Wiles' theorem and the arithmetic of elliptic curves

H. Darmon

September 9, 2007

Contents

Thanks to the work of Wiles [**?**], completed by Taylor–Wiles [**?**] and extended by Diamond [**?**], we now know that all elliptic curves over the rationals (having good or semi-stable reduction at 3 and 5) are modular. This breakthrough has far-reaching consequences for the arithmetic of elliptic curves. As Mazur wrote in [**?**], "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 expository article explores some of the implications of Wiles' theorem for the theory of elliptic curves, with particular emphasis on the Birch and Swinnerton-Dyer conjecture, now the main outstanding problem in the field.

1 Prelude: plane conics, Fermat and Gauss

In a volume devoted to Wiles' proof of Fermat's Last Theorem, what better place to begin this discussion than the Diophantine equation

$$C : x^2 + y^2 = 1, \tag{1}$$

which also figured prominently in Diophantus' treatise, and prompted Fermat's famous marginal comment, more than 350 years ago?

The set $C(\mathbf{Q})$ of rational solutions to equation (1) is well understood, thanks to the parametrization

$$(x, y) = \left(\frac{t^2 - 1}{t^2 + 1}, \frac{2t}{t^2 + 1}\right), \tag{2}$$

giving the classification of Pythagorean triples well-known to the ancient Babylonians. The integer solutions are even simpler: there are $N_{\mathbf{Z}} = 4$ integer lattice points $(\pm 1, 0)$ and $(0, \pm 1)$ on the circle of radius 1.

It has become a dominant theme in number theory that curves such as C ought to be studied over various fields, such as the real or complex numbers, the finite fields $\mathbf{F}_p$, and the p-adic fields $\mathbf{Q}_p$, for each prime p.

The solutions to (1) in $\mathbf{R}^2$ describe the locus of points on the circle of radius 1. A natural measure of the size of this solution set is the circumference of the circle: $N_{\mathbf{R}} = 2\pi$.

The set of $\mathbf{F}_p$-valued solutions $C(\mathbf{F}_p)$ is finite, of cardinality N_p. Let $a_p = p - N_p$. Is there a convenient formula for N_p, or equivalently, for

a_p? Letting t run over the values $t = 0, 1, 2, \ldots, p-1, \infty \in \mathbf{P}_1(\mathbf{F}_p)$ in the parametrization (2) gives $p+1$ distinct points in $C(\mathbf{F}_p)$, with one important caveat: if $t^2 + 1 = 0$ has a solution $t_0 \in \mathbf{F}_p$, then the values $t = \pm t_0$ do not give rise to points over $\mathbf{F}_p$. Hence, if p is odd:

$$a_p = \begin{cases} +1 & \text{if } -1 \text{ is a square mod } p; \\ -1 & \text{if } -1 \text{ is not a square mod } p. \end{cases}$$

The condition which determines the value of a_p might seem subtle to the uninitiated. But much the opposite is true, thanks to the following result which is due to Fermat himself:

Theorem 1.1 (Fermat) *If p is an odd prime,*

$$a_p = \begin{cases} +1 & \textit{if } \ p \equiv 1 \pmod 4 \\ -1 & \textit{if } \ p \equiv 3 \pmod 4, \end{cases}$$

and $a_2 = 0$.

(The computational advantage of this formula is obvious. It now suffices to glance at the last two decimal digits of p to determine whether N_p is equal to $p-1$ or $p+1$.)

Let

$$L(C/\mathbf{Q}, s) = \prod_p (1 - a_p p^{-s})^{-1}$$

be the "Hasse-Weil zeta-function" associated to C. Thanks to Fermat's theorem 1.1, one has:

Corollary 1.2 *The Hasse-Weil L-function $L(C/\mathbf{Q}, s)$ is equal to a Dirichlet L-series $L(s, \chi)$, where $\chi : (\mathbf{Z}/4\mathbf{Z})^\times \longrightarrow \pm 1$ is the unique non-trivial quadratic Dirichlet character of conductor 4. In particular, $L(C/\mathbf{Q}, s)$ has a functional equation and an analytic continuation to the entire complex plane.*

More precisely (cf. [**?**]. ch. 4), setting

$$\Lambda(C/\mathbf{Q}, s) = \left(\frac{4}{\pi}\right)^{s/2} \Gamma\left(\frac{s+1}{2}\right) L(C/\mathbf{Q}, s),$$

we have:

$$\Lambda(C/\mathbf{Q}, s) = \Lambda(C/\mathbf{Q}, 1-s). \qquad (3)$$

The special value $L(C/\mathbf{Q},1)$ is given by:

$$L(C/\mathbf{Q},1) = L(1,\chi) = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \frac{\pi}{4}. \tag{4}$$

Noting the formal equality $L(C/\mathbf{Q},1)\text{“}=\text{”}\prod_p \frac{p}{N_p}$, equation (4) can be rewritten in the suggestive form:

$$\prod_p \frac{N_p}{p} \cdot N_{\mathbf{R}} = 2N_{\mathbf{Z}}, \tag{5}$$

a formula which suggests a mysterious link between the solutions to C over the reals, the finite fields $\mathbf{F}_p$, and the integers. The proof that we have sketched, although quite simple, does little to dispell the mystery.

Another example which was also at the center of Fermat's preoccupations is the Fermat-Pell equation

$$H : x^2 - Dy^2 = 1, \tag{6}$$

where D is a positive square free integer. Assume for simplicity that D is congruent to 1 mod 4.

Defining the integers N_p and $a_p = p - N_p$ as before, one finds that for $p \not| 2D$,

$$a_p = \begin{cases} +1 & \text{if } D \text{ is a square mod } p; \\ -1 & \text{if } D \text{ is not a square mod } p. \end{cases}$$

Extend the definition of a_p by setting $a_p = 0$ if $p|2D$. By Gauss's theorem of quadratic reciprocity:

Theorem 1.3 (Gauss) *Let*

$$\chi_D : (\mathbf{Z}/D\mathbf{Z})^\times \longrightarrow \pm 1$$

be the even (non-primitive) Dirichlet character of conductor $2D$ *defined by* $\chi_D(n) = \left(\frac{n}{D}\right)$. *Then* $a_p = \chi_D(p)$.

