Elliptic curves and p-adic uniformisation

H. Darmon*

September 9, 2007

Elliptic curves. An *elliptic curve* is a curve of genus one with a distinguished rational point. It can be described by a homogeneous equation of the form

$$E : Y^2Z = 4X^3 + aXZ^2 + bZ^3, \tag{1}$$

where the parameters $a, b \in \mathbf{Z}$ satisfy the condition

$$\Delta := -2^{12}(a^3 + 27b^2) \neq 0.$$

The Diophantine theory studies the rational solutions $(X, Y, Z) \in \mathbf{Q}^3$ of equation (1). It is convenient to ignore the trivial solution $(0, 0, 0)$ and to identify solutions if they differ by multiplication by a non-zero scalar. Solutions to (1) are thus viewed as points in the *projective plane* $\mathbf{P}_2(\mathbf{Q})$. Let $E(\mathbf{Q}) \subset \mathbf{P}_2(\mathbf{Q})$ denote this solution set. More generally, if F is any field, let $E(F) \subset \mathbf{P}_2(F)$ be the corresponding set of solutions with values in F. It is identified with the set of $(x, y) \in F^2$ satisfying the associated affine equation

$$y^2 = 4x^3 + ax + b, \tag{2}$$

together with the "point at infinity" corresponding to $(X, Y, Z) = (0, 1, 0)$.

Among all the *projective*[1] curves over $\mathbf{Q}$, the elliptic ones are worthy of special consideration, because they alone are *algebraic groups*: the set $E(\mathbf{Q})$

*This is a transcription of the author's Coxeter-James lecture given at the CMS Winter meeting in Kingston in December 1998. It is a pleasure to thank Massimo Bertolini and Adrian Iovita for many fruitful exchanges over the years, and the Canadian Mathematical Society for its invitation to deliver the Coxeter-James lecture.

[1]I.e., defined by a system of homogeneous equations.

is equipped with a binary composition law

$$E(\mathbf{Q}) \times E(\mathbf{Q}) \longrightarrow E(\mathbf{Q})$$

defined by a system of polynomials with rational coefficients, making $E(\mathbf{Q})$ into a commutative group, with identity element the distinguished point at infinity. The same set of polynomials endows $E(F)$ with a natural addition law[2], admitting a simple geometric description in terms of the *chord and tangent method*: viewing points in $E(F)$ as points on the affine plane by equation (2), one simply sets $P + Q + R = 0$ whenever P, Q, and R lie on the same line. (See for example [ST], ch. I.) The Diophantine study of E is facilitated and enriched by the presence of this extra structure.

The group $E(\mathbf{C})$ is isomorphic to the quotient of $\mathbf{C}$ by a lattice Λ. For a suitable Λ, the inverse isomorphism sends $z \in \mathbf{C}$ to $(\wp(z), \wp'(z)) \in E(\mathbf{C})$, where

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda - 0} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right)$$

is the *Weierstrass $\wp$-function* attached to Λ. The group law on $E(\mathbf{C})$ corresponds to the usual addition law of complex numbers on $\mathbf{C}/\Lambda$. This explicit analytic description yields the structure of $E(\mathbf{C})$ and $E(\mathbf{R})$: the former is a product of two circles, and the latter is either a circle or the product of a group of order 2 with a circle.

The structure of $E(\mathbf{Q})$ lies deeper. In the case of the elliptic curve

$$E : y^2 = x^3 + 877x, \tag{3}$$

Bremner and Cassels [BC] showed that $E(\mathbf{Q})$ is generated by the point $(0, 0)$ of order 2 and the point of *infinite* order

$$(x,y) = \left(\left(\frac{612776083187947368101}{78841535860683900210} \right)^2, \frac{25625626798892680938877683404551308964866915320435660346478694 9}{788415358606839002103} \right).$$

For the elliptic curve[3]

$$y^2 + xy + y = x^3 - 20333x + 203852, \tag{4}$$

[2]provided that the equation (1) remains non-singular over F, which is the case for example if the characteristic of F does not divide Δ.

[3]its equation is not given in Weierstrass form as in equation (1), but can be brought to this form by a simple change of variables.

it turns out that $E(\mathbf{Q})$ is generated by the six points $P_j = (x_j, y_j)$ with

$$\begin{array}{lll} P_1 = (-51, 1078), & P_2 = (3, 376), & P_3 = (165, 1078), \\ P_4 = (-24, 835), & P_5 = (-132, 835), & P_6 = (136, 106). \end{array}$$

In fact, a point in $E(\mathbf{Q})$ can be written *uniquely* as $n_1P_1 + \cdots + n_6P_6$, with $n_j \in \mathbf{Z}$.

In general, how are the rational solutions to equation (2) calculated? As with many fundamental questions in number theory, the first progress dates back to Fermat, who introduced his famous method of *infinite descent* and used it to show that certain elliptic curves, related to the Fermat equation with exponent 4 and 3, have finitely many solutions. Fermat's descent was later adapted by Mordell[4] to prove the following general result about $E(\mathbf{Q})$, which is suggested by the special cases (3) and (4).

