon\phi}^{\Delta P}$ belong to $M_1(K_1(\Delta P), \varepsilon\phi^{-1})$ and $M_1(K_1(\Delta P), \varepsilon\phi)$ respectively. The Whittaker-Fourier coefficients $c^\delta(y)$ of $\widetilde{\mathbf{E}}_{\varepsilon\phi}^{\delta}$ for $\delta|\Delta$ are zero if $y^\infty d_F$ is not integral and otherwise given by*

$$c^\delta(y) = \tilde{a}(\widetilde{\mathbf{E}}_{\varepsilon\phi}^{(\delta)}, y) = \varepsilon\phi^{-1}(y)|y|^{1/2}\kappa(\delta)\phi(\delta)\varepsilon_\delta\phi_\delta(y^\infty d_F)\sigma_{\varepsilon\phi}(y^\infty d_F),$$

where for any integral ideal m of $\mathcal{O}_F[\Delta^{-1}P^{-1}]$,

$$\sigma_{\varepsilon\phi}(m)=\sum_{d|m}\varepsilon\phi(d),$$

the sum likewise running over integral ideals of $\mathcal{O}_F[\Delta^{-1}P^{-1}]$.

(If m is an integral ideal of $\mathcal{O}_F$ prime to P, then $\sigma_{\varepsilon 1}(m)=r(m)$.)

Proof. — This follows from Proposition 3.2.1 once observed that [13, Propostion IV.3.3 (d)]

$$V_{1,1/2}(t)=\begin{cases}0 & \text{if } t<0\\ -2\pi i e^{-2\pi t} & \text{if } t>0.\end{cases}$$

□

Definition 3.2.3. — Let F'_∞ be the maximal abelian extension of F unramified outside p, and let $\mathcal{G}'_F=\mathrm{Gal}(F'_\infty/F)$. We define[21] the **Eisenstein distribution**[22] $\widetilde{\mathbf{E}}_\varepsilon$ on $\mathcal{G}'_F$ by

$$\widetilde{\mathbf{E}}_\varepsilon(\phi)=\frac{D_F^{1/2}\mathbf{N}(\Delta P)^{1/2}}{(-2\pi i)^g}\widetilde{E}^{\Delta P}_{\varepsilon\phi}$$

for any character ϕ of $\mathcal{G}'_F$ of conductor dividing P (it does not depend on the choice of P once we require P to satisfy $v|P\leftrightarrow v|p$). We denote with the same name the distribution induced on the group $\mathcal{G}'$ of §2.3 by

$$\widetilde{\mathbf{E}}_\varepsilon(\mathcal{W})=\widetilde{\mathbf{E}}_\varepsilon(\mathcal{W}|_{F_{\mathbf{A}}^\times}).$$

[21] We *do not* assume that Δ is squarefree when making the definition.
[22] A *distribution* on $\mathcal{G}_F$ is a linear functional on locally constant functions on $\mathcal{G}_F$ – see below for why $\widetilde{\mathbf{E}}_\varepsilon$ is not a measure.

It has values in $\mathbf{M}_1(K_1(N\Delta),\varepsilon)$ and is defined over $\mathbf{Q}_p$.

Remark 3.2.4. — In the case that $\phi = \mathbf{1}$, the constant term of $\widetilde{\mathbf{E}}_\varepsilon(\phi)$ is no longer zero (it is in fact a multiple of $L(1,\varepsilon)$, cf. [45, Lemma 6.2.2]) – for this reason, the constant term of the distribution is unbounded, so that the distribution is not a measure. This difficulty is easy to circumvent as done in [30, §2] following [21]: for a suitable choice of a nonzero ideal C, there is a *measure* $\widetilde{\mathbf{E}}^C_\varepsilon$ whose value on a character ϕ is a nonzero multiple of $\widetilde{\mathbf{E}}_\varepsilon(\phi)$; then one can use $\widetilde{\mathbf{E}}^C_\varepsilon$ rather than $\widetilde{\mathbf{E}}_\varepsilon$, and remove the factor afterwards from the formulas. Since this method is by now standard, we will be content with the present caveat and treat the $\widetilde{\mathbf{E}}_\varepsilon$ as if they were measures in what follows.

4. The p-adic L-function

4.1. Rankin–Selberg convolution. — Let f, g be modular forms of common level M, weights k_f, k_g, and characters ψ_f, ψ_g respectively. We define a normalised Dirichlet series

$$D^M(f,g,s) = L^M(2s-1,\psi_f\psi_g)\sum_m a(f,m)a(g,m)\mathbf{N}m^{-s},$$

where the imprimitve L-function $L^M(s,\varphi)$ of a Hecke character φ of conductor dividing M is as in (3.1.1).

When f and g are primitive forms of level N_f, N_g (that is, normalized new eigenforms at those levels), for a prime $\wp \nmid N_f$ denote by $\alpha^{(1)}_\wp$, $\alpha^{(2)}_\wp$ the two roots of the $\wp^{\text{th}}$ Hecke polynomial of f

$$P_{\wp,f}(X) = X^2 - a(f,\wp)X + \psi_f(\wp)\mathbf{N}\wp^{k_f-1},$$

and by $\beta_{\wp}^{(1)}$, $\beta_{\wp}^{(2)}$ the analogous quantities for g. Then the degree four Rankin–Selberg L-function $L(f\times g,s)$ with unramified Euler factors at $\wp$ given by

$$\prod_{i,j=1}^{2}\left(1-\alpha_{\wp}^{(i)}\beta_{\wp}^{(j)}\mathbf{N}\wp^{-s}\right)^{-1}$$

equals the above Dirichlet series

$$L(f\times g,s)=D^{N_fN_g}(f,g,s)$$

if N_f and N_g are coprime.

Suppose now for simplicity that $k_f=2$, $k_g=1$, and f is a cusp form (not necessarily primitive). The Rankin–Selberg convolution method[23] gives

(4.1.1)

$$\langle f^{\rho},gE_1^M(s;\psi_f\psi_g)\rangle_M=D_F^{s+1}\left[\frac{\Gamma(s+1/2)}{(4\pi)^{s+1/2}}\right]^{[F:\mathbf{Q}]}D^M(f\times g,s+1/2),$$

where $\langle\cdot,\cdot\rangle_M$ is the Petersson inner product (1.1.2).

4.2. Convoluted measure and the p-adic L-function. — Consider the convolution 'measure'[24] $\Theta*\widetilde{\mathbf{E}}_{\varepsilon,N}$ on $\mathscr{G}'$ defined by $\Theta*\widetilde{\mathbf{E}}_{\varepsilon,N}(\mathscr{W})=\Theta(\mathscr{W})\widetilde{\mathbf{E}}_{\varepsilon,N}(\overline{\mathscr{W}})$ for any character $\mathscr{W}:\mathscr{G}\to\mathbf{Z}_p^{\times}$, where $\widetilde{\mathbf{E}}_{\varepsilon,N}=[N]\widetilde{\mathbf{E}}_{\varepsilon}$. We deduce from it the 'measure'

$$\Phi(\mathscr{W})=\mathrm{Tr}_{\Delta}[\Theta*\widetilde{\mathbf{E}}_{\varepsilon,N}(\mathscr{W})]=\mathrm{Tr}_{\Delta}[\Theta(\mathscr{W})[N]\widetilde{\mathbf{E}}_{\varepsilon}(\overline{\mathscr{W}})] \tag{4.2.1}$$

[23] See or [19, Ch. V], or [45, Lemma 6.1.3] for a setting close to ours

[24] The same caveat of Remark 3.2.4 applies here. Namely, the convolution $\Theta*\widetilde{\mathbf{E}}_{\varepsilon,N}$ is only a distribution, but its value on any character $\mathscr{W}$ is a simple nonzero multiple of the value of a *measure* $\Theta*\widetilde{\mathbf{E}}^{C}_{\varepsilon,N}$ on $\mathscr{W}$, so that we can "pretend" it to be a measure. Similarly for Φ.

on $\mathscr{G}'$, which is a kind of p-adic kernel of the Rankin–Selberg L-function as will be made precise below. It is valued in $\mathbf{S}_2(K_0(N), \mathbf{C}_p)$.

Definition 4.2.1. — The p-adic Rankin–Selberg L-function is the bounded element of $\mathbf{Q}_p[\![\mathscr{G}]\!]$ defined by

$$L_p(f_E, \mathscr{W}) = D_F^{-2} H_p(f) L_{f_0}(\Phi(\mathscr{W}))$$

for any character $\mathscr{W} : \mathscr{G} \to \mathbf{Z}_p^\times$, where

$$H_p(f) = \prod_{\wp | p} \left(1 - \frac{1}{\alpha_\wp(f)^2}\right)\left(1 - \frac{\mathbf{N}\wp}{\alpha_\wp(f)^2}\right). \tag{4.2.2}$$

Functional equation. — Consider the restriction of $L_p(f_E)$ to $\mathbf{Z}_p[\![\mathscr{G}]\!]$. Let

$$\Lambda_p(f_E)(\mathscr{W}) = \mathscr{W}(N)^{1/2} \mathscr{W}(\mathfrak{D}) L_p(f_E, \mathscr{W}),$$

where the square root is taken in $1 + p\mathbf{Z}_p$. Let ι denote the involution of $\mathbf{Z}_p[\![\mathscr{G}]\!]$ defined by $\lambda^\iota(\mathscr{W}) = \lambda(\overline{\mathscr{W}}^c)$, where $\mathscr{W}^c(\sigma) = \mathscr{W}(c\sigma c)$.

Then we have a functional equation

$$\Lambda_p(f_E)^\iota = (-1)^g \varepsilon(N) \Lambda_p(f_E), \tag{4.2.3}$$

which can be proven as in [29, §5.2] using the interpolation property of §4.4 and the functional equation for the complex Rankin–Selberg L-fucntion.

In particular, if $\varepsilon(N) = (-1)^{g-1}$ and $\mathscr{W}$ is an anticyclotomic character (i.e., $\mathscr{W}\mathscr{W}^c = 1$), we have

$$L_p(f_E)(\mathscr{W}) = 0.$$

4.3. Toolbox. — Here is a collection of technical lemmas that we need to prove the soundness of our construction of the p-adic L-function. The

reader is invited to skip to §4.4 where the results are put to use, and return here as the need arises.

Lemma 4.3.1. — *Let P be an ideal of $\mathcal{O}_F$ such that $v|P$ if and only if $v|p$. We have*

$$\langle W_{NP} f_0^{\rho}, f_0\rangle_{NP} = \alpha_P(f)(-1)^g \tau(f) H_p(f)\langle f, f\rangle_N$$

with $H_p(f)$ as in (4.2.2) *and*

$$\alpha_P(f) = \prod_{\wp|p} \alpha_\wp(f)^{v_\wp(P)}$$

The proof is similar to that of [29, Lemme 27].

Lemma 4.3.2. — *For a character φ of conductor dividing M and an ideal N prime to M, let $E_\varphi^M = E_1^M(\mathrm{g}, 1/2; \varphi)$, $\widetilde{E}_\varphi^M = W_M E_\varphi^M$. We have*

$$W_M[N]\widetilde{E}_\varphi^M = E_\varphi^{MN} + E^{\mathrm{old}}$$

where the form E^{old} is old at N (in particular, it is orthogonal to newforms of level N).

Proof. — It is easy to see that $W_M[N]\widetilde{E}_\varphi^M = \varphi(N)[N]E_\varphi^M$. Then we are reduced to showing that

$$[N]E_\varphi^M = \varphi^{-1}(N)E_\varphi^{MN} + E^{\mathrm{old}}$$

which can be done by the calculation appearing in the proof of [45, Lemma 6.1.4] □

Lemma 4.3.3. — *Let f and f_0 be as usual, and let g be a primitive form of weight* 1, *level prime to p. If $G_\wp(\mathbf{N}\wp^{-s})$ denotes the $\wp^{\mathrm{th}}$ of Euler factor of*

$L(g,s)$ at the prime $\wp|p$, then

$$D(f_0,g,1)=G_p(\alpha_\wp(f)^{-1})D(f,g,1).$$

The proof is identical to that of [29, Lemme 23]. The result will be applied to the case $g=\Theta(\mathscr{W})$, in which case the Euler factor at $\wp$ is $\prod_{\mathfrak{p}|\wp}(1-\overline{\mathscr{W}}(\mathfrak{p})\mathbf{N}\wp^{-s})$ and we use the notation

$$V_p(f,\overline{\mathscr{W}})=\prod_{\wp|p}\prod_{\mathfrak{p}|\wp}\left(1-\frac{\overline{\mathscr{W}}(\mathfrak{p})}{\alpha_\wp(f)}\right) \tag{4.3.1}$$

for $G_p(\alpha_\wp(f)^{-1})$.

Lemma 4.3.4. — *With notation as in §4.1, we have*

$$D([\Delta]f,\Theta(\mathscr{W}),1)=\mathscr{W}(\mathfrak{D})D(f,\Theta(\mathscr{W}),1).$$

The proof is an easy calculation (cf. [27, § I.5.9]).

4.4. Interpolation property. — We manipulate the definition to show that the p-adic L-function $L_p(f_E)(\mathscr{W})$ of Definition 4.2.1 interpolates the special values of the complex Rankin–Selberg L-function $L(f_E,\mathscr{W},s)$ defined in the Introduction. Notice that if $\mathscr{W}$ is ramified at all places dividing p, then

$$L(f_E,\mathscr{W},s)=L(f\times\Theta(\mathscr{W}),s)$$

where $L(f\times\Theta(\mathscr{W}),s)$ is as in §4.1. (Otherwise, the above equality remains true for the Dirichlet series associated to f and $\Theta(\mathscr{W})$ after removing from $L(f_E,\mathscr{W},s)$ the Euler factors at p, and likewise for the interpolation result just below.)

Theorem 4.4.1. — *Let $\mathscr{W}: \mathscr{G}' \to \overline{\mathbf{Q}}^\times$ be a finite order character of conductor $\mathfrak{f}$. Assume that $v|\mathfrak{f}$ if and only if $v|p$. Then we have*

$$L_p(f_E)(\mathscr{W}) = \frac{\tau(\mathscr{W})\mathbf{N}(\Delta(\mathscr{W}))^{1/2} V_p(f,\mathscr{W})\overline{\mathscr{W}}(\mathfrak{D})}{\alpha_{\mathfrak{N}(\mathfrak{f}(\mathscr{W}))}(f)\Omega_f} L(f_E, \overline{\mathscr{W}}, 1),$$

where $\Omega_f = (8\pi^2)^g \langle f, f\rangle_N$, $\tau(\mathscr{W})$ is as in Remark 2.4.2 and the other factors are defined in §4.3.

