The Birch-Swinnerton-Dyer conjecture
from a naive point of view

by

Don B. Zagier

Max-Planck-Institut für Mathematik
Gottfried-Claren-Str. 26
D-5300 Bonn 3, FRG

Department of Mathematics
University of Maryland
College Park, MD 20742, USA

MPI/89-48

The Birch–Swinnerton-Dyer conjecture from a naive point of view

DON ZAGIER

Max-Planck-Institut für Mathematik, Bonn
and
University of Maryland

Throughout this paper, E will denote an elliptic curve defined over $\mathbf{Q}$ which we suppose given in $\mathbf{P}^2$ by an equation

$$f(x,y,z)=0 \qquad (f\in\mathbf{Z}[x,y,z] \text{ homogeneous of degree } 3) \tag{1}$$

with f of minimal discriminant Δ. If R is any ring with unit, then $E(R)$ denotes the set of solutions of $f=0$ in $\mathbf{P}^2(R)=\{(x,y,z)\in R^3 \,:\, xR+yR+zR=R\}/R^\times$. (In particular, $E(\mathbf{Z})$ is the same as the Mordell-Weil group $E(\mathbf{Q})$ and not, as sometimes in the literature, the finite set of integral points in the affine model $f(x,y,1)=0$ of E over $\mathbf{Z}$.) The L-series of E is the Dirichlet series given by

$$L(E,s)=\prod_{p \text{ prime}}\frac{1}{1-a(p)p^{-s}+\varepsilon(p)p^{1-2s}}=\sum_{n=1}^{\infty}\frac{a(n)}{n^s} \qquad (\mathrm{Re}(s)>\frac{3}{2}), \tag{2}$$

where $a(p)=p+1-|E(\mathbf{Z}/p\mathbf{Z})|$ ($|A|$ denotes the cardinality of a set A) and $\varepsilon(p)=1$ or 0 depending whether $p\nmid\Delta$ or $p|\Delta$. The **Birch–Swinnerton-Dyer conjecture** consists of the two statements

(A) $L(E,s)$ continues meromorphically to $s=1$ and has a zero there of order exactly $r=\mathrm{rank}_{\mathbf{Z}}E(\mathbf{Q})$ (for $r=0$ this means that $L(E,s)$ is holomorphic and non-zero at $s=1$).

(B) Assuming this, define λ by $L(E,s)\sim\lambda\,(s-1)^r \quad (s\to 1)$. Then

$$\lambda=\frac{c\,R\,\Omega}{T^2}\,|\text{Ш}|, \tag{3}$$

where R is the determinant of the Néron-Tate height pairing with respect to a basis of the free part of $E(\mathbf{Q})$, Ω is the real period (= integral over $E(\mathbf{R})$ of a Néron differential of $E/\mathbf{Z}$), $c=\prod_{p|\Delta}c_p$ where $c_p=[E(\mathbf{Z}_p):E^0(\mathbf{Z}_p)]$ ($E^0(\mathbf{Z}_p)$ is the set of points whose reduction (mod p) belong to the non-singular part of $E(\mathbf{Z}/p\mathbf{Z})$), and Ш is the conjecturally finite Tate-Shafarevich group of E. (Each of these quantities will be discussed in more detail later.)

The Birch–Swinnerton-Dyer conjecture is very famous and has been given various formulations in more learned language, most strikingly one in terms of Tamagawa numbers in a beautiful paper by Spencer Bloch (*Invent. math.* **58** (1980), 65-76). Our purpose in this note is to go the other way, replacing as many as possible of the invariants entering into (3) by numbers that can be defined without any theoretical knowledge (in particular, without knowing what the height pairing is, how the groups $E(\mathbf{Z}_p)$ and $E^0(\mathbf{Z}_p)$ look, or even that there is a group structure on $E(R)$). More precisely, we will try to define these invariants in a purely Diophantine way, merely by *counting* solutions of (1) in $\mathbf{Q}$ and in $\mathbf{Z}/n\mathbf{Z}$. The motivation for doing this, besides the pleasure in finding a formulation of the Birch–Swinnerton-Dyer conjecture which can be explained easily to a non-specialist, comes from the analogy with the situation of zeta-functions of quadratic fields, where by defining everything in terms of counting solutios of equations one can get an easy proof of the Dirichlet class number formula. This analogy is described in §1. In §2 we interpret the numbers r and RT^{-2} (and, very vaguely, $RT^{-2}\Omega$) in terms of counting rational solutions of (1). In §3 we express the quantity $c^{-1}\lambda$ (assuming **(A)**) in terms of the number of solutions of (1) in $\mathbf{Z}/n\mathbf{Z}$ for all n. Finally, in §4 we make some remarks about the interpretation of Ш in terms of solutions of (1) and related equations over various other rings.