Theorem *The group $E(\mathbf{Q})$ is a finitely generated abelian group, i.e.,*

$$E(\mathbf{Q}) \simeq T \oplus \mathbf{Z}^r,$$

where T is a finite group (identified with the torsion subgroup *of $E(\mathbf{Q})$).*

The integer r is called the *rank* of $E(\mathbf{Q})$: it represents the minimal number of solutions needed to generate a finite index subgroup of $E(\mathbf{Q})$ by repeated application of the chord and tangent law. The rank depends in a subtle way on E, and can get quite large[5].

Unfortunately, the proof of Mordell's theorem, based on Fermat's descent, is not *effective*; it is not known whether Fermat's descent procedure always terminates eventually. The following basic question remains open.

Question: *Is there an* algorithm *to compute $E(\mathbf{Q})$?*

[4]The proof was then further generalized by Weil to cover *abelian varieties* over *number fields*. An abelian variety is a projective (commutative) algebraic group; it is a natural *higher dimensional* generalization of elliptic curves. For this reason Mordell's theorem is often referred to as the Mordell-Weil theorem, and $E(\mathbf{Q})$ is called the *Mordell-Weil group* attached to E.

[5]It is expected that it can get arbitrarily large, although this is not proved. The record so far is an elliptic curve of rank ≥ 22 [Fe].

What is desired is a deterministic recipe which, given a and b in equation (1) (say) yields a description of $E(\mathbf{Q})$. The torsion subgroup T can be calculated without difficulty; the key challenge arises in computing the rank r and a system of generators for $E(\mathbf{Q})$.

The Birch and Swinnerton-Dyer conjecture. Further insights about $E(\mathbf{Q})$ may be gleaned by studying E over other fields, such as the finite field $\mathbf{F}_p$ with p elements consisting of the residue classes modulo a prime p. The set $E(\mathbf{F}_p)$ is finite. A simple heuristic argument suggests that its cardinality N_p is roughly $p+1$. Indeed Hasse proved that the "error term" $a_p := p + 1 - N_p$ satisfies

$$|a_p| \leq 2\sqrt{p}.$$

Reduction of solutions modulo p gives a natural map $E(\mathbf{Q}) \longrightarrow E(\mathbf{F}_p)$. One might expect the presence of a large supply of rational points in $E(\mathbf{Q})$ to have an impact on the size of $E(\mathbf{F}_p)$ on average. Compelled by this insight, Birch and Swinnerton-Dyer studied the asymptotic behaviour of $\prod_{p<X} N_p/p$ as X gets large. On the basis of numerical experiments, they were led to conjecture that

$$\prod_{p<X} N_p/p \simeq C_E(\log X)^r, \tag{5}$$

where C_E is a constant depending only on E. This striking conjecture asserts that the rank r - an *a priori* subtle *global* invariant of the arithmetic of E over $\mathbf{Q}$ - can be read off from the asymptotic behaviour of the N_p, reflecting information about E of a "local" nature.

Subsequently, following a suggestion of Davenport, Birch and Swinnerton-Dyer gave a more sophisticated formulation of the conjecture in terms of the *Hasse-Weil L-function* $L(E,s)$ attached to E. Let s be a complex variable, and for $p \not| \Delta$ let

$$L(E,p,s) := (1 - a_p p^{-s} + p^{1-2s})^{-1} \tag{6}$$

be the *local L-function* attached to E at p. There is also a simple definition of $L(E,p,s)$ for the finite set of primes dividing Δ, whose precise nature need not concern us here. (See for example [Si].) The L-function $L(E,s)$ of E over $\mathbf{Q}$ is then defined by the "Euler product"

$$L(E,s) := \prod_p L(E,p,s), \tag{7}$$

where the product is taken over all the primes. The Hasse bound $|a_p| \leq 2\sqrt{p}$ implies that it converges absolutely in the right half-plane $\mathrm{Re}(s) > 3/2$. In particular, the point $s = 1$ is outside the domain of absolute convergence. However, the fundamental Shimura-Taniyama conjecture, which will be discussed further below, implies that $L(E, s)$ has an *analytic continuation* to all of $\mathbf{C}$.

Noting the identity (for $p \not| \Delta$)

$$L(E, p, 1) = \frac{p}{N_p},$$

and comparing it with the quantity occuring in (5), Birch and Swinnerton–Dyer were led to conjecture that the rank r should be reflected in the order of vanishing of $L(E, s)$ at $s = 1$.

Conjecture BSD: *The function $L(E, s)$ satisfies*

$$\mathrm{ord}_{s=1} L(E, s) = r.$$

A more precise version of this conjecture expresses $L^{(r)}(E, 1)$, the rth derivative of $L(E, s)$ at $s = 1$, in terms of various quantities associated to E over $\mathbf{Q}$, most notably a "regulator term" measuring the arithmetic complexity of a system of generators for $E(\mathbf{Q})$, and the order of a *conjecturally finite* group known as the Shafarevich-Tate group of E, and denoted by the Cyrillic letter $\underline{III}$. The precise definition of this group would take the reader too far afield; suffice it to say that $\underline{III}(E)$ measures the *difficulty* of computing $E(\mathbf{Q})$ by Fermat's descent method. In particular, its finiteness implies that Fermat's descent terminates when applied to E. It is for this reason that the Shafarevich–Tate conjecture, which predicts the finiteness of $\underline{III}(E)$ for all E, is widely viewed as the most important outstanding question in the arithmetic of elliptic curves.