Proof. — Denote $P = \mathfrak{N}(\mathfrak{f}(\mathscr{W}))$, $\Delta(\mathscr{W}) = \Delta P$, $\phi = \mathscr{W}|_{F_{\mathbf{A}}^\times}$. The result follows from the definition and the following calculation.

$$\begin{aligned}
L_{f_0}(\Phi(\mathscr{W})) &= \frac{\langle W_{NP} f_0^\rho, \mathrm{Tr}_\Delta[\Theta(\mathscr{W})\widetilde{\mathbf{E}}_{\varepsilon,N}(\overline{\mathscr{W}})]\rangle_{NP}}{\langle W_{NP} f_0^\rho, f_0\rangle_{NP}} \\
\text{(L. 4.3.1)} &= \frac{\langle W_{N\Delta} f_0^\rho, W_{\Delta(\mathscr{W})}\Theta(\mathscr{W}) W_{\Delta(\mathscr{W})}\widetilde{\mathbf{E}}^{\Delta(\mathscr{W})}_{\varepsilon\phi^{-1},N}\rangle_{N\Delta(\mathscr{W})}}{\alpha_P(f)(-1)^g\tau(f)H_p(f)\Omega_f} \\
\text{(L. 4.3.2)} &= \frac{(-i)^g\tau(\mathscr{W})D_E}{\alpha_P(f)(-1)^g\tau(f)H_p(f)\Omega_f}\langle W_N[\Delta]f_0^\rho, \Theta(\overline{\mathscr{W}})\mathbf{E}^{N\Delta(\mathscr{W})}_{\varepsilon\phi^{-1}}\rangle_{N\Delta(\mathscr{W})} \\
&= \frac{(-i)^g\tau(\mathscr{W})}{\alpha_P(f)H_p(f)\Omega_f}\langle [\Delta]f_0^\rho, \Theta(\overline{\mathscr{W}})\mathbf{E}^{N\Delta(\mathscr{W})}_{\varepsilon\phi^-}\rangle_{N\Delta(\mathscr{W})} \\
\text{(eq. (4.1.1))} &= \frac{\tau(\mathscr{W})D_F^2\mathbf{N}(\Delta(\mathscr{W}))^{1/2}}{\alpha_P(f)H_p(f)\Omega_f} D^{N\Delta(\mathscr{W})}([\Delta]f_0, \Theta(\overline{\mathscr{W}}), 1) \\
\text{(L. 4.3.3)} &= \frac{\tau(\mathscr{W})D_F^2\mathbf{N}(\Delta(\mathscr{W}))^{1/2}V_p(f,\mathscr{W})}{\alpha_P(f)H_p(f)\Omega_f} D^{N\Delta(\mathscr{W})}([\Delta]f, \Theta(\overline{\mathscr{W}}), 1) \\
\text{(L. 4.3.4)} &= \frac{\tau(\mathscr{W})D_F^2\mathbf{N}(\Delta(\mathscr{W}))^{1/2}V_p(f,\mathscr{W})\overline{\mathscr{W}}(\mathfrak{D})}{\alpha_P(f)H_p(f)\Omega_f} L(f_E, \overline{\mathscr{W}}, 1)
\end{aligned}$$

where we have used various results from §1.3, and the fact that in our case $f^\rho = f$ as f has trivial character. □

4.5. Fourier expansion of the measure. — Consider the restriction of Φ to $\mathscr{G}$, the Galois group of the maximal $\mathbf{Z}_p$-extension of E unramified outside . Any character $\mathscr{W}$ of $\mathscr{G}$ decomposes uniquely as $\mathscr{W} = \mathscr{W}^+\mathscr{W}^-$ with $(\mathscr{W}^+)^c = \mathscr{W}$, $(\mathscr{W}^-)^c = \overline{\mathscr{W}}$ (we say that $\mathscr{W}^+$ is cyclotomic and $\mathscr{W}^-$ is anticyclotomic or dihedral). Since we are interested in the case where $\varepsilon(N) = (-1)^{g-1}$ in which the Φ is zero on the anticyclotomic characters, we study the restriction of Φ to the cyclotomic characters. We can write $\mathscr{W}^+ = \chi \circ \mathfrak{N}$ for a Hecke character $\chi : F^\times \backslash F_{\mathbf{A}}^\times \to 1 + p\mathbf{Z}_p$, and we denote

$$\Theta_\chi = \Theta(\chi \circ \mathfrak{N}), \qquad \Phi_\chi = \Phi(\chi \circ \mathfrak{N}).$$

From now on we assume that $(\Delta, 2) = 1$ and all primes $\wp | p$ are split in E.

Proposition 4.5.1. — *The Fourier coefficients $b(m) = a_p(\Phi_\chi, m)$ of the p-adic modular form Φ_χ are given by*

$$b(m) = \sum_{\substack{n \in F \\ 0 < n < 1 \\ n \in Nm^{-1}\Delta^{-1}}} \chi((1-n)m) \prod_{v|\Delta} \left[\mathbf{1}[v(nm) = 0] + \varepsilon_v((n-1)n)\chi_v^{-2}(nm\wp_v/N) \right] \cdot r((1-n)m\Delta)\sigma_{\varepsilon\chi^{-2}}(nm).$$

Proof. — By (1.3.6), the Fourier coefficients $b(m)$ of $\Phi_\chi = \Theta_\chi \widetilde{\mathbf{E}}_{\varepsilon\chi^2, N}$ is given by

$$b(m) = \sum_{\delta | \Delta} b^\delta(m\delta)$$

with

$$b^\delta(m) = a(\Phi_\chi^{(\delta)}, m) = |y|^{-1} \tilde{a}(\Phi_\chi^{(\delta)}, y) = |y|^{-1} \sum_{n \in F} \tilde{a}(\Theta_\chi^{(\delta)}, (1-n)y) \tilde{a}(\widetilde{\mathbf{E}}_{\varepsilon\chi^2, N}^{(\delta)}, ny)$$
$$= |y|^{-1} \sum_{n \in F} \tilde{a}(\Theta_\chi^{\delta}, (1-n)y) \tilde{a}(\widetilde{\mathbf{E}}_{\varepsilon\chi^2}^{(\delta)}, ny/\pi_N)$$

if $y \in F_{\mathbf{A}}^{\times}$ satisfies $y_\infty > 0$ and $y^\infty d_F = m$.

Then thanks to Proposition 2.5.1 and Corollary 3.2.2, we have[25]:

$$b(m) = \sum_{\delta|\Delta} \sum_{\substack{n \in F \\ 0<n<1}} \varepsilon_\delta((n-1)n)\chi((1-n)m)\chi_\delta^{-2}(nm\delta/N) \cdot r((1-n)m\delta)\sigma_{\varepsilon\chi^{-2}}(nm/N).$$

We interchange the two sums and notice that the term corresponding to δ and n is nonzero only if $n \in Nm^{-1}\Delta^{-1}$ and $\delta_0|\delta$, where

$$\delta_0 = \delta_0(n) = \prod_{\substack{v|\Delta \\ v(nm)=-1}} \wp_v$$

($\wp_v$ being the prime corresponding to v). Now for each n we can rewrite the sum over δ (omitting the factor $\chi((1-n)m)$ which does not depend on δ) as

$$\varepsilon_{\delta_0}((n-1)n)\chi_{\delta_0}^{-2}(nm\delta_0/N) \sum_{\delta'|\Delta/\delta_0} \varepsilon_{\delta'}((n-1)n)\chi_{\delta'}^{-2}(nm\delta'/N)$$
$$= \prod_{v|\delta_0} \varepsilon_{\delta'}((n-1)n)\ \chi_v^{-2}(nm\wp_v) \prod_{v|\Delta/\delta_0} [1 + \varepsilon_v((n-1)n)\chi_v^{-2}(nm\wp_v)].$$

The asserted formula follows. □

Remark 4.5.2. — If $v(nm) = -1$ then $(n-1)\pi_m\pi_v \equiv n\pi_m\pi_v$ in $(\mathcal{O}_{F,v}/\pi_v\mathcal{O}_{F,v})^\times$, so that we actually have

$$\varepsilon_v((n-1)n) = \varepsilon_v((n-1)\pi_m\pi_v)\varepsilon_v(n\pi_m\pi_v) = 1,$$

where $\wp_v$ is the ideal corresponding to v.

[25] Recall that $\kappa(v)^2 = \varepsilon_v(-1)$.

We can now compute the Fourier coefficients of the measure giving the central derivative of the p-adic L-function in the cyclotomic direction. To this end, let

$$\nu\colon \mathrm{Gal}(\overline{\mathbf{Q}}/F) \to 1 + p\mathbf{Z}_p \subset \mathbf{Q}_p^\times$$

be a character of F which is ramified exactly at the primes dividing p, and for $s \in \mathbf{Z}_p$ denote[26] $\Phi(s) = \Phi_{\nu^s}$. Let $\ell_F = \frac{d}{ds}\nu^s|_{s=0}\colon F^\times \backslash F_{\mathbf{A}_\infty^\times}^\times \to \mathbf{Q}_p$ be the associated p-adic logarithm.

Proposition 4.5.3. — *Assume that $\varepsilon(N) = (-1)^{g-1}$. Then $\Phi(0) = 0$ and the Fourier coefficients $b'(m)$ of*

$$\Phi'_\nu = \Phi'(0) = \frac{\partial}{\partial s}\Phi_{\nu^s}|_{s=0}$$

are nonzero only for m integral and nonzero, in which case

$$b'(m) = \sum_v b'_v(m)$$

with the sum running over all finite places v of F and $b'_v(m)$ given for $p|m$ by

1. *If $v = \wp$ is inert in E, then*

$$b'_v(m) = \sum_{\substack{n \in Nm^{-1}\Delta^{-1} \\ (p,nm)=1 \\ \varepsilon_v((n-1)n)=1 \forall v|\Delta \\ 0<n<1}} 2^{\omega_\Delta(n)} r((1-n)m\Delta) r(nm\Delta/N\wp)(v(nm/N)+1)\ell_{F,v}(\pi_v),$$

where

$$\omega_\Delta(n) = \#\{v | (\Delta, nm\Delta)\}.$$

[26] No confusion should arise from our recycling the letter s for this p-adic variable, the complex variable having now exited the scene.

2. If $v = \wp|\Delta$ is ramified in E, then

$$b'_v(m) = \sum_{\substack{n\in Nm^{-1}\Delta^{-1}\\(p,nm)=1\\ \varepsilon_v((n-1)n)=-1\\ \varepsilon_w((n-1)n)=1\forall v\neq w|\Delta\\ 0<n<1}} 2^{\omega_\Delta(n)} r((1-n)m\Delta)r(nm\Delta/N)(v(nm)+1)\ell_{F,v}(\pi_v).$$

3. If v is split in E, then

$$b'_v(m) = 0.$$

Proof. — The vanishing of $\Phi(0) = \Phi_1$ follows from the functional equation (4.2.3) and the sign assumption.

By Proposition 4.5.1, the Fourier coefficient $b_s(m)$ of $\Phi(s) = \Phi_{\nu^s}$ can be expressed as $b_s(m) = \sum_{n\in F} b_{n,s}(m)$ with

$$b_{n,s}(m) = \nu^s((1-n)m)r((1-n)m\Delta)\prod_{v\nmid p\infty}\sigma^n_{s,v}(m/N)$$

where, using Remark 4.5.2:

$$\sigma^n_{s,v}(m) = \begin{cases} \dfrac{1-\varepsilon(nm\wp)\nu(nm\wp)^{-2s}}{1-\varepsilon(\wp)\nu(\wp)^{-2s}} & \text{if } v=\wp\nmid\Delta;\\ 1+\varepsilon_v(n(n-1))\nu(nm\wp)^{-2s} & \text{if } v=\wp|\Delta \text{ and } v(nm)=0;\\ \nu(nm\wp)^{-2s} & \text{if } v=\wp|\Delta \text{ and } v(nm)=-1. \end{cases}$$

Then $b'(m) = \sum_n b'_n(m) = \sum_n\sum_v b'_{n,v}(m)$ with $\sum_n b'_{n,v}(m) = b_v(m)$, and $b'_n(m)$ can be nonzero only if exactly one of the factors $\sigma^n_{s,v}$ vanishes at $s = 0$. If this happens for the place v_0, then the set over which n ranges accounts for the positivity and integrality conditions and the nonvanishing conditions at other places, whereas the condition $(p, nm) = 1$ results from observing that $\lim_{s\to 0}\nu^s(a) = \mathbf{1}[(p,a) = 1]$.

The values of $b'_{n,v}$ can then be determined in each case from the above expressions: for v ramified this is straightforward. For $v = \wp$ inert, notice that if $v(nm/N)$ is odd then $r(nm\Delta/N\wp) = r((nm\Delta/N)^{(\wp)})$, where the superscript denotes prime-to-$\wp$ part; whereas if $v(nm/N)$ is even then $\sigma^n_{0,v}(m/N)$ does not vanish so (n,v) does not contribute to $b'(m)$ and indeed $r(nm/N\wp) = 0$. □

PART II

HEIGHTS

5. Generalities on p-adic heights and Arakelov theory

5.1. The p-adic height pairing. — Let X be a (smooth, projective) curve of genus $g \geq 1$ over a number field E with good reduction at all places above p; or let A be an abelian variety of dimension g over E with good reduction at all places above p. By the work of many authors (Schneider, Perrin-Riou, Mazur–Tate, Coleman–Gross, Zarhin, Nekovář,...) there are p-adic height pairings on the group of degree zero divisors on X and on the Mordell-Weil group of A; we recall their main properties following the construction of Zarhin [44] and Nekovář [26].

Let $\ell : E_{\mathbf{A}}^\times / E^\times \to \mathbf{Q}_p$ be a homomorphism; we call ℓ a *p-adic logarithm* and assume that it is ramified at all $v|p$, that is, $\ell_v : E_v^\times \to \mathbf{Q}_p$ does not vanish identically on $\mathcal{O}_{E,v}^\times$. Let Y denote either $\mathrm{Alb}\, X$ or A, $Y^\vee$ its dual abelian variety, and let $V = V_p Y = T_p Y \otimes_{\mathbf{Z}_p} \mathbf{Q}_p$ be the rational Tate module

of Y, a continuous $\mathrm{Gal}(\overline{E}/E)$-representation.[27] Let $D_{\mathrm{dR}}(V) = H^1_{\mathrm{dR}}(Y^\vee \otimes E_v/E_v)$, equipped with the Hodge filtration. For each $v|p$, let $W_v \subset D_{\mathrm{dR}}(V)$ be a splitting of the Hodge filtration, that is, a complementary subspace to $\Omega^1(Y^\vee \otimes E_v/E_v) \subset D_{\mathrm{dR}}(V)$, which is isotropic[28] for the cup product. When V is ordinary as a $\mathrm{Gal}(\overline{E}_v/E_v)$-representation, there is a canonical choice for W_v, the "unit root" subspace – see [18] for a nice discussion. For the case of interest to us, recall that we have fixed in the Introduction an ordinary form f and a Shimura curve X/F: then we will study the situation where $Y = \mathrm{Alb}\, X$ or $Y = A_f$, the ordinary abelian variety associated to f (see below for the construction of A_f), or their base change to E, and $V_f = V_p A_f) \subset V_p \mathrm{Alb}\, X$. The subspaces W_v will then be assumed chosen compatibly with the canonical choices on V_f.

Theorem 5.1.1. — *Let V be a representation as above, and let $V^*(1)$ denote the twisted dual of V (it is in fact isomorphic to V).*

1. *There is a bilinear symmetric pairing on the Bloch-Kato Selmer group of V*

$$\langle \cdot, \cdot \rangle : H^1_f(E, V) \times H^1_f(E, V^*(1)) \to \mathbf{Q}_p,$$

depending on the auxiliary choices of ℓ and $(W_v)_{v|p}$ (which we usually omit from the notation).

[27] Nekovář [26] defines height pairings for Galois representations in much greater generality than described here.
[28] The isotropy condition ensures that the resulting height pairing is symmetric [26, Theorem 4.1.1 (4)]

2. *The above pairing decomposes as the sum*

$$\langle x, y\rangle = \sum_v \langle x, y\rangle_v$$

of local pairings $\langle\cdot,\cdot\rangle_v$ for v running over the finite primes of E.

3. *(Compatibility.) Let E'_w/E_v be a finite extension, and let $\mathrm{Tr}_{E'_w/E_v} : H^1_f(E'_w, T) \to H^1_f(E_v, T)$ denote corestriction. If $x \in H^1_f(E'_w, V)$, $y' \in H^1_f(E'_w, V^*(1))$, we have*

$$\langle x, \mathrm{Tr}_{E'_w/E_v}(y')\rangle_v = \langle x, y'\rangle_v$$

where $\langle\cdot,\cdot\rangle_w$ is the local pairing associated to $\ell_w = \ell_v \circ N_{E_w/E_v}$ and (if $v|p$) the splitting W_w induced from W_v.

4. *Let $E^\ell_{v,\infty} = \cup_n E^\ell_{v,n}$ be the ramified*[29] *$\mathbf{Z}_p$-extension of F determined by the isomorphism*

$$E_v^\times \supset \mathrm{Ker}(\ell_v) \cong \mathrm{Gal}(E^\ell_{v,\infty}/E) \subset \mathrm{Gal}(E_v^{\mathrm{ab}}/E).$$

induced from class field theory. Assume that V is ordinary as a Galois representation, and let T be a stable lattice in V. Then the module of universal norms

$$N^\ell_\infty(H^1_f(E_v, T)) = \bigcap_n \mathrm{Im}\left[\mathrm{Tr}_{E^\ell_{v,n}/E_v} : H^1_f(E^\ell_{v,n}, T) \to H^1_f(E_v, T)\right]$$

has finite index in $H^1_f(E_v, T)$.

5. *(Boundedness.) In the above ordinary situation, there is a nonzero constant $c \in \mathbf{Z}_p$ such that*

$$\langle x, y\rangle_{v,n} \in p^{-c}\ell_w(E^{\ell,\times}_{v,n})$$

[29] Recall that we choose ℓ_w to be ramified.

if $x \in H^1_f(E_v, T)$, $y \in H^1_f(E_v, T^*(1))$ *and* $\langle\cdot,\cdot\rangle_{v,n}$ *is the local pairing associated to the extension* $E^{\ell}_{v,n}/E_v$ *as in 3.*

We refer to [27, II.1] and references therein for the proof and more details on the construction. A more explicit description of the pairing and its local components will follow from Theorem 5.3.1 below.