Define the Hasse-Weil L-function $L(H/\mathbf{Q},s) = \prod_p (1 - a_p p^{-s})^{-1}$ as before.

Corollary 1.4 *The function* $L(H/\mathbf{Q},s)$ *is equal to the Dirichlet L-series* $L(s,\chi_D)$, *so that it has a functional equation and an analytic continuation to the entire complex plane.*

The precise functional equation, similar to equation (3), can be found in [?], ch. 4.

As before the value $L(H/\mathbf{Q}, 1)$ can be evaluated in closed form (cf. [?], thm. 4.9):

$$L(H/\mathbf{Q}, 1) = L(1, \chi_D) = \sum_{n=1}^{\infty} \frac{\chi_D(n)}{n} = \frac{3}{2\sqrt{D}} \sum_{a=1}^{D-1} \chi_D(a) \log |1 - \zeta_D^a|, \qquad (7)$$

where $\zeta_D = e^{2\pi i/D}$ is a primitive D-th root of unity.

To gain further insight into the arithmetic significance of this special value, one uses the following result of Gauss, which is a primary ingredient in one of his proofs of quadratic reciprocity, and is in fact essentially equivalent to it.

Theorem 1.5 *Every quadratic field is contained in a cyclotomic field generated by roots of unity. More precisely, the quadratic field* $\mathbf{Q}(\sqrt{D})$ *is contained in* $\mathbf{Q}(\zeta_D)$*, and the homomorphism of Galois theory*

$$\mathit{Gal}(\mathbf{Q}(\zeta_D)/\mathbf{Q}) = (\mathbf{Z}/D\mathbf{Z})^\times \longrightarrow \mathit{Gal}(\mathbf{Q}(\sqrt{D})/\mathbf{Q}) = \pm 1$$

is identified with the Dirichlet character χ_D*.*

One of the applications of theorem 1.5 is that it gives a natural way of finding units in $\mathbf{Q}(\sqrt{D})$, and thereby solving Pell's equation. Indeed, the cyclotomic field $\mathbf{Q}(\zeta)$ is equipped with certain natural units, the so-called *circular units.* These are algebraic integers of the form $(1 - \zeta_D^a)$ if D is not prime, and of the form $\frac{1-\zeta_D^a}{1-\zeta_D}$ if D is prime, with $a \in (\mathbf{Z}/D\mathbf{Z})^\times$. In particular, theorem 1.5 implies that the expression

$$u_D = \prod_{a=1}^{D} (1 - \zeta_D^a)^{3\chi_D(a)}$$

is an element of norm 1 in the quadratic field $\mathbf{Q}(\sqrt{D})$, and in fact, in the ring $\mathbf{Z}[\sqrt{D}]$. Hence, formula (7) can be rewritten:

$$L(H/\mathbf{Q}, 1) = \frac{1}{\sqrt{4D}} \log |x_0 + y_0\sqrt{D}|, \qquad (8)$$

where (x_0, y_0) is an integer solution to equation (6). The non-vanishing of $L(1, \chi_D)$, (or, equivalently, by the functional equation, of $L'(0, \chi_D)$) implies that this solution is non-trivial.

Remark: A natural generalization of theorem 1.5, the Kronecker-Weber theorem, states that every abelian extension of the rationals is contained in a cyclotomic field. The norms of circular units always give a subgroup of finite index in the group of units of L.

2 Elliptic curves and Wiles' theorem

Let $E/\mathbf{Q}$ be an elliptic curve over the rationals of conductor N, given by the projective equation

$$y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz^2 + a_6z^3. \tag{9}$$

By the Mordell-Weil theorem, the Mordell-Weil group $E(\mathbf{Q})$ is a finitely generated abelian group,

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

where T is the finite torsion subgroup of $E(\mathbf{Q})$. Paraphrasing a remark of Mazur ([**?**], p. 186), there are resonances between the problem of studying integer points on plane conics and rational points on elliptic curves. In the basic trichotomy governing the study of curves over $\mathbf{Q}$, these Diophantine problems correspond to the only classes of curves having Euler characteristic equal to 0. (The Euler characteristic $\chi(X)$ depends only on the Riemann surface $X(\mathbf{C})$ which is topologically equivalent to a compact surface of genus g with s points removed; it is defined by

$$\chi(X) = (2 - 2g) - s.)$$

2.1 Wiles' theorem and $L(E/\mathbf{Q}, s)$

If p is a prime of good reduction for E, let N_p be the number of distinct solutions to equation (9) in $\mathbf{P}^2(\mathbf{F}_p)$, and set

$$a_p = p + 1 - N_p.$$

Further, set $a_p = 1$ if $E/\mathbf{Q}_p$ has split multiplicative reduction, $a_p = -1$ if $E/\mathbf{Q}_p$ has non-split multiplicative reduction, and $a_p = 0$ otherwise. Define the Hasse-Weil L-function $L(E/\mathbf{Q}, s)$ by the formula

$$L(E/\mathbf{Q}, s) = \prod_{p \nmid N} (1 - a_p p^{-s} + p^{1-2s})^{-1} \prod_{p | N} (1 - a_p p^{-s})^{-1}.$$

To study the elliptic curve E along the lines of section 1, one needs a better understanding of the coefficients a_p, allowing an analysis of the L-function $L(E/\mathbf{Q}, s)$. This is precisely the content of Wiles' theorem, stated here in a form which is analogous to theorems 1.1 and 1.3.

Theorem 2.1 ([?],[?], [?]) *Assume that E has good or semi-stable reduction at 3 and 5. Then the coefficients a_p are the Fourier coefficients of a modular form f of weight 2 and level N which is an eigenform for all the Hecke operators T_p.*

This result gives has the following elliptic curve analogue of corollary 1.4.