Concerning the Birch-Swinnerton Dyer conjecture, Tate wrote [Ta1]

> "This remarkable conjecture relates the behaviour of a function L, at a point where it is not at present known to be defined, to the order of a group $\underline{III}$, which is not known to be finite."

This quote accurately summarized the state of knowledge (or perhaps, ignorance) on the question, until around 1987, when the results of Gross-Zagier and Kolyvagin, and then Wiles, led to dramatic breakthroughs[6].

The Shimura-Taniyama conjecture and Wiles' theorem. The fact that *a priori* $L(E,s)$ is not even known to be *defined* at $s=1$ presents an obvious obstacle to tackling the Birch and Swinnerton-Dyer conjecture. In 1993, Wiles established the analytic continuation of $L(E,s)$ for a large class of elliptic curves by relating E (and its L-function) to modular forms.

Given an integer N, let $\Gamma_0(N)$ be the group of matrices in $\mathbf{SL}_2(\mathbf{Z})$ which are upper triangular modulo N. It acts as a discrete group of Mobius transformations on the Poincaré upper half-plane

$$\mathcal{H} := \{z \in \mathbf{C} | Im(z) > 0\}.$$

A *cusp form of weight 2 for* $\Gamma_0(N)$ is an analytic function f on $\mathcal{H}$ satisfying the relation

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 f(z), \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N), \tag{8}$$

together with suitable growth conditions on the boundary of $\mathcal{H}$. For example, the invariance in equation (8) implies that f is periodic of period 1, and one requires that it can be written as a power series in $q = e^{2\pi i z}$ with no constant term:

$$f(z) = \sum_{n=1}^{\infty} \lambda_n q^n.$$

The Dirichlet series

$$L(f,s) = \sum \lambda_n n^{-s}$$

is called the L-function attached to f. A direct calculation reveals that $L(f,s)$ is essentially the *Mellin transform* of f:

$$\Lambda(f,s) := \Gamma(s) L(f,s) = (2\pi)^s \int_0^{\infty} f(iy) y^{s-1} dy. \tag{9}$$

[6]Prior to this one should also mention the work of Coates and Wiles establishing partial results towards the Birch and Swinnerton-Dyer conjecture for elliptic curves with *complex multiplication* – a restricted class, but one which has played an important role in the development of the theory.

The space of cusp forms of weight 2 on $\Gamma_0(N)$ is a finite-dimensional vector space and is preserved by the involution W_N defined by

$$W_N(f)(z) = Nz^2 f(\frac{-1}{Nz}).$$

Hecke showed that if f lies in one of the two eigenspaces for this involution (with eigenvalue $w = \pm 1$) then $L(f, s)$ satisfies the functional equation:

$$\Lambda(f, s) = -w\Lambda(f, 2 - s). \tag{10}$$

In particular, $L(f, s)$ has an analytic continuation to all of $\mathbf{C}$.

The curve E is said to be *modular* if there exists a cusp form f of weight 2 on $\Gamma_0(N)$ for some N such that

$$L(E, s) = L(f, s).$$

Taniyama and Shimura conjectured in the fifties that every elliptic curve over $\mathbf{Q}$ is modular. This important conjecture gives a framework for proving the analytic continuation and functional equation for $L(E, s)$, and illustrates a deep relationship between objects arising in arithmetic, such as E, and objects, such as f, which are part of an ostensibly different circle of ideas – related to Fourier analysis on groups, and the (infinite-dimensional) representation theory of adelic groups, as described in the ambitious Langlands program.

The conjecture of Shimura-Taniyama, as refined by Weil, predicts that the integer N is equal to the so-called *arithmetic conductor* of E. This integer can be computed effectively in terms of an equation defining E, and is divisible only by the primes dividing Δ, but with different exponents in general. From now on, the letter N will be used to denote the conductor of E.

Thanks to the work of Wiles [Wi], Taylor-Wiles [TW] and its extensions [Di], [CDT], one now knows that E is modular, at least provided that E satisfies a mild technical restriction.

Theorem STW. *If* 27 *does not divide* N*, then* E *is modular.*

Complex uniformisation. The modularity of E can also be formulated as a statement about the complex uniformisation of the Riemann surface $E(\mathbf{C})$. (Cf. [Ma2]).

Theorem STW$_\infty$: *If* 27 *does not divide* N*, then there is a complex analytic uniformisation*

$$\varphi_\infty : \mathcal{H}/\Gamma_0(N) \longrightarrow E(\mathbf{C}).$$

The classical uniformisation theorem of complex analysis states that every Riemann surface is expressible as a quotient of $\mathcal{H}$ by the action of *some* discrete subgroup Γ of $\mathbf{SL}_2(\mathbf{R})$. The above statement lies deeper. Its arithmetic content comes from the fact that it makes a precise statement about the nature of Γ, and relates it to the arithmetic of E over $\mathbf{Q}$. Groups like $\Gamma_0(N)$ which are defined by simple congruence conditions on the matrix entries, are examples of what are called called *arithmetic subgroups* of $\mathbf{SL}_2(\mathbf{Z})$.

Evidence for the Birch–Swinnerton-Dyer conjecture. As Mazur writes in [Ma1],

> " It has been abundantly clear for years that one has a much more tenacious hold on the arithmetic of an elliptic curve $E/\mathbf{Q}$ if one supposes that it is [...] parametrized [by a modular curve]."