Let $\mathrm{Div}(X)$ be the group of divisors on the curve X, $\mathrm{Div}^0(X)$ the subgroup of degree zero divisors, and similarly $\mathrm{CH}(X) = \mathrm{Div}^0(X)/\sim$, $\mathrm{CH}(X)_0 = \mathrm{Div}^0(X)/\sim$ the Chow group of zero-cycles modulo rational equivalence and its subgroup of degree zero elements. Let $V(X) = V_p \mathrm{Alb}\, X$.

Via the Abel-Jacobi map

$$\mathrm{CH}(X)_0 \to H^1_f(E, V(X)),$$

the above pairing induces a ***p*-adic height pairing**

$$\langle\cdot,\cdot\rangle : \mathrm{CH}(X)_0 \times \mathrm{CH}(X)_0 \to \mathbf{Q}_p.$$

In the case of an abelian variety A, via the Abel-Jacobi map $A^\vee(E) \cong \mathrm{CH}^1(A)_0 \to H^1_f(E, V_p A)$ and the similar one for $A(E)$, together with the canonical isomorphism $V_p A^*(1) \cong V_p A^\vee$, we similarly have an induced height pairing

$$\langle\cdot,\cdot\rangle : A^\vee(E) \times A(E) \to \mathbf{Q}_p.$$

By construction when $A = \mathrm{Alb}\, X$ this pairing coincides with the one induced from the one on X via the identification $\mathrm{Alb}\, X(E) \cong \mathrm{CH}(X)_0$ and the canonical autoduality of A.

5.2. p-adic Arakelov theory – local aspects. — Here and in §5.3 we summarize the main results of Besser [3], who develops the p-adic analogue of classical Arakelov theory.

Metrized line bundles. — Let X_v be a proper smooth variety over the finite extension E_v of $\mathbf{Q}_p$, and fix a ramified local p-adic logarithm $\ell_v : E_v^\times \to \mathbf{Q}_p$ which we extend to $\overline{\mathbf{Q}}_p^\times$ by $\ell_v|_{E_v'^\times} = \ell_v \circ N_{E_v'/E_v}$ for any finite extension E_v'/E_v.

A *metrized line bundle* $\widehat{\mathscr{L}} = (\mathscr{L}, \log_{\mathscr{L}})$ on X_v is a line bundle on X_v together with a choice of a *log function* $\log_{\mathscr{L},v}$ on the total space of $\mathscr{L}_v = \mathscr{L}|_{X\otimes E_v}$ minus the zero section (which will also be viewed as a function on the nonzero sections of $\mathscr{L}_v$). A log function is the analogue in the p-adic theory of the logarithm of a metric on the sections of $\mathscr{L}$ on an infinite fibre of $\mathscr{X}$. It is a Coleman function having a certain analytic property[30] and the following algebraic property. If the p-adic logarithm ℓ_v factors as

$$\ell_v = t_v \circ \log_v \tag{5.2.1}$$

for some $\log_v : E_v^\times \to E_v$ and some $\mathbf{Q}_p$-linear $t_v : E_v \to \mathbf{Q}_p$, then for any nonzero section s of $\mathscr{L}_v$ and rational function $f \in E(X_v)$ we have

$$\log_{\mathscr{L},v}(fs) = \log_v(f) + \log_{\mathscr{L},v}(s). \tag{5.2.2}$$

Adding a constant to a log function produces a new log function; this operation is called *scaling*.

[30] For which we refer to [3, Definition 4.1].

One can define a notion of $\overline{\partial}\partial$-operator on Coleman functions, and attach to any log function $\log_{\mathscr{L}}$ on $\mathscr{L}$ its *curvature* $\overline{\partial}\partial \log_{\mathscr{L}} \in H^1_{\mathrm{dR}}(X_v) \otimes \Omega^1(X_v)$; its cup product is the first Chern class of $\mathscr{L}$.

Log functions on a pair of line bundles induce in the obvious way a log function on their tensor product, and similarly for the dual of a line bundle. If $\pi : X_v \to Y_v$ is a morphism, then a log function on a line bundle on Y_v induces in the obvious way a log function on the pullback line bundle on X. If moreover π is a finite Galois cover with Galois group G, and $\mathscr{L}$ is a line bundle on X_v with log function $\log_{\mathscr{L}}$ and associated curvature β, then the norm line bundle $N_\pi \mathscr{L}$ on Y_v with stalks

$$(N_\pi \mathscr{L})_y = \bigotimes_{x \mapsto y} \mathscr{L}_x^{\otimes e(x|y)}$$

has an obvious candidate log function $N_\pi \log_{\mathscr{L}}$ obtained by tensor product. The latter is a genuine log function (i.e. it satisfies the analytic property alluded to above) when there is a form $\alpha \in H^1_{\mathrm{dR}}(Y_v) \otimes \Omega^1(Y_v)$ such that

$$\sum_{\sigma \in G} \sigma^* \beta = \pi^* \alpha,$$

in which case the curvature of $N_\pi \log_{\mathscr{L}}$ is $\alpha / \deg \pi$ ([3, Proposition 4.8]).

The canonical Green function. — Now let X_v / E_v be a curve of genus $g \geq 1$ with good reduction above p. Choose a splitting $W_v \subset H^1_{\mathrm{dR}}(X_v)$ of the Hodge filtration as in §5.1, which we use to identify $W_v \cong \Omega^1(X_v)^\vee$; we then define a canonical element

$$\mu_{X_v} = \frac{1}{g} \mathrm{id} \in \mathrm{End}\, \Omega^1(X_v) \cong W_v \otimes \Omega^1(X_v)$$

and similarly for the self-product $X_v \times X_v$ (denoting π_1, π_2 the projections)

$$\Phi = \begin{pmatrix} \frac{1}{g} & -1 \\ -1 & \frac{1}{g} \end{pmatrix} \in \mathrm{End}\left(\pi_1^*\Omega^1(X_v) \oplus \pi_2^*\Omega^1(X_v)\right) \hookrightarrow H^1_{\mathrm{dR}}(X_v \times X_v) \otimes \Omega^1(X_v \otimes X_v).$$

The first Chern class of Φ is the class of the diagonal $\Delta \subset X_v \times X_v$.

Let $\mathbf{1}$ denote the canonical section of the line bundle $\mathscr{O}(\Delta)$ on $X_v \times X_v$. Given any log function $\log_{\mathscr{O}(\Delta)}$ on $\mathscr{O}(\Delta)$ with curvature Φ, we can consider the function G on $X_v \times X_v$ given by

$$G(P,Q) = \log_{\mathscr{O}(\Delta)}(\mathbf{1})(P,Q).$$

It is a Coleman function with singularities along Δ; we call G a *Green function* for X_v.

A Green function G induces a log function on any line bundle $\mathscr{O}(D)$ on X_v by

$$\log_{\mathscr{O}(D)}(\mathbf{1})(Q) = \sum n_i G(P_i, Q)$$

if $D = \sum n_i P_i$ and $\mathbf{1}$ is the canonical section of $\mathscr{O}(D)$. A log function $\log_{\mathscr{L}}$ on the line bundle $\mathscr{L}$ and the resulting metrized line bundle $(\mathscr{L}, \log_{\mathscr{L}})$ are called *admissible* with respect to G if for one (equivalently, any) nonzero rational section s of $\mathscr{L}$, the difference $\log_{\mathscr{L}}(s) - \log_{\mathrm{div}(s)}$ is a constant. Such a constant is denoted by $\iota_{\log_{\mathscr{L}}}(s)$, or $\iota_{\log_v}(s)$ in the case of the trivial line bundle with the log function $\log_v$. It is the analogue of the integral of the norm of s. It follows easily from the definitions that any isomorphism of admissible metrized line bundles is an isometry up to scaling.

Let ω_{X_v} be the canonical sheaf on X_v. The canonical isomorphism $\omega_{X_v} \cong \Delta^*\mathscr{O}(-\Delta)$ gives another way to induce from G a log function $\log^G_{\omega_{X_v}}$ on ω_{X_v}, namely by pullback (and the resulting metrized line bundle has curvature

$(2g-2)\mu_{X_v}$). The requirement that this log function be admissible, together with a symmetry condition, leads to an almost unique choice of G.

Proposition 5.2.1 ([3, Theorem 5.10]). — *There exists a unique up to constant symmetric Green function G with associated curvature Φ such that $(\omega_{X_v}, \log^G_{\omega_{X_v}})$ is an admissible metrized line bundle with respect to G.*

In the following we will arbitrarily fix the constant implied by the Proposition. In our context, the canonical Green function thus determined is, in a suitable sense, defined over E_v [3, Proposition 8.1].

5.3. p-adic Arakelov theory – global aspects. — Let E be a number field with ring of integers $\mathcal{O}_E$. Let $\mathcal{X}/\mathcal{O}_E$ be an arithmetic surface with generic fibre X, that is, $\mathcal{X} \to \mathcal{O}_E$ is a proper regular relative curve and $\mathcal{X} \otimes_{\mathcal{O}_E} E = X$. We assume that $\mathcal{X}$ has good reduction at all place $v|p$, and denote $X_v = \mathcal{X} \otimes E_v$. Fix choices of a ramified p-adic logarithm ℓ and Hodge splittings W_v as in §5.1.

Arakelov line bundles and divisors. — An *Arakelov line bundle* on $\mathcal{X}$ is a pair

$$\widehat{\mathcal{L}} = (\mathcal{L}, (\log_{\mathcal{L}_v})_{v|p})$$

consisting of a line bundle $\mathcal{L}$ on $\mathcal{X}$ together with admissible (with respect to the Green functions of Proposition 5.2.1) log functions $\log_{\mathcal{L}_v}$ on $\mathcal{L}_v = \mathcal{L}|_{X_v}$. We denote by $\mathrm{Pic}^{\mathrm{Ar}}(\mathcal{X})$ the group of isometry classes of Arakelov line bundles on $\mathcal{X}$.

The group $\mathrm{Div}^{\mathrm{Ar}}(\mathcal{X})$ of *Arakelov divisors* on $\mathcal{X}$ is the group of formal combinations

$$D = D_{\mathrm{fin}} + D_\infty$$

where D_{fin} is a divisor on $\mathscr{X}$ and $D_\infty = \sum_{v|p} \lambda_v X_v$ is a sum with coefficients $\lambda_v \in E_v$ of formal symbols X_v for each place $v|p$ of E. To an Arakelov line bundle $\widehat{\mathscr{L}}$ and a nonzero rational section s of $\mathscr{L}$ we associate the Arakelov divisor

$$\widehat{\mathrm{div}}(s) = (s)_{\mathrm{fin}} + (s)_\infty$$

where $(s)_{\mathrm{fin}}$ is the usual divisor of s and $(s)_\infty = \sum_{v|p} \iota_{\log_{\mathscr{L}_v}}(s_v) X_v$. The group $\mathrm{Prin}^{\mathrm{Ar}}(\mathscr{X})$ of *principal* Arakelov divisors on $\mathscr{X}$ is the group generated by the $\widehat{\mathrm{div}}(h)$ for $h \in E(\mathscr{X})^\times$. The Arakelov Chow group of $\mathscr{X}$ is

$$\mathrm{CH}^{\mathrm{Ar}}(\mathscr{X}) = \mathrm{Div}^{\mathrm{Ar}}(\mathscr{X})/\mathrm{Prin}^{\mathrm{Ar}}(\mathscr{X}),$$

and we have an isomorphism

$$\mathrm{Pic}^{\mathrm{Ar}}(\mathscr{X}) \cong \mathrm{CH}^{\mathrm{Ar}}(\mathscr{X})$$

given by $\widehat{\mathscr{L}} \to [\widehat{\mathrm{div}}(s)]$ for any rational section s of $\mathscr{L}$.

The p-adic Arakelov pairing. — Most important for us is the existence of a pairing on $\mathrm{CH}^{\mathrm{Ar}}(\mathscr{X})$, extending the p-adic height pairing of divisors of §5.1. Let $(\cdot,\cdot)_v$ denote the ($\mathbf{Z}$-valued) intersection pairing of cycles on $\mathscr{X}_v$ with disjoint support. The pairing that we consider is the *negative* of the intersection pairing defined by Besser.

Theorem 5.3.1 **(Besser** [3]**).** — *Let $\mathscr{X}/\mathscr{O}_E$ be an arithmetic surface with good reduction above p. For any choice of ramified p-adic logarithm $\ell : E_{\mathbf{A}}^\times/E^\times \to \mathbf{Q}_p$ and Hodge splittings $(W_v)_{v|p}$ as above, there is a symmetric bilinear pairing*[(31)]

$$\langle\cdot,\cdot\rangle^{\mathrm{Ar}} : \mathrm{CH}^{\mathrm{Ar}}(\mathscr{X}) \times \mathrm{CH}^{\mathrm{Ar}}(\mathscr{X}) \to \mathbf{Q}_p$$

[(31)] The notation of [3] is $D_1 \cdot D_2$ for $\langle D_1, D_2\rangle^{\mathrm{Ar}}$.

satisfying:

1. *If D_1 and D_2 are finite and of degree zero on the generic fibre, and one of them has degree zero on each special fibre of $\mathscr{X}$, then*

$$\langle D_1, D_2\rangle^{\mathrm{Ar}} = \langle D_{1,E}, D_{2,E}\rangle,$$

   where $D_{i,E} \in \mathrm{Div}^0(X)$ is the generic fibre of D_i and $\langle\cdot,\cdot\rangle$ denotes the height pairing of Theorem 5.1.1 associated with the same choices of ℓ and $\mathbb{W}_v$.
2. *If $D_{1,\mathrm{fin}}$, $D_{2,\mathrm{fin}}$ have disjoint supports on the generic fibre, then*

$$\langle D_1, D_2\rangle^{\mathrm{Ar}} = \sum_v \langle D_1, D_2\rangle_v^{\mathrm{Ar}}$$

   where the sum runs over all finite places of E, and the local Arakelov pairings are defined by

$$\langle D_1, D_2\rangle_v^{\mathrm{Ar}} = -(D_1, D_2)_v \ell_v(\pi_v)$$

   for $v \nmid p$ and below for $v|p$.

   If moreover we are in the situation of 1., then for each place v

$$\langle D_1, D_2\rangle_v^{\mathrm{Ar}} = \langle D_{1,E}, D_{2,E}\rangle_v.$$

3. *In the situation of 2., if moreover $D_1 = \widehat{\mathrm{div}}(h)$ is the Arakelov divisor of a rational function h, then*

$$\langle D_1, D_2\rangle_v^{\mathrm{Ar}} = \ell_v(h(D_{2,\mathrm{fin}}))$$

   for all places v.

For completeness, we give the description of the local pairing at $v|p$ of divisors with disjoint supports. If $\ell_v = t_v \circ \log_v$ as in (5.2.1) and G_v is the

Green function on $X_v \times X_v$, we have $\langle D, X_w\rangle_v^{\mathrm{Ar}} = 0$ if $v \neq w$, $\langle X_v, X_v\rangle_v^{\mathrm{Ar}} = 0$, $\langle D, \lambda_v X_v\rangle_v^{\mathrm{Ar}} = (\deg D_E) t_v(\lambda_v)$ and if D_1, D_2 are finite divisors with images $D_{1,v} = \sum n_i P_i$, $D_{2,v} = \sum m_j Q_j$ in X_v then

$$\langle D_1, D_2\rangle_v^{\mathrm{Ar}} = -\sum_{i,j} n_i m_j t_v(G_v(P_i, Q_j)).$$

In fact, in [3] it is proved directly that the global Arakelov pairing and its local components at p coincide with the global and local height pairings of Coleman–Gross [10]. The latter coincide with the Zarhin–Nekovář pairings by [2].

6. Heegner points on Shimura curves

In this section we describe our Shimura curve and construct Heegner points on it, following [45, §§1-2], to which we refer for the details (see also [46, §5], and [8] for the original source of many results on Shimura curves). We go back to our usual notation, so F is a totally real number field of degree g, N is an ideal of $\mathcal{O}_F$, E is a CM extension of F of discriminant Δ coprime to $2Np$, and ε is its associated Hecke character.