1. The prototype: zeta functions of quadratic fields

There is a well-known analogy between the Birch–Swinnerton-Dyer conjecture and Dirichlet's class number formula for number fields in which $\sqrt{R}$ corresponds to the regulator, T to the number of roots of unity, and Ш to the class group of the field. Like the Birch–Swinnerton-Dyer conjecture, Dirichlet's formula can be stated in a much more abstract language than the original formulation. However, it can also be stated—and, at least for quadratic fields, proved—in an entirely elementary way purely in terms of counting solutions of equations. In this section we describe how this goes, since it is the model for what we would like to do for equation (3).

Let D be an integer and consider the set of all binary quadratic forms $Q(u,v) = au^2 + buv + cv^2$ with a, b, $c \in \mathbf{Z}$ and discriminant $b^2 - 4ac = D$. The group $SL(2,\mathbf{Z})$ acts on such forms; we denote by $[Q]$ and by $\mathrm{Aut}(Q)$ the orbit and isotropy group, respectively, of Q under this action. If n is a natural number, then $\mathrm{Aut}(Q)$ acts on the set of representations of n by Q ($=$ pairs $(u,v) \in \mathbf{Z}^2$ with $Q(u,v) = n$). We denote by $r_Q(n)$ the number of inequivalent representations under this action (this can be an integer ≥ 0 or, *a priori*, ∞) and by $r_Q^*(n)$ the number of inequivalent primitive representations (representations with u and v coprime). Then the basic identity is

$$\sum_{[Q]} r_Q^*(n) = \left|\{b \ (\mathrm{mod}\ 2n) \, : \, b^2 \equiv D \ (\mathrm{mod}\ 4n)\}\right| \tag{4}$$

(in particular , $r_Q^*(n)$ is finite, and in fact bounded by n).

We will prove (4) by a very simple counting argument which works in a uniform way for all D, positive, negative, or zero. First, however, we show how it implies the analogue of the Birch–Swinnerton-Dyer formula (3) in the case that D is the discriminant of a quadratic field (i.e. $D \neq 1$ is square-free and congruent to 1 modulo 4 or $D/4$ is square-free and congruent to 2 or 3 modulo 4). The right-hand side of (4) is clearly a multiplicative function of n. If $(\frac{D}{p}) = 1$ (or $p = 2$, $D \equiv 1 \pmod 8$), it is 2 for all $n = p^\nu > 1$; if $(\frac{D}{p}) = -1$ (or $p = 2$, $D \equiv 5 \pmod 8$), it is 0 for all $n = p^\nu > 1$; and if $p|D$ it is 1 for $n = p$ and 0 for $n = p^\nu$, $\nu > 1$. Using this and the obvious identity $r_Q(n) = \sum_{e^2|n} r^*_Q(n/e^2)$, we easily deduce from (4) the equivalent formula

$$\sum_{[Q]} r_Q(n) = \sum_{d|n} \chi_D(d), \tag{5}$$

where χ_D is defined as the totally multiplicative function on $\mathbb{N}$ with $\chi_D(p) = (\frac{D}{p})$ for p prime. By quadratic reciprocity one knows that χ_D is a periodic function of average 0; it follows that the series $L(1, \chi_D) = \sum \chi_D(n)\, n^{-1}$ converges. Taking the average $(= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (\cdots))$ of both sides of (5), we obtain

$$\sum_{[Q]} \langle r_Q \rangle = L(1, \chi_D). \tag{6}$$

But it is easily seen by counting lattice points in a sector of an ellipse or hyperbola that the average value $\langle r_Q \rangle$ of r_Q has a value $\kappa(D)$ independent of Q, given by $\frac{2\log\varepsilon}{\sqrt{D}}$ if $D > 0$ (where $\varepsilon = \frac{1}{2}(t + u\sqrt{D})$ for the smallest positive integer solution of Pell's equation $t^2 - Du^2 = \pm 4$ if such a solution exists, $\kappa(D) = \infty$ otherwise) and by $\frac{2\pi}{w\sqrt{-D}}$ if $D < 0$ (where $w = 2$, 4 or 6 is the number of integral solutions of $t^2 - Du^2 = 4$). Thus (6) says that $L(1, \chi_D) = h(D)\kappa(D)$, where $h(D)$ is the (possibly infinite) number of equivalence classes $[Q]$ of discriminant D. Since $L(1, \chi_D)$ is *a priori* known to be finite (but might be zero), while $h(D)$ and $\kappa(D)$ are *a priori* known to be non-zero (but might be infinite), this gives in one blow the finiteness of the class number $h(D)$, the existence of a solution of Pell's equation for non-square $D > 0$, the non-vanishing of $L(1, \chi_D)$, and the Dirichlet class number formula.