This sentiment is supported by the following result, following from the work of Kolyvagin [Ko] and earlier work of Gross and Zagier [GZ].

Theorem GZK. *Let* E *be an elliptic curve over* $\mathbf{Q}$ *of rank* r*. Suppose that* E *is modular, and that* $ord_{s=1}L(E,s) \leq 1$*. Then*

$$ord_{s=1}L(E,s) = r,$$

and the Shafarevich-Tate conjecture is true for E*.*

The theorem (or rather, its proof) even supplies a procedure for computing $E(\mathbf{Q})$, based on the theory of complex multiplication, which relies on the modularity of E and is more efficient that the descent method of Fermat.

Higher order zeroes. The following question remains as the ultimate challenge concerning the Birch and Swinnerton-Dyer conjecture.

Question: What if $\mathrm{ord}_{s=1}L(E,s) > 1$?

In this case the relation between r and the order of vanishing of $L(E,s)$ at $s=1$ remains mysterious. It appears that the inequality

$$\mathrm{ord}_{s=1}L(E,s) \geq r \tag{11}$$

ought to be easier to prove than the reverse inequality[7]. However, even this "easy half" of the Birch Swinnerton-Dyer conjecture seems out of reach for now. The process whereby the presence of "many" rational points in $E(\mathbf{Q})$ forces higher vanishing of $L(E,s)$ at $s=1$ is simply not understood.

There are elliptic curves E for which the sign $-w$ in the functional equation (10) is -1, and for which $r \geq 3$, such as the elliptic curve

$$y^2 = 4x^3 - 28x + 25 \tag{12}$$

of conductor $N = 5077$. In this case $L(E,s)$ vanishes to odd order, and theorem GZK implies that $L'(E,1) = 0$. Hence

$$\mathrm{ord}_{s=1}L(E,s) \geq 3. \tag{13}$$

But this is basically as far as one can go! Indeed the following question remains open:

Question: *Is there an elliptic curve E over $\mathbf{Q}$ with $ord_{s=1}L(E,s) > 3$?*

In his undergraduate summer project [Gh], Alexandru Ghitza evaluated the first few derivatives of $L(E,s)$ at $s=1$ for the curve of rank 6 given by equation (4). In this case $L(E,1) = 0$ and the sign $-w$ in the functional equation for $L(E,s)$ is 1, so that $L(E,s)$ vanishes to even order ≥ 2. Ghitza's numerical calculations (performed on a high-speed computer with an accuracy of around four significant digits after the decimal point) produced

$$\begin{aligned} L''(E,1) &\simeq -0.0000195, \\ L^{(4)}(E,1) &\simeq -0.00000027, \\ L^{(6)}(E,1) &\simeq 717.6663612. \end{aligned}$$

[7]For example, it is known to hold in the function field case, by work of Tate[Ta2]. The reverse inequality seems inextricably linked with questions related to the Shafarevich–Tate conjecture.

This strongly suggests that $L(E,s)$ vanishes to order 6 at $s=1$, as predicted by the Birch and Swinnerton-Dyer conjecture, but it appears to be an *extremely difficult* theoretical problem to prove that $L''(E,1)=0$, even for this specific curve!

Using the known elliptic curves with $r \geq 22$, note that a proof of (11) would imply the existence of L-functions for which $\mathrm{ord}_{s=1} L(E,s) \geq 22$.

The work of Goldfeld. Producing L-functions $L(E,s)$ with high order zeroes at $s=1$ has a number of applications. For example, Goldfeld [Go] showed that the existence of a suitable[8] elliptic curve E for which $ord_{s=1} L(E,s) \geq r$ implies the following asymptotic lower bounds on the growth of the class number $h(D)$ of the imaginary quadratic field of discriminant D:

$$h(D) \geq c(\log |D|)^{r-2-\epsilon}.$$

The importance of this estimate lies in the fact that the constant c is *effective*, and can be calculated in terms of the elliptic curve E. Goldfeld's work, using the elliptic curve of equation (12), led to the unconditional estimate

$$h(D) \geq \frac{1}{55}(\log |D|)^{1-\epsilon},$$

and thus to a solution of the celebrated class number problem of Gauss. Note the key role played in this estimate by equation (13), which is based in turn on theorem GZK.

More recently, Ram Murty has informed me that an analogue of a conjecture of Polya about the Riemann zeta-function $\zeta(s)$, which was subsequently shown to be false, ought to be true after replacing $\zeta(s)$ by L-functions $L(E,s)$ admitting a high-order zero at $s=1$. (Cf. [Mu].)

p-adic analysis. In the face of the difficulties associated with understanding the complex L-function, it has proved fruitful to replace the complex variable s by a p-adic one, and the classical Hasse-Weil L-function by a p-adic analogue.

In addition to the usual "archimedean" distance $d_\infty(x,y)=|x-y|$, the rational numbers are equipped (for each prime p) with the p-adic distance

[8] by "suitable" it is meant that the number of primes dividing N to odd order should be odd, if $w=1$, and even, if $w=-1$.

$d_p(x,y) = p^{-\mathrm{ord}_p(x-y)}$, according to which two rational numbers are declared to be close to each other if (the numerator of) their difference is divisible by a high power of p. Ostrowski's theorem asserts that the usual absolute value and the p-adic ones, as p ranges over the primes, give a complete list of metrics (up to a suitable equivalence) which are compatible with the field structure on $\mathbf{Q}$.