6.1. Shimura curves. — Let B be a quaternion algebra over F which is ramified at all but one infinite place. Then we can choose an isomorphism $B \otimes \mathbf{R} \cong M_2(\mathbf{R}) \oplus \mathbf{H}^{g-1}$, where $\mathbf{H}$ is the division algebra of Hamilton quaternions. There is an action of $B^\times$ on $\mathfrak{H}^{\pm} = \mathbf{C} - \mathbf{R}$ by Möbius transformations via the map $B^\times \to \mathbf{GL}_2(\mathbf{R})$ induced from the above isomorphism. For each open subgroup K of $\widehat{B}^\times = (B \otimes_F \widehat{F})^\times$ which is compact modulo $\widehat{F}^\times$ we then

have a **Shimura curve**

$$M_K(\mathbf{C}) = B^\times \backslash \mathfrak{H}^\pm \times \widehat{B}^\times / K,$$

where $\mathfrak{H}^\pm = \mathbf{C}\backslash\mathbf{R}$. Unlike modular curves, the curves M_K do not have a natural moduli interpretation. However, by the work of Carayol [8], $M_K(\mathbf{C})$ has a finite map[32] to another (unitary) Shimura curve $M'_{K'}(\mathbf{C})$ which, if the level K' is small enough, has an interpretation as the moduli space of certain quaternionic abelian varieties with extra structure. As a consequence, the curve $M_K(\mathbf{C})$ has a canonical model M_K defined over F (it is connected but not, in general, geometrically connected), and a proper regular integral model[33] $\mathscr{M}_K$ over $\mathscr{O}_F$; if v is a finite place where B is split, then $\mathscr{M}_K$ is smooth over $\mathscr{O}_{F,v}$ if K_v is a maximal compact subgroup of B_v and K^v is sufficiently small.

The order R and the curve X. — Assume that $\varepsilon(N) = (-1)^{g-1}$. Then the quaternion algebra $\mathbf{B}$ over $\mathbf{A}_F$ ramified exactly at all the infinite places and the finite places $v|N$ such that $\varepsilon(v) = -1$ is *incoherent*, that is, it does not arise via extension of scalars from a quaternion algebra over F. On the other hand, for any embedding $\tau\colon F \hookrightarrow \mathbf{R}$, there is a *nearby* quaternion algebra $B(\tau)$ defined over F and ramified at τ and the places where $\mathbf{B}$ is ramified. Fix any embedding $\rho\colon E \to B$, and let R be an order of $\widehat{B} = \widehat{B}(\tau)$ which contains $\rho(\mathscr{O}_E)$ and has discriminant N (this is constructed in [45, §1.5.1]). Then the curve X over F of interest to us is the (compactification of) the curve M_K defined above for the subgroup $K = \widehat{F}^\times \widehat{R}^\times \subset \widehat{B}$; that is, for each

[32] Which is an embedding if $K \supset \widehat{F}^\times$.

[33] In the modular curve case $F = \mathbf{Q}$, $\varepsilon(v) = 1$ for all $v|N$, M_K and $\mathscr{M}_K$ are proper only after the addition of finitely many cusps. (We caution the reader that Carayol [8] uses the notation $\mathscr{M}_K$ to denote instead the set of geometrically connected components of M_K.)

embedding $\tau: F \to \mathbf{C}$, we have

$$X(\mathbf{C}) = B(\tau)^\times \backslash \mathfrak{H}^\pm \times \widehat{B}^\times / \widehat{F}^\times \widehat{R}^\times \cup \{\text{cusps}\}. \tag{6.1.1}$$

The finite set of cusps is nonempty only in the classical case where $F = \mathbf{Q}$, $\varepsilon(v) = 1$ for all $v|N$ so that $X = X_0(N)$. In what follows we will not aggravate the notation with the details of this particular case, which poses no additional difficulties and is already treated in the original work of Perrin-Riou [30].

We denote by $\mathscr{X}$ the canonical model of X over $\mathscr{O}_F$.

Hecke correspondences. — Let m be an ideal of $\mathscr{O}_F$ which is coprime to N and the ramification set of B. Let K_m (respectively K_1) be the the group of those $g \in \widehat{\mathscr{O}}_B$ such that $g_v = 1$ away from m and $\det g$ generates m (respectively, is invertible) at the places dividing m. Then the Hecke operator $T(m)$ on X is defined by

$$T(m)[(z,g)] = \sum_{\gamma \in K_m/K_1} [(z, g\gamma)]$$

under the complex description (6.1.1).

Let $\mathbf{T}'_N$ be the algebra generated by the $T(m)$. Then by the Jacquet–Langlands correspondence, $\mathbf{T}'_N$ is a quotient of the Hekce algebra on Hilbert modular forms $\mathbf{T}_N$ (hence the names $T(m)$ are justified). It acts by correspondences on $X \times X$, and taking Zariski closures of cycles on $\mathscr{X} \times \mathscr{X}$ extends the action to $\mathscr{X}$.

6.2. Heegner points. — The curve X defined above has a distinguished collection of points defined over algebraic extensions of E: we briefly describe it, referring the reader to [45, §2] for more details.

A point y of X is called a **CM point** with multiplication by E if it can be represented by $(\sqrt{-1}, g) \in \mathfrak{H}^{\pm} \times \widehat{B}^{\times}$ via (6.1.1). The order

$$\mathrm{End}(y) = g\widehat{R}g^{-1} \cap \rho(E)$$

in $E = \rho(E)$ is defined independently of the choice of g, and

$$\mathrm{End}(y) = \mathscr{O}_E[c] = \mathscr{O}_F + c\mathscr{O}_F$$

for a unique ideal c of $\mathscr{O}_F$ called the **conductor** of y. We say that the point $y = [(\sqrt{-1}, g)]$ has the *positive orientation* if for every finite place v the morphism $t \to g^{-1}\rho(t)g$ is $R_v^{\times}$-conjugate to ρ in $\mathrm{Hom}(\mathscr{O}_{E,v}, R_v)/R_v^{\times}$.[34] Let Y_c be the set of positively oriented CM points of conductor c. By the work of Shimura and Taniyama, it is a finite subscheme of X defined over E, and the action of $\mathrm{Gal}(\overline{Q}/E)$ is given by

$$\sigma([(\sqrt{-1}, g)]) = [(\sqrt{-1}, \mathrm{rec}_E(\sigma)g)],$$

where $\mathrm{rec}_E : \mathrm{Gal}(\overline{E}/E) \to \mathrm{Gal}(\overline{E}/E)^{\mathrm{ab}} \xrightarrow{\sim} E^{\times}\backslash\widehat{E}^{\times}$ is the reciprocity map of class field theory. If $y = [(\sqrt{-1}, g)]$ has conductor c, then the action factors through

$$\mathrm{Gal}(H[c]/E) \cong E^{\times}\backslash\widehat{E}^{\times}/\widehat{F}^{\times}\widehat{\mathscr{O}}_E[c]^{\times}$$

where $H[c]$ is the ring class field of E of conductor c; the action of this group on Y_c is simply transitive.

[34] This set has two elements only if B is ramified at v (the other element is called the negative orientation at v); otherwise it has one element the condition at v is empty.

For each nonzero ideal c of $\mathcal{O}_F$, let $u(c) = [\mathcal{O}_E[c]^\times : \mathcal{O}_F^\times]$ and define the divisor

$$\eta_c = u(c)^{-1} \sum_{y \in Y_c} y. \tag{6.2.1}$$

Let $\eta = \eta_1$. By the above description of the Galois action on CM points , each divisor η_c is defined over E.

A **Heegner point** $y \in X(H)$ is a positively oriented CM point with conductor 1. We can use the embedding $\iota\colon X \to J(X) \otimes \mathbf{Q}$ to define the point

$$[z] = \iota(\eta) = [\eta] - h[\xi] \in J(X)(E) \otimes \mathbf{Q}$$

where h is a number such that $[z]$ has degree zero in each geometrically connected component of X, and $[\xi]$ is the Hodge class of the Introduction (see below for more on the Hodge class).

Arakelov Heegner divisors. — The Heegner divisor on X can be refined to an Arakelov divisor $\hat{z}$ having degree zero on each irreducible component of each special fibre. On a suitable Shimura curve $\widetilde{X} \xrightarrow{\pi} X$ of deeper level away from $N\Delta_{E/F}$, we can give an explicit description of the pullback $\hat{\widetilde{z}}$ of $\hat{z}$ and of the Hodge class as follows.

After base change to a suitable quadratic extension F' of F, we have an embedding $\widetilde{X} \hookrightarrow \widetilde{X}'$ of $\widetilde{X} = M_{\widetilde{K}}$ into the unitary Shimura curve $\widetilde{X}' = M'_{\widetilde{K'}}$ parametrizing abelian varieties of dimension $4[F : \mathbf{Q}]$ with multiplication by the ring of integers $\mathcal{O}_{B'}$ of $B \otimes_F F'$ and some extra structure. Then by the Kodaira-Spencer map, we have an isomorphism $\omega_{\widetilde{X}'} \cong \det \mathrm{Lie}\, \mathscr{A}^\vee$, where $\mathscr{A} \to \widetilde{X}'$ is the universal abelian scheme and the determinant is that of an

$\mathscr{O}_F$-module of rank 4 (the structure of $\mathscr{O}_F$-module coming from the multiplication by $\mathscr{O}_{B'}$ on $\mathscr{A}$). This gives a way[35] of extending the line bundle $\omega_{\widetilde{X}'}$ to the integral model $\widetilde{\mathscr{X}'}$ and to a line bundle $\mathscr{L}$ on $\widetilde{\mathscr{X}}$. For each finite place $v|p$ we endow $\mathscr{L}|_{\widetilde{X}_v}$ with the canonical log functions $\log_{\mathscr{L},v}$ coming from the description $\mathscr{L}|_{\widetilde{X}_v} = \omega_{\widetilde{X}_v}$ and a fixed choice of Hodge splittings on $\widetilde{X}$. We define $[\hat{\tilde{\xi}}] \in \mathrm{CH}^{\mathrm{Ar}}(\widetilde{\mathscr{X}}) \otimes \mathbf{Q}$ to be the class of $(\mathscr{L}, (\log_{\mathscr{L}})_{v|p})$ divided by its degree, $[\tilde{\xi}]$ to be its finite part, and $\hat{\tilde{\xi}}$ to be any Arakelov divisor in its class.

Then the Heegner Arakelov divisor $\hat{\tilde{z}} \in \mathrm{Div}^{\mathrm{Ar}}(\mathscr{X} \otimes \mathscr{O}_E)$ is described by

$$\hat{\tilde{z}} = \hat{\tilde{\eta}} - h\hat{\tilde{\xi}} + Z, \tag{6.2.2}$$

where $\hat{\tilde{\eta}}$ is the Zariski closure in $\mathscr{X} \otimes \mathscr{O}_E$ of the pullback of η to $\widetilde{X}$, and Z is a finite vertical divisor uniquely determined by the requirement that $\hat{\tilde{z}}$ should have degree zero on each irreducible component of each special fibre.

6.3. Hecke action on Heegner points. — Recall from §1.4 the spaces of Fourier coefficients $\mathscr{D}_{\mathscr{N}} \subset \mathscr{S}$, the arithmetic functions $\sigma_1, r \in \mathscr{D}_N$, and the space $\overline{\mathscr{S}}$ which is a quotient of $\mathscr{S}/\mathscr{D}_N$. The action of Hecke operators on the Arakelov Heegner divisor is described as follows.

Proposition 6.3.1. *— Let m be an ideal of $\mathscr{O}_F$ coprime to N. We have*

1. *$T(m)\eta = \sum_{c|m} r(m/c)\eta_c$.*
2. *Let $\eta_c^0 = \sum_{\mathscr{O}_F \neq d|c} \eta_d$, and let $T^0(m)\eta = \sum_{c|m} \varepsilon(c)\eta^0_{m/c}$. Then η and $T^0(m)\eta$ have disjoint support and if m is prime to $N\Delta$ then $T(m)\eta = T^0(m)\eta + r(m)\eta$.*

[35] See [45, §4.1.3, §1] for more details on this construction.

3. *$T(m)[\xi] = \sigma_1(m)[\xi]$ and $m \mapsto T(m)[\hat{\tilde{\xi}}]$ is zero in $\overline{\mathscr{S}} \otimes \mathrm{CH}^{\mathrm{Ar}}(\widetilde{\mathscr{X}})$.*
4. *The arithmetic function $m \mapsto T(m)Z$ is zero in $\overline{\mathscr{S}} \otimes \mathrm{Div}^{\mathrm{Ar}}(\mathscr{X})$.*

Proof. — Parts 1., 2. and 4. are proved in [45, §4]. For part 3., we switch to the curve $\widetilde{X}$. By definition $[\hat{\tilde{\xi}}]$ is a multiple of the class of the Arekelov line bundle $\mathscr{L} = \det \mathrm{Lie}\, \mathscr{A}^\vee$ on $\widetilde{\mathscr{X}}$ with the canonical log functions on $\mathscr{L}_v \cong \omega_{\widetilde{X}_v}$, where $\mathscr{A} \to \mathscr{X}$ is the universal abelian scheme. We view $T(m)$ as a finite algebraic correspondence of degree $\sigma_1(m)$ induced by the subscheme $\widetilde{\mathscr{X}}_m \subset \widetilde{\mathscr{X}} \times \widetilde{\mathscr{X}}$ of pairs $(A, A/D)$ where D is an 'admissible submodule' of A of order m (see [45, §1.4] for the definition and more details). If $p_1, p_2 \colon \widetilde{\mathscr{X}}_m \to \widetilde{\mathscr{X}}$ are the two projections, then we have

$$T(m)\mathscr{L} = N_{p_1} p_2^* \mathscr{L},$$

and the log functions on $T(m)\mathscr{L}|_{\widetilde{X}_v}$ are the ones induced by this description.[36]

Let $\pi \colon \mathscr{A}_1 \to \mathscr{A}_2$ be the universal isogeny over $\widetilde{\mathscr{X}}_m$. As $p_i^* \mathscr{L} = \det \mathrm{Lie}\, \mathscr{A}_i^\vee$, we have an induced a map

$$\psi_m = N_{p_2} \pi^* \colon T(m)\mathscr{L} \to N_{p_2} p_2^* \mathscr{L} = \mathscr{L}^{\sigma_1(m)},$$

and [45, §4.3] shows that $\psi_m(T(m)\mathscr{L}) = c_m \mathscr{L}^{\sigma_1(m)}$ where $c_m \subset \mathscr{O}_F$ is an ideal with divisor $[c_m]$ on $\mathrm{Spec}\, \mathscr{O}_F$ such that $m \to [c_m]$ is a σ_1-derivative (§1.4), hence zero in $\overline{\mathscr{S}} \otimes \mathrm{Div}(\mathrm{Spec}\, \mathscr{O}_F) \subset \overline{\mathscr{S}} \otimes \mathrm{Div}^{\mathrm{Ar}}(\mathscr{X})$.

[36] The necessary conditions on the curvature detailed in §5.2 are met since in this case the curvature of $T(m)\mathscr{L}_v$ is easily seen to be $T(m)\mu_{\widetilde{X}_v} = \sigma_1(m)\mu_{\widetilde{X}_v}$.

We complete the proof by showing that, for each $v|p$, the difference of log functions

$$\psi_m^* \log_{\mathscr{L}_v^{\sigma_1(m)}} - \log_{T(m)\mathscr{L}_v} \tag{6.3.1}$$

on the line bundle $T(m)\mathscr{L}_v$ on $\widetilde{X}_v$ is also a σ_1-derivative when viewed as a function of m. It is enough to show this after pullback via p_1 on $\widetilde{X}_m$, where (denoting pulled back objects with a prime) the map ψ'_m decomposes as

$$\psi'_m = \otimes_D \pi_D^* \colon \otimes_D \det \mathrm{Lie}\,(\mathscr{A}'/D)^\vee \to \det \mathrm{Lie}\,\mathscr{A}'^{\vee \otimes \sigma_1(m)}$$

where the tensor product runs over admissible submodule schemes of order m of $\mathscr{A}'$ (since base change via p_1 splits the cover p_1, there are exactly $\sigma_1(m)$ of those). Now the difference (6.3.1) is the sum of the $\sigma_1(m)$ differences

$$(\pi_D^*)^* \log_{\mathscr{L}} - \log_{\mathscr{L}},$$

which are all the same since they are permuted by the Galois group of p_1. As π_D^* acts by multiplication by $(\deg \pi_D)^{1/2} = \mathbf{N}(m)^2$, by (5.2.2) each of these differences is $2\log_v \mathbf{N}(m)$ so that (6.3.1) equals

$$2\sigma_1(m) \log_v \mathbf{N}(m)$$

which is indeed a σ_1-derivative. □

Notice that $[\hat{\tilde{\xi}}]$ is only defined as a divisor *class*. Now we pick any divisor $\hat{\xi}$ in its class and *define* the divisor $T(m)\hat{\tilde{\xi}}$ to be a divisor in the class $T(m)[\hat{\tilde{\xi}}]$ such that $m \mapsto T(m)\hat{\tilde{\xi}}$ is zero in $\overline{\mathscr{S}} \otimes \mathrm{Div}^{\mathrm{Ar}}(\mathscr{X})$. This is legitimate for our purposes since the global Arakelov pairing is defined on divisor classes.