Corollary 2.2 (Hecke) *The L-function $L(E/\mathbf{Q}, s)$ is equal to the L function $L(f, s)$ attached to the eigenform f. In particular, it has an analytic continuation and a functional equation.*

More precisely, setting

$$\Lambda(E/\mathbf{Q}, s) = N^{s/2}(2\pi)^{-s}\Gamma(s)L(E/\mathbf{Q}, s),$$

we have

$$\Lambda(E/\mathbf{Q}, s) = \int_0^\infty f\left(\frac{iy}{\sqrt{N}}\right) y^s \frac{dy}{y}, \tag{10}$$

and

$$\Lambda(E/\mathbf{Q}, s) = w\Lambda(E/\mathbf{Q}, 2 - s), \tag{11}$$

where $w = \pm 1$ can be computed as a product of local signs. For example:

Proposition 2.3 *If $E/\mathbf{Q}$ is a semistable curve, then w is equal to $(-1)^{s+1}$, where s is the number of primes of split multiplicative reduction for $E/\mathbf{Q}$.*

Remark: The statements of Wiles' theorem given in theorem 2.1 and corollary 2.2 bear a strong ressemblance to theorem 1.3 and corollary 1.4 respectively. This is only fitting, as Wiles' theorem is a manifestation of a non-abelian reciprocity law for $\mathbf{GL}_2$, having its roots ultimately in the fundamental quadratic reciprocity law of Gauss.

More germane to the discussion of section 1, Wiles' achievement allows one to make sense of the special values $L(E/\mathbf{Q}, s) = L(f, s)$ even when s is outside the domain $\{\mathrm{Real}(s) > \frac{3}{2}\}$ of absolute convergence of the infinite product used to define $L(E/\mathbf{Q}, s)$. This is of particular interest for the point $s = 1$, which is related conjecturally to the arithmetic of $E/\mathbf{Q}$ by the Birch and Swinnerton-Dyer conjecture.

Conjecture 2.4 *The Hasse-Weil L-function $L(E/\mathbf{Q}, s)$ vanishes to order r (=rank$(E/\mathbf{Q})$) at $s = 1$, and*

$$L^{(r)}(E/\mathbf{Q}, s) = \#\underline{III}(E/\mathbf{Q}) \left(\det\left(\langle P_i, P_j\rangle\right)_{1\le i,j\le r}\right) \#T^{-2} \left(\int_{E(\mathbf{R})} \omega\right) \prod_p m_p,$$

where $\underline{III}(E/\mathbf{Q})$ is the (conjecturally finite) Shafarevich-Tate group of $E/\mathbf{Q}$, the points $P_1, \ldots, P_r$ are a basis for $E(\mathbf{Q})$ modulo torsion, $\langle\ ,\ \rangle$ is the Néron-Tate canonical height, ω is the Néron differential on E, and m_p is the number of connected components in the Néron model of $E/\mathbf{Q}_p$.

Motivated by this conjecture, one calls the order of vanishing of $L(E/\mathbf{Q}, s)$ at $s = 1$ the *analytic rank* of $E/\mathbf{Q}$, and denotes it r_{an}.

If E is a semistable elliptic curve, then the formula for w given in proposition 2.3 implies that $t + r_{an}$ is always even, where t denotes the number of analytic uniformizations (complex and p-adic) with which $E/\mathbf{Q}$ is endowed. Hence, a corollary of conjecture 2.4 is the following parity conjecture for the rank:

Conjecture 2.5 *If $E/\mathbf{Q}$ is semistable, then the integer $r + t$ is even.*

A great deal of theoretical evidence is available for conjecture 2.4 when the analytic rank r_{an} is equal to 0 or 1. By contrast, very little is known when $\mathrm{ord}_{s=1} L(E/\mathbf{Q}, s) > 1$, and so conjecture 2.4, and even conjecture 2.5, remain very mysterious. (Some numerical evidence has been gathered for certain specific elliptic curves, such as the curve of rank 3 and conductor 5077, cf. [**?**].)

2.2 Geometric versions of Wiles' theorem

To tackle conjecture 2.4 requires an explicit formula for the leading term of $L(E/\mathbf{Q}, s)$ at $s = 1$. There are such formulae where the analytic rank is 0 or 1. In deriving them, essential use is made of the following geometric version of Wiles' theorem, which may be seen as a direct analogue of theorem 1.5:

Theorem 2.6 *Suppose that E has good or semistable reduction at 3 and 5. Then the elliptic curve E is uniformized by the modular curve $X_0(N)$, i.e., there is a non-constant algebraic map defined over $\mathbf{Q}$:*

$$\phi : X_0(N) \longrightarrow E.$$

Here $X_0(N)$ is the usual modular curve which is a (coarse) moduli space classifying elliptic curves together with a cyclic subgroup of order N. Its complex points can be decribed analytically as a compactification of thequotient

$$Y_0(N)_{/\mathbf{C}} = \mathcal{H}/\Gamma_0(N),$$

where $\Gamma_0(N)$ is the usual congruence subgroup of level N of $\mathbf{SL}_2(\mathbf{Z})$, and $\mathcal{H}$ is the complex upper half plane of complex numbers τ with $\mathrm{Im}(\tau) > 0$.

The pull-back of the Néron differential ω on E is an integer multiple of the differential $2\pi i f(\tau)d\tau = f(q)\frac{dq}{q}$, where f is the modular form given in theorem 2.1 and $q = e^{2\pi i\tau}$:

$$\phi^*\omega = cf(q)\frac{dq}{q}. \qquad (12)$$

The integer c is called the *Manin constant* associated to ϕ. When the degree of ϕ is minimal (among all possible maps $X_0(N) \longrightarrow E'$ with E' isogenous to E) it is conjectured that $c = 1$.