Just as $\mathbf{R}$ is the completion of $\mathbf{Q}$ with respect to the usual metric, the field $\mathbf{Q}_p$ is the completion of $\mathbf{Q}$ with respect to d_p. It has a greater arithmetic complexity that $\mathbf{R}$, in the sense that its algebraic closure $\bar{\mathbf{Q}}_p$ is of *infinite degree* over $\mathbf{Q}_p$, unlike $\mathbf{C}$ over $\mathbf{R}$. As a consequence, $\bar{\mathbf{Q}}_p$ is not a complete field[9]. The role of $\mathbf{C}$ is played by a larger field, denoted $\mathbf{C}_p$, the *completion* of $\bar{\mathbf{Q}}_p$ with respect to the p-adic metric.

The *p-adic upper half plane* $\mathcal{H}_p$ is defined as

$$\mathcal{H}_p := \mathbf{P}_1(\mathbf{C}_p) - \mathbf{P}_1(\mathbf{Q}_p) = \mathbf{C}_p - \mathbf{Q}_p.$$

Note that replacing $\mathbf{C}_p$ by $\mathbf{C}$, and $\mathbf{Q}_p$ by $\mathbf{R}$, yields two copies of the usual Poincaré upper half plane. In the p-adic setting, $\mathbf{C}_p - \mathbf{Q}_p$ does not split naturally into two disjoint connected pieces, so that it is more natural to work with $\mathcal{H}_p$ in its entirety.

The space $\mathcal{H}_p$ is endowed with a rich theory of "p-adic analytic functions" which mirrors the complex-analytic theory. By analogy with the complex case, it could be tempting to define an "analytic" function on $\mathcal{H}_p$ as a $\mathbf{C}_p$-valued function which admits a power series expansion in each open disk. In the p-adic setting, however, two open discs are either disjoint or one is contained in the other! The space of "analytic functions" according to this definition turns out to be too large and not "rigid" enough to yield a useful theory: for example, the principle of analytic continuation fails.

A fruitful function theory, obeying many of the principles of classical complex analysis, is obtained by replacing open discs by so-called *affinoid* sets, which are made up of a closed p-adic disc with a number of open disks deleted. The affinoids cover $\mathcal{H}_p$ and can be used to define a sheaf of *rigid analytic functions* which enjoys many of the same formal properties as the sheaf of complex analytic functions on $\mathcal{H}$.

[9]for example, if ζ_n is a primitive p^nth root of unity, then $\sum_{n=1}^{\infty} \zeta_n p^n$ does not have a limit in $\bar{\mathbf{Q}}_p$ even though its partial sums form a Cauchy sequence.

The group $\mathbf{SL}_2(\mathbf{Q}_p)$ acts on $\mathcal{H}_p$ by fractional linear transformations, just as $\mathbf{SL}_2(\mathbf{R})$ acts on $\mathcal{H}$. If Γ is a discrete subgroup of $\mathbf{SL}_2(\mathbf{Q}_p)$, the quotient $X_\Gamma := \mathcal{H}_p/\Gamma$ inherits a p-adic topology and becomes a *rigid analytic curve*: it is the analogue, in the p-adic realm, of a Riemann surface.

The p-adic uniformisation theory of Mumford addresses the question of which curves $X/\mathbf{C}_p$ can be be written as a quotient $\mathcal{H}_p/\Gamma$, for $\Gamma \subset \mathbf{SL}_2(\mathbf{Q}_p)$. Unlike the complex case, not every curve over $\mathbf{C}_p$ can be so uniformized. Mumford identifies a simple necessary and sufficient condition[10] for X to admit a p-adic uniformisation. Curves over $\mathbf{C}_p$ with this property are called *Mumford curves*. An elliptic curve over $\mathbf{Q}$ is a Mumford curve precisely when its conductor is exactly divisible by p. The p-adic uniformisation theory of elliptic curves with $p||N$ was developped by Tate, and later generalized by Mumford to curves of higher genus.

Rigid analytic Shimura-Taniyama. Let E be an elliptic curve over $\mathbf{Q}$ with $27 \not| N$, so that E is modular in the sense of theorem STW. The following result is a p-adic analogue of theorem STW_∞, and follows by combining the result of Wiles with earlier work of Eichler-Shimizu-Jacquet-Langlands, Shimura, and Cerednik-Drinfeld. (Cf. for example the work of Jordan-Livné [JL].)