7. Heights of Heegner points

Let Ψ be the modular form of level N with Fourier coefficients given by the p-adic height pairing $\langle z, T(m)z\rangle$ (it is a modular form because of Lemma 1.3.1 and the fact that the quaternionic Hecke algebra $\mathbf{T}'_N$ is a quotient of $\mathbf{T}_N$, as explained at the end of §??). We will compute the heights of Heegner points, with the goal of showing (in §8) that $L_{f_0}(\Phi')$ and $L_{f_0}(\Psi)$ are equal up to the action of some Hecke operators. The proof of the main theorem will follow.

The strategy is close to that of Perrin-Riou, namely we separate the local contributions to Ψ from primes above p, writing $\Psi \sim \Psi_{\text{fin}} + \Psi_p$; using the computations of [45, 46] we show that Φ' and Ψ_{fin} are "almost" equal, while the contribution of Ψ_p is shown to vanish. The absence of cusps however poses some difficulties, that we circumvent through the use of p-adic Arakelov theory.

We will work throughout in the space $\overline{\mathscr{S}} \supset \overline{\mathscr{S}}_N$ of §1.4, using the simbol $\sim$ to denote equality there; we abuse notation by using the same name for a modular form and its image in $\overline{\mathscr{S}}_N$.

The height pairings (and the accompanying Arakelov pairings) on the base change of X to E that will be considered are the ones associated to choices of Hodge splittings on $V_w = H^1_{\text{dR}}(X_w/E_w)$ $(w|p)$ compatible with the canonical choices on the ordinary subrepresentations $V_{f,w}$, and to a "cyclotomic" p-adic logarithm given by $\ell = \ell_F \circ \mathfrak{N}\colon E^\times \backslash E^\times_{\mathbf{A}^\infty} \to \mathbf{Q}_p$ for some

$$\ell_F\colon F^\times \backslash F^\times_{\mathbf{A}^\infty} \to \mathbf{Q}_p.$$

(These data will be omitted from the notation).

As mentioned before, the Shimura curve X and its integral model $\mathscr{X}$ may not be fine enough for the needs of Arakelov and intersection theory, so that we may need to pass to a Shimura curve of deeper level $\widetilde{\mathscr{X}} \overset{\pi}{\to} \mathscr{X}$ and consider the pullbacks $\widetilde{\eta}$ of the divisors η, etc. Then notation such as $\langle \hat{\eta}, T^0(m)\hat{\eta}\rangle^{\mathrm{Ar}}$ is to be properly understood as $\langle \hat{\widetilde{\eta}}, T^0(m)\hat{\widetilde{\eta}}\rangle^{\mathrm{Ar}}/\deg \pi$.

7.1. Local heights at places not dividing p. — The next two results will be used to show the main identity.

Lemma 7.1.1. — *In the space $\overline{\mathscr{S}}_N$ we have*

$$\langle z, T(m)z\rangle = \langle \hat{z}, T(m)\hat{z}\rangle^{Ar} \sim \langle \hat{\eta}, T^0(m)\hat{\eta}\rangle^{Ar}.$$

Proof. — First observe that by Lemma 1.3.1, the first member indeed belongs to $\overline{\mathscr{S}}_N$. The first equality is a consequence of Theorem 5.3.1.1 and the construction of $\hat{z}$. The second part follows from expanding the second term for m prime to $N\Delta$ according to (6.2.2) and observing that the omitted terms are zero in $\overline{\mathscr{S}}_N$ by Proposition 6.3.1. □

We can therefore write

$$\Psi \sim \sum_w \Psi_w = \sum_v \Psi_v = \Psi_{\mathrm{fin}} + \Psi_p \tag{7.1.1}$$

in $\overline{\mathscr{S}}$, with the first sum running over the finite places w of E, the second sum running over the finite places v of F, and

$$\Psi_w(m) = \langle \hat{\eta}, T^0(m)\hat{\eta}\rangle_w^{\mathrm{Ar}}, \quad \Psi_v = \sum_{w|v} \Psi_w, \quad \Psi_{\mathrm{fin}} = \sum_{v\nmid p} \Psi_v, \quad \Psi_p = \sum_{v|p} \Psi_v.$$

(We are exploiting the fact that for m prime to $N\Delta$ the divisors $\hat{\eta}$ and $T^0(m)\hat{\eta}$ have disjoint supports so that we can apply Theorem 5.3.1.2.)

For each prime $\wp$ of F above p, we define an operator[37] on $\mathscr{S}$

$$\mathscr{R}_\wp = (U_\wp - 1), \quad \mathscr{R}_p = \prod_{\wp|p} \mathscr{R}_\wp$$

Proposition 7.1.2. — *The arithmetic function $\mathscr{R}_p^3 \Psi_p$ belongs to $\overline{\mathscr{S}}_N \subset \overline{\mathscr{S}}$, and we have*

$$L_{f_0}(\mathscr{R}_p^3 \Psi_p) = 0.$$

The proof of this crucial fact will occupy §7.2.

Proposition 7.1.3. — *In the space $\overline{\mathscr{S}}$ we have*

$$\Psi_{\mathrm{fin}} \sim \sum_{v \nmid p} \Psi_v + h,$$

where h is a modular form which is killed by L_{f_0}; the sum runs over the finite places of F and the summands are given by:

1. *If $v = \wp$ is inert in E, then*

$$\Psi_v(m) = \sum_{\substack{n \in N m^{-1} \Delta^{-1} \\ \varepsilon_{F,v}((n-1)n) = 1 \forall v | \Delta \\ 0 < n < 1}} 2^{\omega_\Delta(n)} r((1-n)m\Delta) r(nm\Delta/N\wp)(v(nm/N)+1) \ell_{F,v}(\pi_v).$$

2. *If $v = \wp | \Delta$ is ramified in E, then*

$$\Psi_v(m) = \sum_{\substack{n \in N m^{-1} \Delta^{-1} \\ \varepsilon_v((n-1)n) = -1 \\ \varepsilon_w((n=1)n) = 1 \forall v \neq w | \Delta \\ 0 < n < 1}} 2^{\omega_\Delta(n)} r((1-n)m\Delta) r(nm\Delta/N)(v(nm)+1) \ell_v(\pi_v).$$

3. *If v is split in E, then*

$$\Psi_v(m) = 0.$$

[37] This is different from the operator bearing the same name in [30].

Proof. — For m prime to $N\Delta$ we have $\Psi_{\mathrm{fin}}(m) = \sum_{w\nmid p}\langle\hat{\eta}, T^0(m)\hat{\eta}\rangle_w^{\mathrm{Ar}}$ (the sum running over all finite places w of E). By Theorem 5.3.1.2, up to the factor $\ell_{F,v}(\pi_v)$ (which equals $\ell_w(\pi_w)$ or its half for each place w of E above v), each term is given by an intersection multiplicity $(\hat{\eta}, T(m)\hat{\eta})_w$, which is computed by Zhang.

When $v(N)\leq 1$ for all v which are not split in E, the result is summarised in [45, Proposition 5.4.8]; in this case, the values obtained there are equivalent to the asserted ones by [45, Proposition 7.1.1 and Proposition 6.4.5], and there is no extra term h. In fact (and with no restriction on N), these values appear also as local components ${}^{\mathrm{C}}\Phi'_v$ at finite places of a form ${}^{\mathrm{C}}\Phi'$ of level N which is a kernel of the Rankin–Selberg convolution for the central derivative $L'(f_E, 1)$ of the complex L-function.

In general, [46, Lemma 6.4.3] proves that[38]

$$\frac{\Psi_v}{\ell_{F,v}(\pi_v)} \sim \frac{{}^{\mathrm{C}}\Phi'^{\sharp}_v}{\log \mathbf{N}(|\pi_v|^{-1})} + {}_vh, \tag{7.1.2}$$

where ${}_vh$ is a modular form with zero projection onto the f-eigenspace (see the discussion at the very end of [46]; the forms ${}_vh$ come from intersections at bad places), and ${}^{\mathrm{C}}\Phi'^{\sharp}$ is a form of level $N\Delta$ which is a kernel for the complex Rankin–Selberg convolution in level $N\Delta$ (in particular, it is modular and $\mathrm{Tr}_\Delta({}^{\mathrm{C}}\Phi'^{\sharp}) = {}^{\mathrm{C}}\Phi'$). Applying the operator Tr_Δ in (7.1.2) we recover the asserted formula. □

[38] We are adapting the notation to our case. In [46], the form f is denoted by ϕ, the functions ${}_vh$ are denoted by ${}_vf$.

7.2. Local heights at p. — Here we prove Propostion 7.1.2 which shows that the contirbution of the p-part of the height is negligible. We fix a a prime $\wp$ of F dividing p.

Let the divisors η_c be as in (6.2.1), and denote

$$H_s = H[\wp^s], \qquad u_s = u(\wp^s).$$

***Proposition 7.2.1* (Norm relation).** — *Let $m = m_0\wp^n$ be an ideal of F with m_0 prime to $\wp N$. We have*

$$[T(m\wp^{r+2}) - 2T(m\wp^{r+1}) + T(m\wp^r)](\eta) = u_{n+r+2}^{-1} T(m_0)\mathrm{Tr}_{H_{n+r+2}/E}([y']),$$

as divisors on X, where $y' \in X(H_{n+r+2})$ is any CM point of conductor $\wp^{n+r+2}$.

Proof. — By the multiplicativity of Hecke operators it is enough to prove the statement for $m_0 = 1$. A simple computation based on Proposition 6.3.1 shows that the left-hand side is equal to η_{n+r+2}. Since the Galois action of $\mathrm{Gal}(H_{n+r+2}/E)$ is simply transitive on $Y_{\wp^{n+r+2}}$, the right-hand side is also equal to η_{n+r+2}. □

Lemma 7.2.2. — *Let v a place of E dividing $\wp$, and let $h \in E_v(X)$ be a rational function whose reduction at v is defined and nonzero. Then we have*

$$e_\wp \circ \mathscr{R}^3_\wp \langle \widehat{\mathrm{div}}(h), T^0(m)\hat{\eta}\rangle_v^{Ar} = e_\wp \circ \mathscr{R}^3_\wp \langle \mathrm{div}(h), T(m)z\rangle_v = 0$$

Proof. — We will show that $U_\wp^s \mathscr{R}^3_\wp \langle \widehat{\mathrm{div}}(h), T^0(m)\hat{\eta}\rangle_v^{\mathrm{Ar}}$ tends to 0 in the p-adic topology, thereby proving the vanishing of the first expression; the proof for the second expression is similar, cf. [30, Lemme 5.4]. We may

assume m prime to $\wp N\Delta$. As $\mathscr{R}^2_\wp r(m)=0$, Proposition 7.2.1 gives

$$U^s_\wp \mathscr{R}^2_\wp \eta = u_{s+2}^{-1}\mathrm{Tr}_{H_{s+2}/E} y_{s+2}$$

where $y_{s+2}\in Y_{\wp^{s+2}}$. For s large enough the divisor of h is supported away from y_s and its conjugates. Then by Theorem 5.3.1.3 we have

$$\begin{aligned} U^s_\wp \mathscr{R}^2_\wp \langle \widehat{\mathrm{div}}(h), T^0(m)\hat{\eta}\rangle_v^{\mathrm{Ar}} &= u_{s+2}^{-1}\ell_v(h(T^0(m)y_{s+2})) \\ &= u_{s+2}^{-1}\sum_{w|v}\ell_v(N_{H_{s+2,w}/E_v}h(y_{s+2})), \end{aligned}$$

where w runs over the places of H above v (which are identified with the places of H_{s+2} above v, since H_{s+2}/H is totally ramified above $\wp$).

For any $w|v$ we have

$$\mathscr{R}_\wp \ell_v N_{H_{s+2,w}/E_v} h(y_{s+2}) = \ell_v \circ N_{H_w/E_v}(N_{H_{s+3,w}/H_w}(y_{s+3})/N_{H_{s+2,w}/H}(y_{s+2})).$$

Suppose that

(∗) the w-adic valuation of $N_{H_{s,w}/H_w}(h(y_s))$ is independent of s.

Then each w-summand in the expression of interest is the product of u_{s+2}^{-1} (which is eventually constant in s) and the p-adic logarithm of a unit which is a norm from an extension of E_v whose ramification degree grows linearly in s; hence its p-adic valuation grows linearly in s, which proves the Lemma.

It remains to prove (∗). This can be done using the moduli interpretation of CM points on X (see [45, §§1-2]) and a degenerate case of Gross's theory of quasi-canonical liftings, parallelling the proof of [30, Lemme 5.5]. As the generalisation to our case presents no difficulties but a precise exposition

would require a lengthy discussion of the moduli interpretation which we do not need elsewhere, we omit the details. □

Lemma 7.2.3. — *Let $e'_\wp$ denote the operator $e'_\wp = e_\wp \circ \mathscr{R}^3_\wp$ on $\overline{\mathscr{S}}_N$.*

1. *The operator $e'_\wp$ extends to the subspace of $\overline{\mathscr{S}}$ generated by the functions*

$$m \mapsto \langle D, T(m)z\rangle_v$$

for $D \in \mathrm{Div}^0(X)(E_v)$. The result only depends on the class $[D] \in J(X)(E_v)$, it is denoted by

$$e'_\wp \langle [D], T(m)z\rangle_v,$$

and is a bounded element of $\overline{\mathscr{S}}_N$.

2. *The arithmetic function*

$$\Psi_v : m \mapsto \langle \hat{\eta}, T^0(m)\hat{\eta}\rangle_v^{\mathrm{Ar}}$$

belongs to the subspace defined in 1., and the value of the operator $e'_\wp$ on it is

$$e'_\wp \langle \hat{\eta}, T^0(m)\hat{\eta}\rangle_v^{\mathrm{Ar}} = e'_\wp \langle [z], T(m)z\rangle_v.$$

Proof. — Since the height pairing is bounded and bounded subsets of $\mathbf{Q}_p$ are compact, we can find a sequence of integers (r_k) such that the sequence of arithmetic functions $U_\wp^{r_k} \mathscr{R}^3_\wp \langle D, T(m)z\rangle_v$ converges for all D to a bounded element of $\overline{\mathscr{S}}$, and this gives the desired extension of $e'_\wp$. The independence on the choice of D in the class $[D]$ follows from Lemma 7.2.2.

To prove that the result is modular, we observe that by Proposition 7.2.1 we have, for each m prime to $N\wp$:

$$U_\wp^{r_k} \mathscr{R}^3_\wp \langle D, T(m)z\rangle_v \sim U_\wp^{r_k} \mathscr{R}^3_\wp \langle \widehat{D}, T(m)\hat{\eta}\rangle_v^{\mathrm{Ar}}$$

$$= u_{r_k+2}^{-1} \langle \widehat{D}, T(m)(\hat{y}_{r_k+3} - \hat{y}_{r_{k+2}}) \rangle_v^{\mathrm{Ar}}$$

for $y_n = \mathrm{Tr}_{H_n/E} \hat{y}'_n \in \mathrm{Div}(X)$ with y'_n a CM point of conductor $\wp^n$. Therefore the m^{th} coefficient of $e'_\wp \langle [D], T(m)z \rangle_v$ is the value at $T(m)$ of the linear form on $\mathbf{T}_{N\wp}$

$$T \mapsto \lim_{k\to\infty} u_{r_k+2}^{-1} \langle \widehat{D}, T(\hat{y}_{r_k+3} - \hat{y}_{r_{k+2}}) \rangle_v^{\mathrm{Ar}}.$$

Then the modularity follows from Lemma 1.3.1.