When theorem 2.6 is satisfied, the elliptic curve E is also uniformized by other arithmetic curves, the Shimura curves associated to indefinite quaterion algebras. Although somewhat less studied than classical modular curves, they are endowed with a similarly rich arithmetic structure. They play an important role in Ribet's fundamental "lowering the level" results (cf. the article of Edixhoven in this volume). It is also likely that a deeper understanding of the arithmetic of elliptic curves might be achieved by considering the collection of all modular and Shimura curve parametrizations simultaneously. (See for example the remarks in [**?**].)

Let $N = N^+N^-$ be a factorization of N such that N^- is square-free, is the product of an even number of primes, and satisfies $\gcd(N^+, N^-) = 1$. Let B be the indefinite quaternion algebra which is ramified exactly at the primes dividing N^-, and let R be a maximal order in B. The algebra B is unique up to isomorphism, and any two maximal orders in B are conjugate. (For more on the arithmetic of quaternion algebras over $\mathbf{Q}$, see [**?**].) The Shimura curve X_{1,N^-} is defined as a (coarse) moduli space for abelian surfaces with quaternionic multiplication by R, i.e., abelian surfaces A equipped with a map

$$R \longrightarrow \mathrm{end}(A).$$

The curve X_{N^+,N^-} is a (coarse) moduli space for abelian surfaces with quaternionic multiplication by R, together with a subgroup scheme generically isomorphic to $\mathbf{Z}/N^+\mathbf{Z} \times \mathbf{Z}/N^+\mathbf{Z}$ and stable under the action of R. Shimura showed that the curves X_{1,N^-} and X_{N^+,N^-} have canonical models over $\mathbf{Q}$. Let J_{N^+,N^-} be the Jacobian of X_{N^+,N^-}. By a theorem of Jacquet-Langlands [**?**], it is isogenous to a factor of the Jacobian $J_0(N)$ corresponding to the forms of level N which are *new* at the primes dividing N^-, and hence we have:

Theorem 2.7 *Suppose that E has good or semistable reduction at* 3 *and* 5*. Then E is a factor of the Jacobian J_{N^+,N^-}, i.e., there is a non-constant algebraic map ϕ_{N^+,N^-} defined over $\mathbf{Q}$:*

$$\phi_{N^+,N^-} : J_{N^+,N^-} \longrightarrow E.$$

A nice account of the theory of Shimura curves can be found in [**?**] and [**?**].

Remark: The case where $N^- = 1$ corresponds to the case of the usual modular curves. In this case, the algebra B is the matrix algebra $M_2(\mathbf{Q})$, the order R can be chosen to be $M_2(\mathbf{Z})$, and an abelian surface with endomorphisms by R is isomorphic to a product $A = E \times E$, where E is an elliptic curve. The level N structure on A corresponds to a usual level N structure on E, so that the curve $X_{N,1}$ is isomorphic to $X_0(N)$.

In general, there is considerable freedom in choosing the map ϕ_{N^+,N^-}. One rigidifies the situation by requiring that ϕ_{N^+,N^-} be *optimal*, i.e., that its kernel be a (connected) abelian subvariety of J_{N^+,N^-}. This can always be accomplished, if necessary by replacing E by another elliptic curve in the same isogeny class.

Likewise, we will always assume in the next section that the morphism ϕ of theorem 2.6 sends the cusp $i\infty$ to the identity of E, and that the map induced by ϕ on Jacobians is optimal.

3 The special values of $L(E/\mathbf{Q}, s)$ at $s = 1$

We now review some of the information on the leading term of $L(E/\mathbf{Q}, s)$ at $s = 1$ which can be extracted from the knowledge that E is modular.

3.1 Analytic rank 0

Theorem 3.1 *There is a rational number M such that*

$$L(E/\mathbf{Q}, 1) = M \int_{E(\mathbf{R})} \omega,$$

where ω is a Néron differential on E.

Proof: Let $0, i\infty$ be the usual cusps in the extended upper half-plane, and let ϕ be the modular parametrization of theorem 2.6. The theorem of Manin-Drinfeld says that the divisor $(i\infty) - (0)$ is torsion in $J_0(N)$, and hence, if ϕ sends $i\infty$ to the point at infinity on E, then $\phi(0)$ is a torsion point in E. By composing ϕ with an isogeny, assume without loss of generality that $\phi(0) = \phi(i\infty)$ is the identity element in $E(\mathbf{Q})$. Then the modular parametrization ϕ induces a map from the interval $[0, i\infty]$ (with the points 0 and $i\infty$ identified) to the connected component $E_0(\mathbf{R})$ of $E(\mathbf{R})$. Let M_0 be the winding number of this map between two circles. By the formula for $L(E/\mathbf{Q}, 1)$ of equation (10),

$$L(E/\mathbf{Q}, 1) = 2\pi i \int_0^{i\infty} f(\tau) d\tau = \frac{1}{c} \int_0^{i\infty} \phi^* \omega = \frac{M_0}{c} \int_{E_0(\mathbf{R})} \omega = M \int_{E(\mathbf{R})} \omega,$$

where $M = \frac{M_0}{c}[E(\mathbf{R}) : E_0(\mathbf{R})]^{-1}$.

The reader should compare theorem 3.1 with equation (4), which also expresses the special value $L(C/\mathbf{Q}, 1)$ as a rational multiple of the period 2π.

While theorem 3.1 gives some evidence for the Birch and Swinnerton-Dyer conjecture, proving that the value of $L(E/\mathbf{Q}, 1)$ is the correct one "up to rational multiples", it does not shed much light on the relation between M and arithmetic quantities associated to E such as the rank of $E/\mathbf{Q}$ and the order of $\underline{III}(E/\mathbf{Q})$.