Theorem $\mathbf{STW}_p$. *Suppose that $p||N$, so that E is Mumford curve over $\mathbf{Q}_p$. Then there exists a discrete* arithmetic *subgroup Γ of $\mathbf{SL}_2(\mathbf{Q}_p)$ and a rigid analytic uniformisation of $E(\mathbf{C}_p)$:*

$$\varphi_p : \mathcal{H}_p/\Gamma \longrightarrow E(\mathbf{C}_p).$$

The key word in this theorem is the word *arithmetic*. It means that the groups involved in the uniformisation are analogous to $\Gamma_0(N)$. The definition of these groups is somewhat more involved. Rather than provide a complete definition, here is an example which gives the flavour of the general case. Let

$$B := \mathbf{Q} + \mathbf{Q}i + \mathbf{Q}j + \mathbf{Q}k$$

[10] X should have a model over $\mathcal{O}$ (the ring of integers of $\mathbf{C}_p$) whose special fiber is a union of projective lines intersecting transversally at ordinary double points.

be the ring of Hamilton quaternions with coefficients in $\mathbf{Q}$, and let

$$R = \mathbf{Z}[i, j, k, \frac{1+i+j+k}{2}]$$

be Hurwitz's maximal order. If p is an odd prime, then $B \otimes \mathbf{Q}_p$ is isomorphic to the ring $M_2(\mathbf{Q}_p)$ of two by two matrices with entries in $\mathbf{Q}_p$; after choosing such an isomorphism, the group

$$\Gamma = R[1/p]_1^\times \tag{14}$$

of elements of norm 1 in $R[1/p]$ can be viewed as a subgroup of $\mathbf{SL}_2(\mathbf{Q}_p)$. This Γ is an example of a p-adic arithmetic subgroup of $\mathbf{SL}_2(\mathbf{Q}_p)$; in fact, if E is an elliptic curve of conductor $2p$, then $E(\mathbf{C}_p)$ is uniformized by $\mathcal{H}_p/\Gamma$.

The pull-back to $\mathcal{H}_p$ of a suitable invariant differential ω on E yields a Γ-invariant differential $f(z)dz$ on $\mathcal{H}_p$. The function f is a *rigid analytic modular form* of weight two on $\mathcal{H}_p$, i.e., a rigid analytic function on $\mathcal{H}_p$ satisfying the transformation property analogous to (8)

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 f(z), \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma. \tag{15}$$

Schneider's p-adic L-functions. By analogy with the construction of $L(f, s)$, the following goal seems natural.

Goal. *Attach to a rigid analytic modular form f a p-adic L-function $L_p(f, s)$, by a process of p-adic Mellin transform.*

What is desired here is a $\mathbf{C}_p$-valued function of the variable $s \in \mathbf{C}_p$ which is rigid analytic, at least in a neighbourhood of $s = 1$. A definition along those lines was proposed by Schneider [Sch], by associating to f a p-adic measure μ_f on $\mathbf{P}_1(\mathbf{Q}_p)$, the p-adic boundary of $\mathcal{H}_p$. This measure behaves like the boundary measure attached to f, and indeed it satisfies the following analogue of the Poisson inversion formula [Te] which allows f to be recovered from μ_f:

$$f(z) = \int_{\mathbf{P}_1(\mathbf{Q}_p)} \frac{d\mu_f(t)}{z-t}.$$

By analogy with formula (9) for the complex L-function, Schneider proposed defining

$$L_p(f,s) := \int_{\mathbf{Z}_p^\times} \langle x \rangle^{s-1} d\mu_f(x), \tag{16}$$

where $\mathbf{Z}_p^\times \subset \mathbf{P}_1(\mathbf{Q}_p)$ is the group of p-adic units and $\langle x \rangle = x/(\lim_{n\to\infty} x^{p^n})$.

This definition does not lead to a satisfactory theory of p-adic L-functions, because the definition of $L_p(f,s)$ is sensitive to the identification of B_p with $M_2(\mathbf{Q}_p)$ used to make Γ act on $\mathcal{H}_p$ and $\mathbf{P}_1(\mathbf{Q}_p)$. It appears that even the order of vanishing of $L_p(f,s)$ depends on these choices, and so it is doubtful that a conjecture analogous to conjecture BSD can be formulated for Schneider's $L_p(f,s)$.

The Iovita-Spiess construction. A definition of a p-adic L-function which is modelled on Schneider's approach, but does lead to a fruitful theory, was proposed by Adrian Iovita in a graduate course at McGill University, and independently by Michael Spiess. The Iovita-Spiess construction is best explained in the special case of the group Γ of equation (14). (The full details are given in [BDIS].) Let K be a maximal commutative subalgebra of the quaternion algebra B. It is isomorphic to a quadratic imaginary field in which the prime 2 is either inert or ramified. Let $\mathcal{O}$ be the ring of integers of K. Replacing K by a conjugate subalgebra, one may assume that

$$K \cap \Gamma = (\mathcal{O}_K[1/p])^\times.$$

In fact, if the p-class group $\mathrm{Pic}(\mathcal{O}[1/p])$ is trivial, the subalgebra K with this property is unique up to conjugation by elements of Γ. Assume for simplicity that this is the case. The identification $B_p = M_2(\mathbf{Q}_p)$ yields an action of $K_p^\times := (K \otimes \mathbf{Q}_p)^\times$ on the boundary $\mathbf{P}_1(\mathbf{Q}_p)$ of $\mathcal{H}_p$, having at most two fixed points and acting transitively on the complement Ω. A choice of base point in Ω thus yields a continuous map

$$\eta : K_p^\times/\mathbf{Q}_p^\times \longrightarrow \Omega \subset \mathbf{P}_1(\mathbf{Q}_p).$$