For the second part we may argue as in the proof of Lemma 7.1.1: in $\overline{\mathscr{S}}$ we have

$$\begin{aligned} \mathscr{R}_\wp^3 \langle \hat{\eta}, T^0(m)\hat{\eta} \rangle_v^{\mathrm{Ar}} &\sim \mathscr{R}_\wp^3 \langle \hat{\eta} + \widehat{\mathrm{div}}(h), T^0(m)\hat{\eta} \rangle_v^{\mathrm{Ar}} \\ &\sim \mathscr{R}_\wp^3 \langle \hat{z} + \widehat{\mathrm{div}}(h), T(m)\hat{z} \rangle_v^{\mathrm{Ar}} = \mathscr{R}_\wp^3 \langle z + \mathrm{div}(h), T(m)z \rangle_v. \end{aligned}$$

□

Proof of Propostion 7.1.2. — Let $V = J(X)(E) \otimes \overline{\mathbf{Q}}$ and write $[z] = z_f + z_{f^\perp}$, with $z_f, z_{f^\perp} \in V$ such that $T(m)z_f = a(f,m)z_f$ for m prime to Np and that the modular q-expansion $m \mapsto T(m)z_{f^\perp} \in \mathscr{S}_N \otimes V$ is orthogonal to f. Let

$$e'_\wp \Psi_v[f](m) = e'_\wp \langle z_f, T(m)z \rangle_v, \qquad e'_\wp \Psi_v[f^\perp](m) = \langle z_{f^\perp}, T(m)z \rangle_v;$$

then by Lemma 7.2.3 we have

$$e'_\wp \Psi_v = e'_\wp \Psi_v[f] + e'_\wp \Psi_v[f^\perp]$$

in $d\overline{\mathscr{S}}_{N\wp}$. It is easy to see that

$$L_{f_0}(e'_\wp \Psi_v[f^\perp]) = 0$$

(cf. [30, Lemme 5.9]). We now show that $e'_{\wp}\Psi_v[f]=0$. The ordinarity assumption and Theorem 5.1.1.4 (cf. [30, Exemple 4.12]) imply that z_f is "almost" a universal norm in the totally ramified $\mathbf{Z}_p$-extension $E^{\ell}_{v,\infty}$ of E_v: that is, after perhaps replacing z_f by an integer multiple, for each layer $E^{\ell}_{v,n}$ we have

$$z_f = \mathrm{Tr}_n(z_n)$$

for some $z_n \in J(X)(E^{\ell}_{v,n})\otimes\overline{\mathbf{Q}}$, where $\mathrm{Tr}_n = \mathrm{Tr}_{E^{\ell}_{v,n}/E_v}$. Then we have

$$e'_{\wp}\Psi_v[f](m) = e'_{\wp}\langle \mathrm{Tr}(z_n), T(m)z_f\rangle_v = e'_{\wp}\langle z_n, T(m)z_f\rangle_{v,n}.$$

where $\langle\cdot,\cdot\rangle_{v,n}$ is the local height pairing on $\mathrm{Div}^0(X)(E^{\ell}_{v,n})$ associated to the logarithm $\ell_{n,v} = \ell_v \circ N_{E^{\ell}_{v,n}/E_v}$. By Theorem 5.1.1, these height pairings form a compatible bounded family: the compatibility implies the second equality just above, and the boundedness means that they take values in $c^{-1}\mathrm{Im}(\ell_n)\subset \mathbf{Z}_p$ for a uniform nonzero constant $c\in\mathbf{Z}_p$. As the extension $E^{\ell}_{v,n}/E_v$ has ramification degree p^n, we have for some nonzero $c'\in\mathbf{Z}_p$

$$e'_{\wp}\Psi_v[f](m) \in c^{-1}\mathrm{Im}(\ell_n) \subset c'^{-1}p^n\mathbf{Z}_p$$

for all n; therefore $e'_{\wp}\Psi_v[f]=0$.

We conclude that

$$L_{f_0}(\mathscr{R}^3_p\Psi_v) = L_{f_0}(e\mathscr{R}^3_p\Psi_v) = \prod_{\wp'\neq\wp}(\alpha_{\wp'}(f)-1)^3 L_{f_0}(e'_{\wp}\Psi_v) = 0.$$

This completes the proof of Propostion 7.1.2.

PART III

MAIN THEOREM AND CONSEQUENCES

8. Proof of the main theorem

In this section we prove Theorem B. First, notice that when $\mathscr{W}$ is anticyclotomic (that is, $\mathscr{W}^c = \overline{\mathscr{W}}$), both sides of the formula are zero: indeed, $L_p(f_E)(\mathscr{W}^s)$ vanishes identically by the functional equation (4.2.3), and

$$\langle z_f, z_f \rangle_{\mathscr{W}} = \langle z_f^c, z_f^c \rangle_{\mathscr{W}\circ c} = -\langle z_f, z_f \rangle_{\mathscr{W}}$$

since by the work of Shimura [34], $[z]^c = W_N[z]$ and $W_N z_f = (-1)^g \tau(f) z_f$, where $\tau(f) = \pm 1$. Therefore it suffices to prove the formula when

$$\mathscr{W} = \mathscr{W}^+ = \nu \circ \mathfrak{N}$$

for some Hecke character ν of F valued in $1 + p\mathbf{Z}_p$.

8.1. Basic case. — First we prove the formula when $\Delta_{E/F}$ is totally odd and each prime $\wp$ of F dividing p is split in E.

Let $\Psi_{\mathscr{W}} \in \overline{\mathscr{S}}_N$ denote the modular form with coefficients $\langle [z], T(m)[z] \rangle_{\mathscr{W}}$, where $\langle \cdot, \cdot \rangle_{\nu\circ\mathfrak{N}}$ is the height pairing on $J(X)(E)$ associated to the p-adic logarithm $\ell_F \circ \mathfrak{N}$, with

$$\ell_F = \frac{d}{ds} \nu^s|_{s=0} : F^\times \backslash F_{\mathbf{A}^\infty}^\times \to \mathbf{Q}_p$$

We compare the Fourier coefficients of $\Phi'_{\mathscr{W}}$ and $\Psi_{\mathscr{W}}$.

Lemma 8.1.1. — *In the space $\overline{\mathscr{S}}_N$ of §1.5, we have*

$$\left(\prod_{\wp|p} U_\wp^4 - U_\wp^2\right)\Phi'_{\mathscr{W}} \sim \left(\prod_{\wp|p}(U_\wp - 1)^4\right)\Psi_{\mathscr{W},\mathrm{fin}} + h,$$

where h is a form killed by L_{f_0}.

Proof. — The coefficients of $\Phi'_{\mathscr{W}}$ and $\Psi_{\mathscr{W},\mathrm{fin}}$ are computed in Proposition 4.5.3 and Proposition 7.1.3. Up to the form h, they look "almost" the same, in that in each case the m^{th} Fourier coefficient is given by a sum of terms indexed by n in a certain finite set, except that in the left-hand side we have the additional restriction $(p, nm) = 1$. We can rewrite the indexing parameter n relative to the right-hand side as $n_0 \prod_{\wp|p} \wp^{t_\wp}$ with $(p, n_0 m) = 1$ and each $t_\wp \geq 0$. Now a simple calculation based on the observation that $r(m_0\wp^t) = r(m_0)(t+1)$ when $\wp \nmid m_0$ shows that the contribution of the terms with some $t_\wp > 0$ vanishes and that the contribution of the remaining terms gives the right-hand side. The details are as in [30, Proof of Proposition 3.20]. □

To proceed to the conclusion, notice that since the functional L_{f_0} is bounded, $L_{f_0} \circ \Phi$ is also a measure. In particular[(39)], it commutes with limits, so that

$$L'_{p,\mathscr{W}}(f_E)(1) = L_{f_0}\left(\frac{d}{ds}\Phi(\mathscr{W}^s)|_{s=0}\right).$$

Since by Proposition 7.1.2 we have

$$L_{f_0}\left(\prod_{\wp|p}(U_\wp - 1)^4 \Psi_{\mathscr{W},p}\right) = 0,$$

[(39)]Recall that a measure is a bounded or equivalently *continuous* functional on continuous functions.

we find for $\mathscr{W} = \nu \circ \mathfrak{N}$ (with $\alpha_{\wp} = \alpha_{\wp}(f)$):

$$\begin{aligned} D_F^2 \prod_{\wp}(\alpha_{\wp}^4 - \alpha_{\wp}^2) L'_{p,\mathscr{W}}(f_E, \mathbf{1}) &= \prod_{\wp}(\alpha_{\wp}^4 - \alpha_{\wp}^2)\left(1 - \frac{1}{\alpha_{\wp}^2}\right)\left(1 - \frac{\mathbf{N}\wp}{\alpha_{\wp}^2}\right) L_{f_0}(\Phi'_{\mathscr{W}}) \\ &= \prod_{\wp}(\alpha_{\wp} - 1)^4 \left(1 - \frac{1}{\alpha_{\wp}^2}\right)\left(1 - \frac{\mathbf{N}\wp}{\alpha_{\wp}^2}\right) L_{f_0}(\Psi_{\mathscr{W}}) \\ &= \prod_{\wp}(\alpha_{\wp} - 1)^4 \left(1 - \frac{1}{\alpha_{\wp}^2}\right) \langle z_f, z_f \rangle_{\mathscr{W}}. \end{aligned}$$

Here, besides the definition of $L_p(f_E)$ (Definition 4.2.1) we have used various properties of the functional L_{f_0} from Lemma 1.5.1 and the observation that the projection onto the f-component of the modular form $\Psi_{\mathscr{W}} \in S_2(K_0(N), \mathbf{Q}_p)$ is $\mathbf{1}_f(\Psi_{\mathscr{W}}) = \langle z_f, z_f \rangle_{\mathscr{W}}$.

This completes the proof of Theorem B when $(\Delta_{E/F}, 2) = 1$ and all primes $\wp | p$ split in E.

8.2. Reduction to the basic case. — The general case, where E is only assumed to satisfy $(\Delta_{E/F}, Np) = 1$, can be reduced to the previous one under the assumption

$$L'_{p,\mathscr{W}}(f_E, \mathbf{1}) \neq 0$$

by the following argument due to Kobayashi [22, Proof of Theorem 5.9] using the complex Gross–Zagier formula (which is known with no restrictions on Δ) and the factorisation $L_p(f_E, \chi \circ \mathfrak{N}) \sim L_p(f, \chi) L_p(f_\varepsilon, \chi)$.

Indeed, by the factorisation the orders of vanishing at the central point of the factors of $L_p(f_E, \nu^s \circ \mathfrak{N})$, will be one (say for $L_p(f)$) and zero (say for

$L_p(f_\varepsilon)$). Then, by the first part of Theorem C[40], the orders of vanishing of $L(f,s)$ and $L(f_\varepsilon,s)$ at $s=1$ will also be one and zero. Moreover the Heegner point $z_{f,E'}$ attached to f and any E' also satisfying $L(f_{\varepsilon_{E'/F}},1)\neq 0$ is non-torsion, and in fact its trace $z_{f,F}=\mathrm{Tr}_{E'/F}(z_{f,E'})$ is non-torsion and $z_{f,E'}$ is up to torsion a multiple of $z_{f,F}$ in $J(X)(E')\otimes\overline{\mathbf{Q}}$. Therefore, by the complex and p-adic Gross–Zagier formulas for a suitable E' satisfying the assumptions of §8.1 and $L(f_{\varepsilon_{E'/F}},1)\neq 0$, we have

$$L'_{p,\nu}(f,\mathbf{1})=\prod_{\wp|p}\left(1-\frac{1}{\alpha_\wp}\right)^2\frac{L'(f,1)}{\Omega_f^+\langle z_{f,F},z_{f,F}\rangle}\langle z_{f,F},z_{f,F}\rangle_\nu$$

where $\langle\cdot,\cdot\rangle_\nu$ is the p-adic height pairing on $J(X)(F)$ attached to ν, and $\langle\cdot,\cdot\rangle$ is the Néron–Tate height (the ratio appearing above belongs to $M_f^\times$ by the Gross–Zagier formula). This allows us to conclude

$$\begin{aligned}
L'_{p,\mathcal{W}}(f_E,\mathbf{1})&=\frac{\Omega_f^+\Omega_{f_\varepsilon}^+}{D_E^{-1/2}\Omega_f}L'_{p,\nu}(f,\mathbf{1})L_p(f_\varepsilon,\mathbf{1})\\
&=D_E^{1/2}\prod_{\wp|p}\left(1-\frac{1}{\alpha_\wp}\right)^2\left(1-\frac{\varepsilon(\wp)}{\alpha_\wp}\right)^2\frac{L'(f,1)L(f_\varepsilon,1)}{\Omega_f\langle z_{f,F},z_{f,F}\rangle}\langle z_{f,F},z_{f,F}\rangle_\nu\\
&=D_F^{-2}\prod_{\wp|p}\left(1-\frac{1}{\alpha_\wp}\right)^2\left(1-\frac{\varepsilon(\wp)}{\alpha_\wp}\right)^2\frac{\langle z_{f,E},z_{f,E}\rangle}{\langle z_{f,F},z_{f,F}\rangle}\langle z_{f,F},z_{f,F}\rangle_\nu\\
&=D_F^{-2}\prod_{\wp|p}\left(1-\frac{1}{\alpha_\wp}\right)^2\left(1-\frac{\varepsilon(\wp)}{\alpha_\wp}\right)^2\langle z_{f,E},z_{f,E}\rangle_{\mathcal{W}}.
\end{aligned}$$

[40]Which can be proved by using the p-adic Gross–Zagier formula attached to a field E' satisfying the assumptions of §8.1.

Remark 8.2.1. — It is natural to conjecture that when $L'_{p,\mathscr{W}}(f_E, \mathbf{1}) = 0$ we should have $\langle z_f, z_f \rangle_{\mathscr{W}} = 0$. However in this case the above argument fails because, without knowledge of the nontriviality of the p-adic height pairing, the vanishing of $L_p(f_E, \mathscr{W}^s)$ to order ≥ 2 does not imply a similar high-order vanishing for $L(f_E, s)$.

9. Periods and the Birch and Swinnerton-Dyer conjecture

As seen in the Introduction, the application of our result to the Birch and Swinnerton-Dyer formula rests on a conjectural relation among the periods of f and the associated abelian variety A. Here we would like to elaborate on this conjecture and its arithmetic consequences. We retain the notation of the Introduction, and set $\dim A = [M : \mathbf{Q}] = d$.

9.1. Real periods. — The conjecture on periods stated in the Introduction can be refined to a conjecture on rationality rather than algebraicity. First we need to precisely define the automorphic periods $\Omega^+_{f^\sigma}$, for $\sigma \in \mathrm{Hom}(M_f, \mathbf{C})$; they are naturally defined as elements of $\mathbf{C}^\times / M^\times$ (see [32] for a modern exposition): one can choose them "covariantly" in order to have $\prod_\sigma \Omega^+_{f^\sigma}$ defined up to $\mathbf{Q}^\times$, or define directly the product as follows. Let $\mathscr{H}_N = Z(\mathbf{A}) \backslash \mathbf{GL}_2(\mathbf{A}) / K_0(N) K_\infty$ be the Hilbert modular variety[41] of level N. Then the perfect pairing of $\mathbf{Q}$-vector spaces

$$H_g(\mathscr{H}_N, \mathbf{Q})^+ \times S_2(K_0(N), \mathbf{Q}) \to \mathbf{C} \tag{9.1.1}$$

(where "+" denotes the intersection of the $+1$-eigenspaces for the complex conjugations) decomposes under the diagonal action of $\mathbf{T}_N$ into $\mathbf{Q}$-rational

[41] We are ignoring cusps, which only appear in the case $F = \mathbf{Q}$ of least interest to us.

blocks parametrised by the Galois-conjugacy classes of eigenforms. Then

$$\prod_\sigma \Omega^+_{f^\sigma} \in \mathbf{C}^\times/\mathbf{Q}^\times$$

is $(2\pi i)^{dg}$ times the discriminant of the pairing on the rational block corresponding to $\{f^\sigma\}_\sigma$. (The individual $\Omega^+_{f^\sigma} \in \mathbf{C}^\times/M^\times$ are defined as the discriminants of (9.1.1) on $\overline{\mathbf{Q}}$-rational $\mathbf{T}_N$-eigenblocks.)

Conjecture 9.1.1. — *We have*

$$\Omega_A \sim \prod_\sigma \Omega^+_{f^\sigma}$$

in $\mathbf{C}^\times/\mathbf{Q}^\times$.

The conjecture is also made by Yoshida [41] up to algebraicity. When A has complex multiplication, it has been proved by Blasius [5]. It is also known when $F = \mathbf{Q}$; before discussing that, let us translate it into a language closer to conjectures of Shimura.