3.2 Analytic rank 1: the Gross-Zagier formula

Assume now that the sign w in the functional equation (11) for $L(E/\mathbf{Q}, s)$ is -1, so that the L-function of $E/\mathbf{Q}$ vanishes to odd order. The L-function $L(E/\mathbf{Q}, s)$ now has an "automatic zero" at $s = 1$, and one might hope for a natural closed form expression for the special value $L'(E/\mathbf{Q}, 1)$.

Rather surprisingly, no really "natural" closed form expression is known. Instead, a formula can only be written down after choosing an auxiliary quadratic imaginary field K. Let K be such a field, D its discriminant, and let χ be the associated odd Dirichlet character. Let $E^{(D)}$ be the quadratic twist of E, relative to the character χ. Consider the L-series

$$L(E/K, s) := L(E/\mathbf{Q}, s)L(E^{(D)}/\mathbf{Q}, s).$$

It can be shown that this L-series has an analytic continuation and a functional equation relating its value at s and $2 - s$, in (at least) two different ways. Since E and $E^{(D)}$ are both modular, each of the two factors on the right has a functional equation and analytic continuation. Alternately, the functional equation for $L(E/K, s)$ can be obtained by expressing $L(E/K, s)$ as the Rankin convolution of the L-series $L(f, s)$ with the L-series of a theta-function of weight 1 associated to the imaginary quadratic field K, and applying Rankin's method. (Cf. [**?**], ch. IV). If K is an arbitrary quadratic field (not necessarily quadratic imaginary) one has

Proposition 3.2 *The sign w_K in the functional equation for $L(E/K, s)$ can be expressed as a product of local signs*

$$w_K = \prod_v w_v,$$

where $w_v = \pm 1$ depends only on the behaviour of E over the completion K_v. In particular,

1. *If E has good reduction at v, then $w_v = 1$;*

2. *If v is archimedean, then $w_v = -1$;*

3. *If E/K_v has split (resp. non-split) multiplicative reduction at v then $w_v = -1$ (resp. $w_v = 1$).*

Heegner Points:
Just as cyclotomic fields are equipped with certain canonical units (the circular units) whose logarithms express the special values of Dirichlet L-series, so modular curves and Shimura curves are equipped with a certain natural set of algebraic points, the *Heegner points* associated to the imaginary quadratic field K, whose heights express first derivatives of the L-functions attached to cusp forms.

Let A be any elliptic curve which has complex multiplication by the maximal order $\mathcal{O}_K$ of K. There are exactly h such curves, where h is the class number of K. They are all defined over the Hilbert class field H of K and are conjugate to each other under the action of $\mathrm{Gal}(H/K)$.

Assume further that all the primes dividing the conductor N are split in the imaginary quadratic field K. By prop. 3.2, this implies that $w_K = -1$, so that the analytic rank of $E(K)$ is odd.

Under this hypothesis the complex multiplication curve A has a rational subgroup of order N which is defined over H. This subgroup is not unique, and choosing one amounts to choosing an integral ideal of norm N in the quadratic field K. Choose such a subgroup C of A. The pair (A, C) gives rise to a point α on $X_0(N)$ which is defined over H. It is called a *Heegner point* on $X_0(N)$ (associated to the maximal order $\mathcal{O}_K$). Let $P_H = \phi(\alpha)$ be the image of α on $E(H)$ by the modular parametrization ϕ of theorem 2.6, and let $P_K = \mathrm{trace}_{H/K} P_H$ be its trace to $E(K)$. The point P_K (up to sign) depends only on the quadratic imaginary field K, not on the choice of A and C. Hence, its Néron-Tate height is canonical.

The fundamental theorem of Gross and Zagier expresses the special value of $L'(E/K, 1)$ in terms of the height of P_K.

Theorem 3.3 $L'(E/K,1) = \left(\int\int_{E(\mathbf{C})} \omega \wedge i\bar{\omega}\right) \langle P_K, P_K\rangle / c^2 u_K^2 |D|^{\frac{1}{2}}$.

The proof of this beautiful theorem, which is quite involved, is given in [?].

Remarks:
1. Theorem 3.3 gives a formula for $L'(E/\mathbf{Q}, 1)L(E^{(D)}/\mathbf{Q}, 1)$, and in this sense does not give a "natural" formula for $L'(E/\mathbf{Q}, 1)$ alone.
2. Theorem 3.3 is also true when $w = 1$. In this case, the twisted L-function $L(E^{(D)}/\mathbf{Q}, s)$ vanishes at $s = 1$, and theorem 3.3 gives a formula for $L(E/\mathbf{Q}, 1)L'(E^{(D)}/\mathbf{Q}, 1)$.

3.3 Some variants of the Gross-Zagier formula

The fundamental formula of Gross and Zagier has been extended and generalized in various directions in the last years. Let us mention very briefly a few of these variants:

A. *Shimura curve analogues*: Assume here for simplicity that E is semistable so that N is square-free, and that K is a quadratic imaginary field of discriminant D with $\gcd(N, D) = 1$. Let $N = N^+N^-$ be the factorization of N such that N^+ is the product of all primes which are split in K, and N^- is the product of the primes which are inert in K. By prop. 3.2, the integer N^- is a product of an even number of prime factors if and only if the sign w_K in the functional equation for $L(E/K, s)$ is -1. Assume that $w_K = -1$. The Gross-Zagier formula given in theorem 3.3 corresponds to the case where $N^+ = N, N^- = 1$. Assume now that $N^- \neq 1$. One can then define the Shimura curve X_{N^+,N^-} as in section 2.2.

The curve X_{N^+,N^-} is equipped with Heegner points defined over the Hilbert class field H of K, which correspond to moduli of quaternionic surfaces with level N^+ structure having complex multiplication by $\mathcal{O}_K$, i.e, quaternionic surfaces A endowed with a map

$$\mathcal{O}_K \longrightarrow \underline{\mathrm{end}}(A),$$

where $\underline{\mathrm{end}}(A)$ denotes the algebraic endomorphisms of A which commute with the quaternionic multiplications. By considering the image in the Mordell Weil group $E(H)$ of certain degree zero divisors supported on Heegner points in $J_{N^+,N^-}(H)$ by ϕ_{N^+,N^-}, one obtains a Heegner point P_K in $E(K)$, which cannot be obtained from the modular curve parametrization ϕ. One expects that that the height of P_K can be expressed in terms of the derivative $L'(E/K, 1)$, in a manner analogous to theorem 3.3. In particular, one expects that P_K is of infinite order in $E(K)$ if and only if $L'(E/K, 1) \neq 0$. Nothing as precise has yet been established, but some work in progress of Keating and Kudla supports this expectation.