Let $\mu_{f,K} := \eta^*(\mu_f)$ be the pullback of Schneider's measure μ_f to a measure on $K_p^\times/\mathbf{Q}_p^\times$, and let $\bar{\mu}_{f,K}$ be the measure obtained by composing $\mu_{f,K}$ with complex conjugation on $K_p^\times$. The invariance of μ_f under Γ translates into the invariance of $\mu_{f,K}$ and $\bar{\mu}_{f,K}$ under the action of $\mathcal{O}[1/p]^\times$, and hence

yields measures on the compact p-adic group $G_\infty := K_p^\times/(\mathbf{Q}_p^\times \mathcal{O}[1/p]^\times)$. Letting $\mu_{f,K}^{(2)}$ be the convolution measure $\mu_{f,K} * \bar{\mu}_{f,K}$, define, in analogy with Schneider's definition (16):

$$L_p(f,K,s) := \int_{G_\infty} \langle \frac{x}{\bar{x}} \rangle^{s-1} d\mu_{f,K}^{(2)}(x).$$

Note the crucial role played by the quadratic imaginary field K in the definition of $L_p(f,K,s)$. In fact, the measure $d\mu_{f,K}^{(2)}$ interpolates special values of the complex L-function $L(f/K,s)$ of f over K. More precisely, if $\chi : G_\infty \longrightarrow \mathbf{C}_p^\times$ is a non-trivial character of finite order, interpreted as an idèle class character in the usual way, there is the interpolation formula

$$\int_{G_\infty} \chi(x) d\mu_{f,K}^{(2)}(x) = \Omega_p \frac{L(f/K,\chi,1)}{\Omega_\infty}, \tag{17}$$

where $\Omega_p \in \mathbf{C}_p$ and $\Omega_\infty \in \mathbf{C}$ are suitable p-adic and complex periods, and $L(f/K,\chi,s)$ is the complex L-function of f over K twisted by the character χ. The complex number $\frac{L(f/K,\chi,1)}{\Omega_\infty}$ turns out to be algebraic and is viewed as an element of $\mathbf{C}_p$ by choosing an embedding of $\bar{\mathbf{Q}}$ in $\mathbf{C}_p$. This interpolation formula follows from a generalization of a formula of Gross [Gr1] for special values of L-series. See [BDIS] for details.

The p-adic Birch and Swinnerton-Dyer conjecture. If E is an elliptic curve over $\mathbf{Q}$ satisfying the conclusion of theorem STW_p, so that it is attached to a rigid analytic modular form f on $\mathcal{H}_p$, define

$$L_p(E,K,s) := L_p(f,K,s).$$

Even before the connection with Schneider's approach was made explicit, the p-adic L-function $L_p(E,K,s)$ could be constructed from a different and more general point of view, which does not rely on p-adic analysis and also allows the definition of $L_p(E,K,s)$ in the good reduction case, where $p \nmid N$. In this level of generality, the p-adic Birch and Swinnerton-Dyer conjecture for the function $L_p(E,K,s)$ was formulated and studied in a series of articles [BD1], [BD2], [BD3], [BD4] and [BD5].

Because of the presence of the field K in the interpolation formula (17), it is natural to expect the order of vanishing of $L_p(E,K,s)$ to be related to

the rank r_K of the Mordell–Weil group $E(K)$. In [BD1], it was conjectured that

$$\mathrm{ord}_{s=1} L_p(E, K, s) \geq r_K.$$

This p-adic variant of the "easy half" (11) of the Birch Swinnerton–Dyer conjecture has recently been proved in [BD6].

Theorem BD. *The p-adic L-function $L_p(E, K, s)$ vanishes to order at least r_K at $s = 1$.*

The proof of theorem BD is based on two ingredients.

1. The theory of congruences between modular forms and the Jacquet-Langlands correspondence. This circle of ideas plays a crucial role in Wiles' proof of theorem STW, and in Ribet's earlier reduction [Ri] of Fermat's Last Theorem to the Shimura-Taniyama conjecture.

2. Kolyvagin's theory of the "Euler systems" of Heegner points, the principal ingredient in the proof of theorem GZK.

Thus, theorem BD relies crucially on the ideas of Gross-Zagier, Kolyvagin, Ribet, and Wiles, which have revolutionized the theory of elliptic curves through the proofs of theorems GZK and STW.

To conclude, here are two natural questions connected with the original Birch and Swinnerton-Dyer conjecture.

1. Theorem BD can be used to exhibit elliptic curves whose p-adic L-function $L_p(E, K, s)$ satisfies

   $$\mathrm{ord}_{s=1} L_p(E, K, s) \geq 22.$$

   Does the existence of such p-adic analytic L-functions with high order zeroes have independent applications to other questions of number theory (or mathematics in general), as in Goldfeld's solution of Gauss's class number problem?

2. Is it possible to replace the rigid analytic L-functions by classical ones in the proof of theorem BD? The proof in [BD6] is based on congruences in an essential way and breaks down entirely when the prime p is replaced by the "place at ∞". In this sense, it sheds *no light* on the original Birch and Swinnerton–Dyer conjecture, even on the "easy inequality".

As Mazur writes in [Ma3],

> "A major theme in the development of number theory has been to try to bring $\mathbf{R}$ somewhat more into line with the p-adic fields; a major mystery is why $\mathbf{R}$ resists this attempt so strenuously. "

An explanation of the mysterious analogy between the complex and p-adic realms would surely lead to deep insights: it is an issue which lies at the heart of the tantalizing and elusive Birch and Swinnerton–Dyer conjecture.