For each $\tau\colon F \to \mathbf{R}$, let $f_{B(\tau)}$ be the Jacquet-Langlands transfer of f to a rational form on the quaternion algebra $B(\tau)/F$ defined in the Introduction (recall that $B(\tau)$ is ramified at all infinite places except τ), and let X be our Shimura curve. Then A is (up to isogeny) a quotient ϕ of $J(X)$, and for each embedding τ we can write

$$\phi^*\omega_A = c_\tau \bigwedge_\sigma 2\pi i f^\sigma_{B(\tau)}(z)\,dz$$

as forms in $H^0(J(X)(\mathbf{C}_\tau),\Omega^d)$, for some $c_\tau \in F^\times$ (since both are generators of a rank one F-vector space). Then we have

$$\int_{A(\mathbf{R}_\tau)} |\omega_A|_\tau \sim \prod_\sigma \Omega^+_{f^\sigma_{B(\tau)}} \text{ in } \mathbf{C}^\times/F^\times,$$

where $\Omega^+_{f^\sigma_{B(\tau)}}$ is $2\pi i$ times the discriminant of the $f^\sigma_{B(\tau)}$-part of the analogue of the pairing (9.1.1) on $X(\mathbf{C}_\tau)$. When choices are made covariantly in τ, we then get $\Omega_A \sim \prod_{\sigma,\tau} \Omega^+_{f^\sigma_{B(\tau)}}$ in $\mathbf{C}^\times/\mathbf{Q}^\times$.

Our conjecture, decomposed into its (σ,τ)-constituents, can then be rewritten as

$$\Omega^+_f \sim \prod_\tau \Omega^+_{f_{B(\tau)}} \text{ in } \mathbf{C}^\times/(MF)^\times. \tag{9.1.2}$$

In this form, this is a stronger version of Shimura's conjecture [37] on the factorisation of periods of Hilbert modular forms up to algebraic factors in terms of P-invariants. The reader is referred to [41] for a discussion of this point.

Notice that (9.1.2) is nontrivial even when $F = \mathbf{Q}$: it asserts that the periods of the transfers of f to any indefinite quaternion algebra have the same transcendental (or irrational) parts. However, in this case the conjecture is known by the work of Shimura [36] (for the algebraicity) and Prasanna [31] (for the rationality).

For general F, Shimura's conjecture on P-invariants is largely proved by Yoshida [42] under an assumption of non-vanishing of certain L-values.

Remark 9.1.2. — It is clear that our conjecture implies that the Birch and Swinnerton-Dyer conjectural formula is true up to a nonzero rational factor when A has analytic M-rank zero. By the complex (respectively, the p-adic)

Gross–Zagier formula, the conjecture for f also implies the complex (respectively, the p-adic) Birch and Swinnerton-Dyer formulas up to a rational factor when A has (p-adic) analytic M-rank one.

9.2. Quadratic periods. — We can formulate a conjecture analogous to Conjecture 9.1.1 for the periods of the base-changed abelian variety $A_E = A \times_{\operatorname{Spec} F} \operatorname{Spec} E$.

Conjecture 9.2.1. — *We have*

$$\Omega_{A_E} \sim \prod_{\sigma} \Omega_{f^\sigma}$$

in $\mathbf{C}^\times/\mathbf{Q}^\times$.

Here the period of A_E is

$$\Omega_{A_E} = \prod_{\tau : E \to \mathbf{C}} \int_{A(\mathbf{C}_\tau)} |\omega_{A_E}|_\tau,$$

where for a differential form $\omega = h(z) dz_1 \wedge \cdots \wedge dz_k$ we have $|\omega|_\tau = |h(z)|_\tau^2 dz_1 \wedge d\overline{z}_1 \wedge \cdots \wedge dz_k \wedge d\overline{z}_k$.

As above, this conjecture can be "decomposed" into

$$\Omega_f \sim \prod_{\tau} \Omega_{f_{B(\tau)}} \text{ in } \mathbf{C}^\times/(MF)^\times. \tag{9.2.1}$$

where $\Omega_{f_{B(\tau)}}$ is π^2 times the Pertersson inner product of $f_{B(\tau)}$. This is essentially Shimura's conjecture on Q-invariants (see [37]). Up to algebraicity it has been proved by Harris [14] under a local condition (a new proof of the same result should appear in forthcoming work of Ichino–Prasanna, yielding rationality and removing the local assumption). Since (9.2.1) is implied

by (9.1.2) for f and f_ε, Harris's result can be seen as evidence for the conjecture on real periods.

We take the opportunity to record an immediate consequence of the conjecture on quadratic periods and the Gross–Zagier formulas.

Theorem 9.2.2. — *If A_E has analytic M-rank ≤ 1, then the complex and the p-adic Birch and Swinnerton-Dyer formulas for A_E are true up to a nonzero algebraic (or rational) factor.*

References

[1] Daniel Bertrand, *Propriétés arithmétiques de fonctions thêta à plusieurs variables*, Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, pp. 17–22, DOI 10.1007/BFb0099438 (French). MR756080 ↑12

[2] Amnon Besser, *The p-adic height pairings of Coleman-Gross and of Nekovář*, Number theory, CRM Proc. Lecture Notes, vol. 36, Amer. Math. Soc., Providence, RI, 2004, pp. 13–25. MR2076563 (2005f:11130) ↑67

[3] ———, *p-adic Arakelov theory*, J. Number Theory **111** (2005), no. 2, 318–371, DOI 10.1016/j.jnt.2004.11.010. MR2130113 (2006j:14029) ↑13, 61, 62, 64, 65, 67

[4] Daniel Bump, Solomon Friedberg, and Jeffrey Hoffstein, *Nonvanishing theorems for L-functions of modular forms and their derivatives*, Invent. Math. **102** (1990), no. 3, 543–618, DOI 10.1007/BF01233440. MR1074487 (92a:11058) ↑11

[5] Don Blasius, *On the critical values of Hecke L-series*, Ann. of Math. (2) **124** (1986), no. 1, 23–63, DOI 10.2307/1971386. MR847951 (88i:11035) ↑8, 89

[6] Daniel Bump, *Automorphic forms and representations*, Cambridge Studies in Advanced Mathematics, vol. 55, Cambridge University Press, Cambridge, 1997. MR1431508 (97k:11080) ↑30

[7] Colin J. Bushnell and Guy Henniart, *The local Langlands conjecture for* GL(2), Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 335, Springer-Verlag, Berlin, 2006. MR2234120 (2007m:22013) ↑31

[8] Henri Carayol, *Sur la mauvaise réduction des courbes de Shimura*, Compositio Math. **59** (1986), no. 2, 151–230 (French). MR860139 (88a:11058) ↑67, 68

[9] William Casselman, *On some results of Atkin and Lehner*, Math. Ann. **201** (1973), 301–314. MR0337789 (49 #2558) ↑24

[10] Robert F. Coleman and Benedict H. Gross, *p-adic heights on curves*, Algebraic number theory, Adv. Stud. Pure Math., vol. 17, Academic Press, Boston, MA, 1989, pp. 73–81. MR1097610 (92d:11057) ↑67

[11] Andrzej Dabrowski, *p-adic L-functions of Hilbert modular forms*, Ann. Inst. Fourier (Grenoble) **44** (1994), no. 4, 1025–1041 (English, with English and French summaries). MR1306548 (96b:11065) ↑3

[12] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), no. 3, 349–366, DOI 10.1007/BF01388432 (German). MR718935 (85g:11026a) ↑6, 99

[13] Benedict H. Gross and Don B. Zagier, *Heegner points and derivatives of L-series*, Invent. Math. **84** (1986), no. 2, 225–320, DOI 10.1007/BF01388809. MR833192 (87j:11057) ↑iii, 4, 10, 46, 104

[14] Michael Harris, *L-functions of* 2×2 *unitary groups and factorization of periods of Hilbert modular forms*, J. Amer. Math. Soc. **6** (1993), no. 3, 637–719, DOI 10.2307/2152780. MR1186960 (93m:11043) ↑91

[15] Haruzo Hida, *On p-adic L-functions of* GL(2) × GL(2) *over totally real fields*, Ann. Inst. Fourier (Grenoble) **41** (1991), no. 2, 311–391 (English, with French summary). MR1137290 (93b:11052) ↑3, 23

[16] H. Hida and J. Tilouine, *Anti-cyclotomic Katz p-adic L-functions and congruence modules*, Ann. Sci. École Norm. Sup. (4) **26** (1993), no. 2, 189–259. MR1209708 (93m:11044) ↑38

[17] Ming-Lun Hsieh, *Eisenstein congruence on unitary groups and Iwasawa main conjecture for CM fields*, preprint available at `http://www.math.ntu.edu.tw/mlhsieh/research.html`. ↑12

[18] Adrian Iovita, *Formal sections and de Rham cohomology of semistable abelian varieties*. part B, Proceedings of the Conference on p-adic Aspects of the Theory of Automorphic Representations (Jerusalem, 1998), 2000, pp. 429–447, DOI 10.1007/BF02834846. MR1809629 (2002g:14026) ↑58

[19] Hervé Jacquet, *Automorphic forms on* GL(2). *Part II*, Lecture Notes in Mathematics, Vol. 278, Springer-Verlag, Berlin, 1972. MR0562503 (58 #27778) ↑48

[20] Nicholas M. Katz, *p-adic properties of modular schemes and modular forms*, Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Springer, Berlin, 1973, pp. 69–190. Lecture Notes in Mathematics, Vol. 350. MR0447119 (56 #5434) ↑20

[21] ———, *The Eisenstein measure and p-adic interpolation*, Amer. J. Math. **99** (1977), no. 2, 238–311. MR0485797 (58 #5602) ↑47

[22] Shinichi Kobayashi, *The p-adic Gross-Zagier formula for elliptic curves at supersingular primes*, Invent. Math. **191** (2013), no. 3, 527–629, DOI 10.1007/s00222-012-0400-9. MR3020170 ↑86

[23] V. A. Kolyvagin, *Finiteness of $E(\mathbf{Q})$ and $SH(E,\mathbf{Q})$ for a subclass of Weil curves*, Izv. Akad. Nauk SSSR Ser. Mat. **52** (1988), no. 3, 522–540, 670–671 (Russian); English transl., Math. USSR-Izv. **32** (1989), no. 3, 523–541. MR954295 (89m:11056) ↑10

[24] V. A. Kolyvagin and D. Yu. Logachëv, *Finiteness of SH over totally real fields*, Izv. Akad. Nauk SSSR Ser. Mat. **55** (1991), no. 4, 851–876 (Russian); English transl., Math. USSR-Izv. **39** (1992), no. 1, 829–853. MR1137589 (93d:11063) ↑10

[25] Yu. I. Manin, *Non-Archimedean integration and p-adic Jacquet-Langlands L-functions*, Uspehi Mat. Nauk **31** (1976), no. 1(187), 5–54 (Russian). MR0417134 (54 #5194) ↑3

[26] Jan Nekovář, *On p-adic height pairings*, Séminaire de Théorie des Nombres, Paris, 1990–91, Progr. Math., vol. 108, Birkhäuser Boston, Boston, MA, 1993, pp. 127–202. MR1263527 (95j:11050) ↑57, 58

[27] ———, *On the p-adic height of Heegner cycles*, Math. Ann. **302** (1995), no. 4, 609–686, DOI 10.1007/BF01444511. MR1343644 (96f:11073) ↑51, 60

[28] A. A. Panchishkin, *Convolutions of Hilbert modular forms and their non-Archimedean analogues*, Mat. Sb. (N.S.) **136(178)** (1988), no. 4, 574–587, 592 (Russian); English transl., Math. USSR-Sb. **64** (1989), no. 2, 571–584. MR965894 (89k:11033) ↑3

[29] Bernadette Perrin-Riou, *Fonctions L p-adiques associées à une forme modulaire et à un corps quadratique imaginaire*, J. London Math. Soc. (2) **38** (1988), no. 1, 1–32, DOI 10.1112/jlms/s2-38.1.1 (French). MR949078 (89m:11043) ↑29, 49, 50, 51

[30] ———, *Points de Heegner et dérivées de fonctions L p-adiques*, Invent. Math. **89** (1987), no. 3, 455–510, DOI 10.1007/BF01388982 (French). MR903381 (89d:11034) ↑iii, 1, 4, 9, 10, 11, 12, 22, 47, 69, 77, 79, 80, 83, 85, 106

[31] Kartik Prasanna, *Arithmetic properties of the Shimura-Shintani-Waldspurger correspondence*, Invent. Math. **176** (2009), no. 3, 521–600, DOI 10.1007/s00222-008-0169-z. With an appendix by Brian Conrad. MR2501296 (2011d:11102) ↑90

[32] A. Raghuram and Naomi Tanabe, *Notes on the arithmetic of Hilbert modular forms*, J. Ramanujan Math. Soc. **26** (2011), no. 3, 261–319. MR2865819 (2012m:11060) ↑88

[33] Peter Schneider, *p-adic height pairings. II*, Invent. Math. **79** (1985), no. 2, 329–374, DOI 10.1007/BF01388978. MR778132 (86j:11063) ↑10

[34] Goro Shimura, *On the real points of an arithmetic quotient of a bounded symmetric domain*, Math. Ann. **215** (1975), 135–164. MR0572971 (58 #27992) ↑84

[35] ———, *The special values of the zeta functions associated with Hilbert modular forms*, Duke Math. J. **45** (1978), no. 3, 637–679. MR507462 (80a:10043) ↑2, 8

[36] ———, *The periods of certain automorphic forms of arithmetic type*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **28** (1981), no. 3, 605–632 (1982). MR656039 (84f:10040) ↑8, 90

[37] ———, *Algebraic relations between critical values of zeta functions and inner products*, Amer. J. Math. **105** (1983), no. 1, 253–285, DOI 10.2307/2374388. MR692113 (84j:10038) ↑90, 91

[38] Christopher Skinner and Eric Urban, *The Iwasawa main conjectures for GL(2)*. preprint. ↑12

[39] Jeanine Van Order, *On the Dihedral Main Conjectures of Iwasawa Theory for Hilbert Modular Eigenforms*, Canad. J. Math., to appear. ↑12

[40] J.-L. Waldspurger, *Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie*, Compositio Math. **54** (1985), no. 2, 173–242 (French). MR783511 (87g:11061b) ↑11, 30

[41] Hiroyuki Yoshida, *On the zeta functions of Shimura varieties and periods of Hilbert modular forms*, Duke Math. J. **75** (1994), no. 1, 121–191, DOI 10.1215/S0012-7094-94-07505-4. MR1284818 (95d:11059) ↑89, 90

[42] ———, *On a conjecture of Shimura concerning periods of Hilbert modular forms*, Amer. J. Math. **117** (1995), no. 4, 1019–1038, DOI 10.2307/2374957. MR1342839 (96d:11056) ↑90

[43] Xinyi Yuan, Shou-Wu Zhang, and Wei Zhang, *The Gross-Zagier Formula on Shimura Curves*, Annals of Mathematics Studies, vol. 184, Princeton University Press, Princeton, NJ, 2012. ↑7, 30, 35

[44] Yuri G. Zarhin, *p-adic heights on abelian varieties*, Séminaire de Théorie des Nombres, Paris 1987–88, Progr. Math., vol. 81, Birkhäuser Boston, Boston, MA, 1990, pp. 317–341. MR1042777 (91f:11043) ↑57

[45] Shou-Wu Zhang, *Heights of Heegner points on Shimura curves*, Ann. of Math. (2) **153** (2001), no. 1, 27–147, DOI 10.2307/2661372. MR1826411 (2002g:11081) ↑iii, 1, 5, 7, 10, 13, 26, 27, 31, 41, 43, 47, 48, 50, 67, 68, 69, 72, 73, 75, 78, 80

[46] ———, *Gross-Zagier formula for* GL_2, Asian J. Math. **5** (2001), no. 2, 183–290. MR1868935 (2003k:11101) ↑iii, 1, 10, 67, 75, 78

[47] ———, *Gross-Zagier formula for* GL(2). *II*, Heegner points and Rankin *L*-series, Math. Sci. Res. Inst. Publ., vol. 49, Cambridge Univ. Press, Cambridge, 2004, pp. 191–214, DOI 10.1017/CBO9780511756375.008. MR2083213 (2005k:11121) ↑iii, 1, 10

AFTERWORD (FOR THE LAYMAN)

0. — This concluding part is written for all the friends, relatives, and the occasional strangers met at parties who have been wondering what, exactly, I have been doing all these years. The most frequent questions were: is your thesis going to be just numbers and formulas, or also words? Did you prove some new theorem or formula? What is your thesis about? Is it useful for something?

The first question needs no answer at this point. The second question has a quick and dirty answer – I proved this formula (that's Theorem B in the Introduction):

$$L'_{p,\mathscr{W}}(f_E,\mathbf{1}) = D_F^{-2}\prod_{\wp|p}\left(1-\frac{1}{\alpha_\wp}\right)^2\left(1-\frac{1}{\varepsilon(\wp)\alpha_\wp}\right)^2\langle z_f, z_f\rangle_{\mathscr{W}}.$$

This, of course, means nothing without context, which brings me to the (longer) answer to the third question.[42] Before going there, however, let me offer two words about whether this is *new*: yes, and no. No, that is, because this work is a generalisation of the results of another mathematician. Yes, in the sense that in this situation those results were not known before (were they, no doctorate would follow), and although perhaps it is not surprising (to the experts) that they are true, proving them so has required some "*new*" ideas – which were in turn inspired from other people's work... But this is not an essay in the philosophy of history of scientific ideas.