B. *Perrin Riou's p-adic analogue*: In [**?**], a formula is obtained (when all primes dividing N are split in K) relating the first derivative of the two-variable p-adic L-function of E/K to the p-adic height of the Heegner point P_K. The calculations of [**?**] are also quite involved, but on a conceptual level they follow those of Gross and Zagier quite closely.

C. *Rubin's p-adic formula*: Let E be an elliptic curve with complex multiplication by $\mathcal{O}_K$. In [**?**], Rubin obtains a formula expressing the derivative of the two-variable p-adic L-function of E/K at a point which lies outside the range of classical interpolation, to the p-adic logarithm in the formal group attached to E over $K \otimes \mathbf{Q}_p$ of a Heegner point in $E(K)$. The proof of this formula uses the theory of elliptic units, as well as the formula of Gross Zagier and Perrin-Riou's p-adic analogue, in an essential way. A striking feature of Rubin's formula is that it allows one to recover a rational point in $E(K)$ as the formal group exponential evaluated on an expression involving the first derivative of a p-adic L-function, in much the same way that, if χ is an even Dirichlet character, exponentiating $L'(0, \chi)$ yields a unit in the real quadratic field cut out by χ.

D. *Formulae for $L(E/K, 1)$ when E is a Tate curve*: Suppose that E has a prime p of multiplicative reduction which is inert in K, and suppose that all other primes dividing N are split in K. Then the sign in the functional equation for $L(E/K, s)$ is 1 by prop. 3.2, and one expects no Heegner point construction yielding a point on $E(K)$. However, there are Heegner points $P_n \in E(H_n)$, where H_n is the ring class field of H of conductor p^n, constructed from elliptic curves with level N structure having complex multiplication by the orders of conductor p^n in $\mathcal{O}_K$. The precise construction is explained in [**?**], where it is shown that these points are trace-compatible, and that $\text{trace}_{H_1/H}(P_1) = 0$. Assume to simplify the exposition that K has unit group $\mathcal{O}_K^\times = \pm 1$ and class number 1, so that $H = K$, and that the group of connected components of the Néron model of E/K at the prime p is trivial. The prime p is totally ramified in H_n/K; let p_n be the unique prime of H_n over p, and let Φ_n be the group of connected components of E/K_n at the prime p_n; one has

$$\Phi_n = \mathbf{Z}/(p+1)\mathbf{Z} \times \mathbf{Z}/p^{n-1}\mathbf{Z}, \quad \Phi_\infty := \lim_{\leftarrow} \Phi_n = \mathbf{Z}/(p+1)\mathbf{Z} \times \mathbf{Z}_p,$$

where the inverse limit is taken with respect to the norm maps.

The main formula of [**?**] relates the image $\bar{P}_n$ of P_n in the group Φ_n to the special value $L(E/K, 1)$. The norm-compatible system of points P_n gives rise to a canonical Heegner element $P_\infty \in \lim_{\leftarrow} E(H_n)$, and hence to an element $\bar{P}_\infty$ in Φ_∞. As a corollary to the main result of [**?**] one obtains:

Theorem 3.4 *The element $\bar{P}_\infty$ is non-torsion if and only if $L(E/K, 1) \neq 0$.*

The calculations involved in the proof of theorem 3.4 are considerably simpler than those of [?] needed to prove theorem 3.3. The main ingredients in this proof are a formula of Gross for the special value $L(E/K,1)$ (generalized somewhat in [?]) and a moduli description due to Edixhoven for the specialization map to the group of connected components of $J_0(N)$. For more details, see [?].

A precursor of theorem 3.4 for Eisenstein quotients can be found in Mazur's article [?].

E. *p-adic analytic construction of Heegner points from derivatives of p-adic L-functions*: Assume for simplicity that E is semi-stable, and that, as before, $E/\mathbf{Q}$ has a prime p of multiplicative reduction which is inert in K, so that it is equipped with the analytic Tate parametrization

$$\Phi_{Tate} : K_p^\times \longrightarrow E(K_p),$$

where $K_p := K \otimes \mathbf{Q}_p$. Assume now that $L(E/K,s)$ has sign -1 in its functional equation. Let H_∞ be the compositum of all the ring class fields of K of conductor p^n, whose Galois group $G_\infty = \mathrm{Gal}(H_\infty/K)$ is canonically isomorphic to an extension of the class group $\Delta = \mathrm{Gal}(H/K)$ by the group $(K_p^\times)_1$ of elements in K_p of norm 1. By a generalization of the work of Gross [?] explained in [?], there exists an element $\mathcal{L}$ in the completed integral group ring $\mathbf{Z}[\![G_\infty]\!] := \lim_\leftarrow \mathbf{Z}[G_n]$ such that

$$|\chi(\mathcal{L})|^2 = \mathcal{L}(E/K,\chi,1)/\int\int_{E(\mathbf{C})} \omega \wedge \bar{\omega} \prod_{\ell | N^-} m_\ell \sqrt{D}, \qquad (13)$$

for all finite order characters $\chi : G_\infty \longrightarrow \mathbf{C}^\times$. The element $\mathcal{L}$ plays the role of the p-adic L-function associated to the anti-cyclotomic $\mathbf{Z}_p$-extension in this setting. (It really might be more accurate to view it as a*square root* of the p-adic L-function.)