References

[BC] A. Bremner, J.W.S. Cassels, *On the equation* $Y^2 = X(X^2 + p)$, Math. Comp. **42** (1984) 257–264.

[BD1] M. Bertolini and H. Darmon, *Heegner points on Mumford-Tate curves.* Invent. Math **126** 413–456 (1996).

[BD2] M. Bertolini and H. Darmon, *A rigid-analytic Gross-Zagier formula and arithmetic applications.* Annals of Math **146** (1997) 111-147.

[BD3] M. Bertolini and H. Darmon, *Heegner points, p-adic L-functions, and the Cerednik-Drinfeld uniformization.* Invent. Math, **131** (1998), no. 3, 453–491.

[BD4] M. Bertolini and H. Darmon, *p-adic periods, p-adic L-functions and the p-adic uniformization of Shimura curves*, Duke Math J., to appear.

[BD5] M. Bertolini and H. Darmon, *Euler systems and Jochnowitz congruences*, Amer. J. Math., to appear.

[BD6] M. Bertolini and H. Darmon, *The p-adic Birch and Swinnerton-Dyer conjecture*, in progress.

[BDIS] M. Bertolini, H. Darmon, A. Iovita, M. Spiess, *Teitelbaum's exceptional zero conjecture in the anticyclotomic setting.*, in progress.

[CDT] B. Conrad, F. Diamond, R. Taylor, *Modularity of certain potentially Barsotti-Tate Galois representations.* J. Amer. Math. Soc., to appear.

[DDT] H. Darmon, F. Diamond, R. Taylor, *Fermat's Last Theorem*, Current Developments in Math, Vol. **1**, pp. 1–157, International Press, 1996.

[Di] F. Diamond, *On deformation rings and Hecke rings*. Ann. of Math. (2) **144** (1996), no. 1, 137–166.

[Fe] S. Fermigier, *Une courbe elliptique définie sur* $\mathbf{Q}$ *de rang* ≥ 22. Acta Arith. **82** (1997), no. 4, 359–363.

[Gh] A. Ghitza, *Heights on* $E(\mathbf{Q})$, CICMA preprint series, August 1997.

[Go] D. Goldfeld, *Gauss's class number problem for imaginary quadratic fields*. Bull. Amer. Math. Soc. (N.S.) **13** (1985), no. 1, 23–37.

[Gr1] B.H. Gross, *Heights and the special values of L-series*. Number theory (Montreal, Que., 1985), 115–187, CMS Conf. Proc., 7, Amer. Math. Soc., Providence, RI, 1987.

[GZ] B.H. Gross, D.B. Zagier. *Heegner points and derivatives of L-series*. Invent. Math. **84** (1986), no. 2, 225–320.

[JL] B.W. Jordan, R. Livne, *Local diophantine properties of Shimura curves*, Math. Ann. **270** (1985), 235-248.

[Ko] Kolyvagin, V. A. *The Mordell-Weil and Shafarevich-Tate groups for Weil elliptic curves*. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **52** (1988), no. 6, 1154–1180, 1327; translation in Math. USSR-Izv. **33** (1989), no. 3, 473–499.

[Ma1] B. Mazur, *p-adic analytic number theory of elliptic curves and Abelian varieties over* Q. Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 1, pp. 369–377. Canad. Math. Congress, Montreal, Que., 1975.

[Ma2] B. Mazur. *Number theory as gadfly*. Amer. Math. Monthly **98** (1991), no. 7, 593–610.

[Ma3] B. Mazur. *On the passage from local to global in number theory*. Bull. Amer. Math. Soc. (N.S.) **29** (1993), no. 1, 14–50.

[MTT] B. Mazur, J. Tate, J. Teitelbaum. *On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer*. Invent. Math. **84** (1986), no. 1, 1–48.

[Mu] R. Murty, *Some remarks on the Riemann hypothesis*, in preparation.

[Ri] K. Ribet, *On modular representations of Gal*($\bar{\mathbf{Q}}/\mathbf{Q}$) *arising from modular forms*, Invent. Math. **100** (1990), 431–476.

[Sch] Schneider, P. *Rigid-analytic L-transforms*. Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), 216–230, Lecture Notes in Math., 1068, Springer, Berlin-New York, 1984.

[Si] J.H. Silverman, The arithmetic of elliptic curves. Corrected reprint of the 1986 original. Graduate Texts in Mathematics, 106. Springer-Verlag, New York, 1992.

[ST] J.H. Silverman, J. Tate, Rational points on elliptic curves. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1992.

[Ta1] J. Tate. *The arithmetic of elliptic curves*. Invent. Math. **23** (1974), 179–206.

[Ta2] J. Tate, *On the conjectures of Birch and Swinnerton-Dyer and a geometric analog*. Séminaire Bourbaki, Vol. 9, Exp. No. 306, 415–440, Soc. Math. France, Paris, 1995.

[Te] J. Teitelbaum, *Values of p-adic L-functions and a p-adic Poisson kernel*. Invent. Math. **101** (1990), no. 2, 395–410.

[TW] R. Taylor and A. Wiles, *Ring theoretic properties of certain Hecke algebras*, Annals of Math. **141**, No. 3, 1995, pp. 553-572.

[Wi] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Annals of Math. **141**, No. 3, 1995, pp. 443-551.