[42]The more practically-minded reader should jump to the end for an answer to the last question.

1. — Esoteric as these pages may look like, the questions they deal with are much older than all the modern-day musings about gravity, or cells, or molecules. In fact, the proper name for this subject, Arithmetic, usually elicits reactions better-known to contemporary artists ("Any child can do that" – arithmetic is identified with primary school mathematics), so that we are stuck with the (uglyish) name of "theory of numbers". The numbers in questions are the integers $0, 1, 2, 3, \ldots$ and their negatives, as well as the fractions (such as $3/4, -7/5, \ldots$) – which go under the name of *rational* numbers. The founding father is usually recognised in Diophantus of Alexandria (3rd century A.D.), who in a treatise named, indeed, *Arithmetica* listed many problems and solutions to equations to be solved in integers or rationals. These questions already had a long history by then, starting with the Pythagorean discovery of the irrationality of the square root of 2 – that is to say, the non-existence of rational solutions to the quadratic equation $x^2 = 2$.

The general theory of quadratic equations of one variable is known to most high schoolers, and reduced to the Pythagorean question of the squareness of the discriminant[(43)] in the system of numbers of interest. Similarly, the study of equations of degree three or four[(44)] in one variable was reduced to the extraction of third and fourth roots by Italian Renaissance mathematicians. That the same cannot be said in general for equations of degree five or more was the early-Nineteenth-century discovery of Abel and independently Galois, who developed a complete theory of the symmetry of those

[(43)]For those struggling to recollect old memories, the discriminant of $ax^2 + bx + c$ is the (in)famous $\Delta = b^2 - 4ac$.
[(44)]Like $x^3 + x + 1 = 0$, or $2x^4 - 5x^2 + 7 = 0$.

equations. This was the end of a beautiful story – and the beginning of a new one, but we will not go there.

2. — Many of the equations studied by Diophantus are rather those in *two* variables, like $x^2+y^2=1$ (more generally one could consider systems of several equations in several variables, but two variables are still enough to give us headaches after almost two millennia). The rational solutions to this equation, like $x=3/5$, $y=4/5$, correspond to Pythagorean triples – triples of whole numbers, such as $(3,4,5)$, which can be the sides of a right-angled triangle[45]. Of course $x^2+y^2=1$ is also the equation for a circle: more precisely, this means that the solutions of this equation in the system of the *real* numbers (that is, the infinite decimals like 0.7163538902...) form a circle in the xy-plane. That the same can be said of the real solutions of $x^2+y^2=3$ should convince the reader that solving equations in the system of rational numbers is considerably more complicated than in the system of real numbers, once he or she is told that $x^2+y^2=3$ has no rational solutions at all![46]

[45]Because they satisfy the requirements of Pythagorean Theorem on the sum of the squares of the legs being the square of the hypothenuse: $3^2+4^2=5^2$. There are infinitely many Pythagorean triples, parametrized by $(m^2-n^2, 2mn, m^2+n^2)$ for any integers $m>n$.

[46]This fact is by no means obvious, although not difficult to show: writing $x=a/c$, $y=b/c$ with a common denominator c, the problem is equivalent to that of solving

$$a^2+b^2=3c^2$$

in integers a,b,c with c not equal to zero. If there is a solution, then the solution with the smallest possible positive c can't have a,b,c all even, since otherwise the halves of a, b and c would give a smaller solution. Now it is easy to see that the square of an even number $(2k)^2=4k^2$ leaves remainder 0 when divided by 4, while the square of an odd number $(2k+1)^2=4(k^2+k)+1$ leaves remainder 1. So looking at the remainders of division by four on either side of our equation, and denoting "$\equiv$" the relation of equality up to the addition of a multiple of 4, we should have $1+0\equiv 3$ or $0+1\equiv 3$ if only one of a or b is odd (so that c is too); or $1+1\equiv 0$ if both a and b are odd (so that c is even). This is clearly not the case, so there is no solution.

We have thus met the idea of attacking the difficult problem of *studying the solutions to an equation in the system of rational numbers* by *first studying them in simpler systems of numbers*, such as the real numbers and the systems of numbers "up to the addition of multiples of a fixed integer", such as the system of numbers up to addition of multiples of 4 used in the last footnote (these are called simply numbers "modulo 4" – note that this is a *finite* number system, containing only the four elements $0, 1, 2, 3$: indeed we have $4 \equiv 0$, and for example $2 + 3 \equiv 4 + 1 \equiv 1$). This idea turns out to be very fruitful, and in fact it suffices to give a complete treatment of equations of degree 2 in two variables, such as $x^2 + 3xy - 7y^2 + 5x - 4 = 0$: there is a solution exactly when there are real solutions[(47)] *and* solutions in the numbers "modulo N" for any integer N; this can be verified, by a human or a computer, in a finite (and quite short) amount of time, and if there is a solution then there are infinitely many.

3. — Finally, we can approach the topic of this thesis. For general equations of degree three or more, the existence of solutions in the real and finite number systems is not enough to guarantee the existence of a solution in the system of rational numbers. Even if we know that there is a solution, in general we still don't know whether there are finitely many or infinitely many others. Actually, for equations of degree at least five, one of the most important recent results in the subject, a theorem of the German mathematician Faltings [12], says that there is always at most a finite number of solutions. (On the other hand, the study of equations of degree four can be reduced to the case of degree three.) So in a sense the most important equations to

[(47)]This may not be the case: consider for example $x^2 + y^2 = -1$. When there are real solutions, they form a conic in the xy-plane: an ellipse, a parabola or a hyperbola.

be studied are those of degree three, called cubic equations, like for example $x^3 + y^3 = 1729$.[48] Here are the main questions: is there an efficient way (an algorithm) to detect whether a given cubic equation has a rational solution or not? If there is a solution, what can be said about how many solutions there are?

The answer to the first question is not known, and this thesis adds (almost) nothing to that. For the second question, it was known to Diophantus that given two solutions one could construct a third one by a geometric method: picturing the two given solutions as points P and Q on the cubic curve $\mathscr{C}$ in the xy-plane corresponding to the equation, the third one is constructed by intersecting the line through P and Q with $\mathscr{C}$ – since the degree is three, there will be three intersections, that is P, Q and the new solution. One can perform the same construction starting from just one point P and using the tangent line to that point (this corresponds to the 'degenerate' case $P = Q$); and then iterate the procedure. Then two things can happen: either one returns to P after a certain number of iterations (in this case P is called a *torsion* point); or one keeps getting new points (i.e. solutions) indefinitely. A fundamental 1922 theorem of Mordell says the following: all the rational solutions to the cubic equation of interest can be *generated from a finite number of points* by performing the geometric construction just described. The smallest possible such number of generating points (excluding the torsion points) is called the *rank* of the cubic: it is 0 or a positive integer, and it is 0

[48] I am choosing this equation because of its curious history: two famous mathematicians, Hardy and Ramanujan, were once meeting at Ramanujan's house. Hardy said upon arriving that his taxicab had a rather unremarkable number, 1729. Ramanujan immediately replied that the number was remarkable indeed, for being the smallest number expressible in two ways as the sum of two cubes: $1729 = 1000 + 729 = 10^3 + 9^3$, and $1729 = 1 + 1728 = 1^3 + 12^3$.

precisely when the number of solutions is finite; otherwise, the rank gives as a basic measure of "how infinite" the set of solutions is.

4. — What does the number of solutions in other number systems tell us about the rank of a cubic equation? After extensive computer simulations (a new thing at the time), the mathematicians Birch and Swinnerton-Dyer found in the 1960s a conjectural answer. Consider for example the equation $x^3 + y^3 = 1729$; we can look at its solutions in the systems of numbers modulo p for various prime numbers p (we don't need to restrict to prime numbers, but they are sufficient, and easier to deal with since one can define the operation of division on the associated number systems); for example, in the system of numbers modulo 5 we have $2^3+1^3 \equiv 9 \equiv 5+4 \equiv 4$, and $1729 \equiv 1725+4 \equiv 4$ so that $x = 2$, y$=$ 1 is a solution in this system. How many solutions can we expect, in general, in the system of numbers modulo p? We are looking at one equation in two variables, so roughly speaking we expect to have one free variable (for example, if every number modulo p had a unique cube root, then we could take x as the free variable and then $y = \sqrt[3]{x^3 - 1729}$) – since the free variables can assume the p values $0, 1, \ldots, p-1$ we then expect about p solutions.

The actual number of solution will vary around p, and here is the idea: if this number is often larger than p, then this should be because of the existence of *rational solutions* which "reduce" to solutions in the system of numbers modulo p.

How does one make this idea precise? If N_p is the number of solutions modulo p to a fixed cubic, a good measure of the discrepancy with the expected number of points is its ratio to p, that is N_p/p: if there are no rational solution (or only finitely many of them) this should always be around 1, while if there are infinitely many then it should often be larger. This information can be packaged into a function of a variable x (a real number),

$$L(x) = \frac{N_2}{2} \cdot \frac{N_3}{3} \cdot \frac{N_5}{5} \cdots \frac{N_{p_x}}{p_x}$$

where p_x is the largest prime number smaller than x. Again, if there are only finitely many rational solutions, each factor should be about 1 so that we expect that $L(x)$ does not grow when x grows. On the other hand if there are infinitely many solutions, Birch and Swinnerton-Dyer observed in their examples that $L(x)$ grew like

$$L(x) \sim C \log(x)^r,$$

for some real constant C and some non-negative integer r; and that r was equal to the *rank* of the curve.[49] They conjectured that the above relation between the growth rate of $L(x)$ and the rank should hold for all cubic equations, and moreover they predicted what the value of C should be (in terms of various quantities associated to this equation).

Fifty years later, this conjecture is still unproven and likely to remain so for a long time – notwithstanding the prize of one million dollars offered by

[49]To get a sense at why this conjecture could not have been made before the advent of computers, notice that $\log(x)$, the inverse function of the exponential function, grows very slowlly. So if in the study of the growth of $L(x)$ one wants to see when it surpasses, say, the value of 9, and if $r = 1$, then one needs to find the number of solutions to his equation modulo primes up to $10^9 = 1,000,000,000$.

a foundation who ranked it among the seven most important unsolved[(50)] mathematical problems. This thesis, as part of a wider circle of ideas developed in the past three decades, makes some progress towards it.

5. — The readers who have followed me to this point now know what we are talking about – yet they may (and should) complain that after several pages they still have received no explanation as to how any of the words in the title of this work relates to the problems I described. In fact, at least two words make sense: I have tried to explain above how the problem of finding rational solutions to an equation can be viewed as that of finding points with rational coordinates on a curve. I will try to explain the other words.

Shimura curves. — Shimura is a Japanese mathematician who studied a certain class of curves, called indeed Shimura curves, which *parametrise* certain other geometric objects (let us call those "Shimura objects"). An example should illustrate what this means: a circle in the plane can be identified by three numbers, the two coordinates of its centre (two real numbers) and the length of its radius (a positive real number); if we consider circles to be equivalent when they can be rigidly moved to coincide, than the radius is the only parameter: that is, the positive real numbers *parametrise* the circles up to equivalence.

Shimura curves are less concrete than other curves given by explicit equations, yet they have a crucial advantage: one can find rational points on them in a natural way, since they correspond to the "rational Shimura objects" – in the above example, circles with rational radius can be thought

(50) One of the seven has been solved since being included in the list – but the million dollars offered for it was refused by the winner.

of as "rational circles" (of course, in the example it is trivial to find rational points on the – rather straight – "curve" constituted by the positive real numbers, but by now my reader will be convinced that for other curves this is not so). Shimura and others made the striking conjecture that *every cubic curve is related to some Shimua curve*, in the precise sense that one can find a transformation (a function) from a Shimura curve to the cubic curve, which sends rational points to rational points. (Here is an example to illustrate the concept: the parabola $y = x^2$ is related to the real line with coordinate t by the function which sends t to the point $(x = t, y = t^2)$ on the parabola – and if t is rational then so are x and y.) The proof of this conjecture by Andrew Wiles in 1993 was the essential ingredient towards his proof of the famous Fermat's Last Theorem, which had been waiting for one for four centuries.

Heegner points. — Heegner was a German high school teacher and amateur mathematician, who in 1952 had the brilliant idea of producing rational points on cubic curves from the natural rational points on the related Shimura curve. This is to date the only systematic way of finding solution to cubic equations.[51] One can raise the question of whether the so-called Heegner points thus obtained are torsion points or not, and here is the answer:[52] suppose that $L(x)$ grows like $C \log(x)^r$; then the point is torsion

[51]Here is an example of its power – relegated to a footnote in order to avoid scaring any readers with the size of the numbers involved. The equation $1063y^2 = x^3 - x$ has rank one, and the simplest non-torsion solution has the x-coordinate $q^2/1063$, where

$$q = \frac{1109186374182976967504702163571228176738233966743464 5}{3173426575447721807352079773209000125228079367778 87}.$$

(This was found as a Heegner point by Noam Elkies. The reader may make a guess on how long it would take to find it with a naive search.)

[52]Due to Gross and Zagier [13]: my more attentive followers may recognize in the latter the name of my *laurea* thesis advisor.

exactly when $r = 1$ (as an exercise, the reader can try to extract what this implies for the conjecture of Birch and Swinnerton-Dyer). Notice that, since r is supposed to equal the rank, it is no surprise that the Heegner point is torsion when $r = 0$, as in this case *every* solution is torsion. On the other hand, when r is at least two, we should expect many (non-torsion) rational points, yet the only known method for finding them fails, and we are at a loss.

p-adic heights. — A measure of the complexity of a rational number written as a reduced fraction m/n is the maximum of the number of digits of m and n. One can refine this notion to study the complexity of a point on a cubic curve – and the resulting measure of complexity of the point is called the *height* of the point. It is a non-negative real number, which is zero precisely when the point is torsion. A study of the height of Heegner points is what has led to the result mentioned just above on the "growth versus rank" part of the Birch and Swinnerton-Dyer conjecture.

The *p-adic height* is another measure of the complexity of a point, which is not a real number but rather a *p-adic number*, that is an element of a certain infinite system of numbers obtained by combining together the systems of numbers modulo $p, p^2, p^3, \ldots$ for a fixed prime number p. Now reading the title should give some ideas as to what this thesis studies – which is useful for understanding the constant C in the growth of $L(x)$. (This constant is important: it is related to the problem of deciding on the very existence of any points on some *other* related cubic curves.)

6. — Finally, a word of explanation on the formula I started this section with (and which is indeed my main result): the Heegner point being studied is z_f (here f is the name of the corresponding cubic curve on which it lies), and $\langle z_f, z_f \rangle_{\mathscr{W}}$ is its p-adic height; on the left-hand side, the quantity $L'_{p,\mathscr{W}}(f_E, \mathbf{1})$ is related to the function $L(x)$ introduced above.

As mentioned before, the results of this work are not new in the study of rational solutions to cubic equations.[53] The novelty here is that they are proved for a certain class of systems of equations in any number of variables which share some of the features of cubic equations of one variable (some of these systems are related to the study of one-variable equations of degree higher than three), and over systems of numbers which are more general than the system of rational numbers.

7. — I have thus far escaped the question of what all this is useful for. Jacobi's noble answer *"pour l'honneur de l'esprit humain"* may not satisfy every palate in an era of tight budgets. It is certainly unclear that the world is a better place after these pages. Yet the same could be said of the vast majority of basic research.[54] *Si parva licet*, nothing could have sounded more abstruse than antimatter when Dirac suggested its existence in 1928; yet today it is used daily in PET[55] scans in every modern hospital. More to the point, no one from Diophantus to Heegner could foresee the advent of the internet; yet the cryptographic methods that protect our electronic transactions and communications are using all the arithmetic knowledge developed in the

[53] They are due to the French mathematician Bernadette Perrin-Riou [30].
[54] Not to mention many other more marketable cultural artifacts: opinion pieces, TV ads, mortgage-backed securities…
[55] Positron Emission Tomography (positrons are the antiparticles of electrons).

past two millennia – and cubic curves have a significant part in it. Whether the ideas of this work will prove useful for improving on those methods, for nothing at all, or for something else which has not been invented yet, is as unclear as it is unpredictable: we will have to wait and see.

DANIEL DISEGNI • *E-mail :* disegni@math.columbia.edu