Note that if χ_{triv} denotes the trivial character, then $\chi_{triv}(\mathcal{L}) = 0$, since $L(E/K,1) = 0$. Hence $\mathcal{L}$ belongs to the augmentation ideal I in the completed group ring $\mathbf{Z}[\![G_\infty]\!]$. Let $\mathcal{L}'$ be the natural projection of $\mathcal{L}$ in $I/I^2 = G_\infty$. One shows (cf. [?]) that $\mathcal{L}'$ belongs to $(K_p^\times)_1 \subset G_\infty$. The element $\mathcal{L}'$ in $K_p^\times$ should be viewed as the first derivative of the p-adic L-function of E/K (in the anticyclotomic direction, at the trivial character).

Let P_K be the Heegner point on $E(K)$ coming from the Shimura curve parametrization ϕ_{N^+,N^-} that was introduced in paragraph A of this section, and let $\bar{P}_K$ be its Galois conjugate. The following theorem is the main result of [**?**]:

Theorem 3.5 *Let w_p be a local sign which is -1 is $E/\mathbf{Q}_p$ has split multiplicative reduction, and 1 is $E/\mathbf{Q}_p$ has non-split multiplicative reduction. Then*

$$\Phi_{Tate}(\mathcal{L}') = \pm(P_K + w_p \bar{P}_K).$$

Note that, since $p|N^-$, the curve X_{N^+,N^-} is never a classical modular curve. Like the formula of Rubin described in paragraph C, theorem 3.5 allows one to recover a global point in $E(K)$ from the first derivative of a p-adic L-function.

The main ingredients in the proof of theorem 3.5 are the explicit construction of $\mathcal{L}$ given in [**?**] and [**?**] and the Cerednik-Drinfeld theory of p-adic uniformization of the Shimura curve X_{N^+,N^-} [**?**], [**?**], [**?**]. The details of the proof are given in [**?**].

Remarks:
1. The formulas described in paragraphs D and E were inspired by some fundamental ideas of Mazur, Tate, and Teitelbaum on p-adic analogues of the Birch and Swinnerton-Dyer conjecture. The connection with this circle of ideas is explained in [**?**].
2. There are many other generalizations of the Gross-Zagier formula which were not mentionned here because they are not directly relevant to modular elliptic curves: for example, the work of Nekovar [**?**] and Zhang [**?**] extending the work of Gross-Zagier and Kolyvagin to modular forms of higher weight, replacing Heegner points by higher-dimensional cycles on Kuga-Sato varieties.
3. In connection with the results described in paragraphs C and E, one should also mention an intriguing result of D. Ulmer [**?**] which constructs global points on certain universal elliptic curves over the function fields of modular curves in characteristic p. Some of the results described above (and, in particular, the formula of paragraph E) should extend to the function field setting; this extension has some tantalizing similarities, as well as differences, with Ulmer's constructions.

4 The Birch and Swinnerton-Dyer conjecture

4.1 Analytic rank 0

For modular elliptic curves of analytic rank 0, one has the following theorem.

Theorem 4.1 *If $L(E/\mathbf{Q},1) \neq 0$, then $E(\mathbf{Q})$ is finite, and so is $\underline{III}(E/\mathbf{Q})$.*

There are now several ways of proving this theorem. We will review the different strategies, giving only the briefest indication of the details of the proofs.

4.1.1 Kolyvagin's proof

It can be divided into three steps.

Step 1 (Non-vanishing lemma): Choose an auxiliary imaginary quadratic field $K/\mathbf{Q}$ such that

1. All primes dividing N are split in K.

2. Under assumption 1, the sign w_K is -1 and the L-function $L(E/K,s)$ necessarily vanishes at $s=1$. One requires in addition that the L-function $L(E/K,s)$ has only a simple zero, i.e., that $L'(E^{(D)}/\mathbf{Q},1) \neq 0$.

The existence of such a quadratic field K follows from the theorems of Bump-Friedberg-Hoffstein [**?**] and Murty-Murty [**?**] on non-vanishing of first derivatives of twists of automorphic L-series.

Step 2 (Gross-Zagier formula): Invoking the Gross-Zagier formula (theorem 3.3) one concludes that the Heegner point $P_K \in E(K)$ is of infinite order. In particular the rank of $E(K)$ is at least 1.

Step 3 (Kolyvagin's descent): In [**?**], Kolyvagin proves the following theorem:

Theorem 4.2 *If the Heegner point P_K is of infinite order, then $E(K)$ has rank 1 and $\underline{III}(E/K)$ is finite.*

Crucial to the proof of theorem 4.2 is the fact that the Heegner point P_K does not come alone. Namely, for each abelian extension L/K such that the Galois group $\mathrm{Gal}(L/\mathbf{Q})$ is dihedral, satisfying $\gcd(\mathrm{Disc}(L/K), ND) = 1$, there is a

Heegner point P_L in $E(L)$ and this system of points is norm-compatible in the sense that, if $L_1 \subset L_2$, then

$$\mathrm{trace}_{L_2/L_1} P_{L_2} = \ell(L_2/L_1) P_{L_1},$$

where $\ell(L_2/L_1) \in \mathbf{Z}[\mathrm{Gal}(L_1/K)]$ is an element whose definition involves the local Euler factors in $L(E/K, s)$ at the primes dividing $\mathrm{Disc}(L_2/L_1)$. Kummer theory allows one to construct Galois cohomology classes $c_L \in H^1(L, T_p(E))$ from the points P_L, where $T_p(E)$ is the p-adic Tate module of E. These classes satisfy the same trace-compatibility properties as the P_L. Kolyvagin calls such a system of cohomology classes an *Euler System* [**?**], and shows that if the "initial" class c_K is non-zero, the rank of $E(K)$ is less than or equal to 1 and $\underline{III}(E/K)$ is finite.

We will not go into the details of Kolyvagin's ingenious argument, referring the reader instead to [**?**] and [**?**] for more details.

4.1.2 A variant

The following variant of Kolyvagin's basic strategy avoids the non-vanishing result of Bump-Friedberg-Hoffstein and Murty-Murty, as well as the formula of Gross and Zagier. It only works, however, for elliptic curves having a prime p of multiplicative reduction, and does not prove the finiteness of $\underline{III}(E/\mathbf{Q})$, but only of the p-primary part of $\underline{III}(E/\mathbf{Q})$.

Step 1 (Non-vanishing lemma): Choose now an auxiliary imaginary quadratic field $K/\mathbf{Q}$ such that

1. The prime p is inert in K, and all the other primes dividing N are split in K.
2. By proposition 3.2, the L-function $L(E/K, s)$ has sign $w_K = 1$ in its functional equation. One requires also that $L(E^{(D)}/\mathbf{Q}, 1) \neq 0$, so that $L(E/K, 1) \neq 0$.

The existence of such a quadratic field K follows a theorem of Waldspurger [**?**] on non-vanishing of the values of twists of automorphic L-series.

Step 2 (A variant of the Gross-Zagier formula): Invoking theorem 3.4, one finds that the element P_∞ has non-trivial image in Φ_∞.

Step 3 (A variant of Kolyvagin's descent): In [**?**], the following theorem is proved:

Theorem 4.3 *If the image $\bar{P}_\infty$ of P_∞ in Φ_∞ is non-torsion, then $E(K)$ has rank* 0 *and $\underline{III}(E/K) \otimes \mathbf{Z}_p$ is finite.*

This theorem is proved by a minor adaptation of Kolyvagin's argument. The entire system of points P_n is now used to construct a cohomology class $c_K \in H^1(K, T_p(E))$, which is part of an Euler system. The non-vanishing of $\bar{P}_\infty$ translates into the non-triviality of the class c_K, and in fact of its image in a certain quotient (the "singular part") of the local cohomology group $H^1(K_p, T_p(E))$. Such a non-triviality is used to uniformly bound the p^n Selmer group of E/K, following the ideas of Kolyvagin.

The details of the argument are explained in [**?**].

4.1.3 Kato's proof

Recently Kato [**?**] has discovered a wholly original proof of theorem 4.1 which does not require the choice of an auxiliary imaginary quadratic field and does not use Heegner points.

Kato's argument constructs cohomology classes $c_L \in H^1(L, T_p(E))$, where L is a cyclotomic extensions of the rationals with discriminant prime to N. These classes are constructed from certain elements introduced by Beilinson, belonging to the K_2 of modular function fields. Defined via explicit modular units (Siegel units), these classes yield elements in $H^1(L, T_p(J_0(N)))$ which are mapped to $H^1(L, T_p(E))$ via the map ϕ of theorem 2.6. (In particular, theorem 2.6 is also crucial to Kato's construction.)

Kato's classes c_L obey norm-compatibility properties similar to those of Kolyvagin, and hence deserve to be viewed as an Euler system [**?**]. The most difficult part of Kato's argument, given in [**?**], is to relate the basic class $c_{\mathbf{Q}} \in H^1(\mathbf{Q}, T_p(E))$ (or rather, its localization in a certain quotient – the "singular part" – of the local cohomology group $H^1(\mathbf{Q}_p, T_p(E))$) to the special value $L(E/\mathbf{Q}, 1)$.

4.2 Analytic rank 1

In the case of analytic rank 1, there is:

Theorem 4.4 *Suppose that $L(E/\mathbf{Q},1)=0$ but that $L'(E/\mathbf{Q},1)\neq 0$. Then $E(\mathbf{Q})$ has rank 1, and $\underline{III}(E/\mathbf{Q})$ is finite.*

This theorem lies somewhat deeper than theorem 4.1. To prove it, one disposes only at present of the basic strategy of Kolyvagin based on the Gross-Zagier formula.

Step 1 (Non-vanishing lemma): Choose an auxiliary imaginary quadratic field $K/\mathbf{Q}$ such that

1. All primes dividing N are split in K.

2. The Hasse-Weil L-function $L(E/K,s)$ has a simple zero at $s=1$, so that $L(E^{(D)}/\mathbf{Q},1)\neq 0$.

The existence of such a quadratic field K follows from the same theorem of Waldspurger [**?**] on non-vanishing of values of twists of automorphic L-series used in step 1 of section 4.1.2.

Step 2 (Gross-Zagier formula): Invoking the Gross-Zagier formula (theorem 3.3) one finds that the Heegner point $P_K\in E(K)$ is of infinite order. In particular the rank of $E(K)$ is at least 1. More precisely, by analyzing the action of complex conjugation on P_K, one finds that P_K (up to torsion) actually belongs to $E(\mathbf{Q})$ in this case, so that the rank of $E(\mathbf{Q})$ is at least 1.

Step 3 (Kolyvagin's descent): By theorem 4.2, one concludes that $E(K)$ has rank 1 and finite Shafarevich-Tate group. Hence the rank of $E(\mathbf{Q})$ is exactly 1, its Shafarevich-tate group $\underline{III}(E/\mathbf{Q})$ is finite, and, as a by-product, $E^{(D)}(\mathbf{Q})$ and $\underline{III}(E^{(D)}/\mathbf{Q})$ are also finite.

Remark: When the sign in the functional equation for $L(E/\mathbf{Q},s)$ is -1, the class $c_{\mathbf{Q}}$ constructed by Kato gives rise to a natural element in the pro-p Selmer group of $E/\mathbf{Q}$, defined as the inverse limit $\lim_{\leftarrow}\mathrm{Sel}(\mathbf{Q},E_{p^n})$. One might expect that this class is non-zero if and only if $L'(E/\mathbf{Q},1)\neq 0$. A proof of this would show that

$$L'(E/\mathbf{Q},1)\neq 0\Rightarrow rank(E(\mathbf{Q}))\leq 1,$$

which represents a part of theorem 4.4. The reverse inequality seems harder to obtain with Kato's methods.