And now to the proof of (4), which is considerably shorter than the discussion of what to do with it. We use the following general principle: if a group G acts on two sets X and Y, and $S \subset X \times Y$ is a subset invariant under the diagonal action of G, then

$$\sum_{x \in X/G} |S_x/G_x| = \sum_{y \in Y/G} |S_y/G_y|, \tag{7}$$

where $S_x = \{y \in Y : (x, y) \in S\}$, $G_x = \{g \in G : gx = x\}$ and similarly for S_y and G_y. The proof of (7) is obvious: just count the number of orbits of G in S in the two possible orders. We apply it to the case where G is $SL(2,\mathbf{Z})$, X the set of quadratic forms Q of discriminant D, Y the set of pairs of coprime integers (u,v), and S the set of pairs $(Q,(u,v)) \in X \times Y$ with $Q(x,y) = n$. Then the left-hand side of (7) is equal to the left-hand side of (4) by definition. On the other hand, Y/G consists of a single element which we can represent by the point $y = (1,0)$. For this choice, S_y is the set of quadratic forms $nu^2 + buv + \frac{b^2-D}{4n}v^2$ with $b^2 \equiv D \pmod{4n}$ and $G_y = \{\left(\begin{smallmatrix}1 & r\\ 0 & 1\end{smallmatrix}\right) : r \in \mathbf{Z}\}$ acts on this by $b \mapsto b + 2rn$. The result follows.

In summary, the identity (4), which can be written in the equivalent form

$$\zeta(2s)^{-1} \sum_{[Q]} \sum_{\substack{(u,v)\in\mathbf{Z}^2/\mathrm{Aut}(Q)\\ Q(u,v)>0}} \frac{1}{Q(u,v)^s} = \sum_{n=1}^{\infty} \frac{|\{b\,(2n) \,:\, b^2 \equiv D\,(4n)\}|}{n^s}, \tag{8}$$

has a simple combinatorial interpretation in terms of counting orbits under a certain group action on the set of integral solutions of a certain Diophantine equation, and gives the Dirichlet class number formula and other arithmetical information by looking at the asymptotic behavior near $s = 1$. If we could interpret the Birch–Swinnerton-Dyer conjecture in the same way, we would obtain a proof of it. This of course we cannot do, but we will at least be able to write some of the invariants it involves in terms of the asymptotics of two Dirichlet series analogous to those occurring in equation (8).

2. The global Diophantine invariants r and RT^{-2} (and Ω)

Let $|\ \ |$ be any continuous norm on $\mathbf{R}^3$, e.g. $|(x,y,z)| = (x^2+y^2+z^2)^{1/2}$ or $|x|+|y|+|z|$ or $\max\{|x|,|y|,|z|\}$. We can consider $|\ \ |$ as defined on $\mathbf{P}^2(\mathbf{Q})$ or $E(\mathbf{Q})$ by identifying these sets with $\mathbf{P}^2(\mathbf{Z}) = \mathbf{Z}^3/\{\pm1\} \subset \mathbf{R}^3/\{\pm1\}$ and $E(\mathbf{Z})$, respectively. Now let

$$\mathcal{N}(B) = \{P \in E(\mathbf{Q}) \,:\, |P| \le B\}$$

be the number of solutions of (1) in coprime integers (x,y,z) of norm $\le B$, two solutions differing only in sign being counted as one. Then we have

Proposition. *The asymptotic growth of $\mathcal{N}(B)$ as $B \to \infty$ is given by*

$$\mathcal{N}(B) \sim \frac{\pi^{r/2}}{(r/2)!}\,\frac{T}{\sqrt{R}}\,(\log B)^{r/2} \qquad (B \to \infty). \tag{9}$$

Here $(r/2)!$ is to be interpreted as $\Gamma(1+\frac{r}{2})$ if r is odd. Note that the assertion of (9) is independent of the choice of norm because any two norms are bounded by multiples of one another and $(\log B + \mathrm{O}(1))^{r/2} \sim (\log B)^{r/2}$.

Proof: We recall how the Néron-Tate height pairing is defined: one shows that there exists a positive-definite quadratic form h on $E(\mathbf{Q})/(\text{torsion})$ such that

$$h(P) = \log|P| + \mathrm{O}(1) \qquad \text{for all } P \in E(\mathbf{Q}); \tag{10}$$

the height pairing is the associated non-degenerate bilinear form. Thus we can consider $E(\mathbf{Q})/(\text{torsion}) \approx \mathbf{Z}^r$ as a lattice Λ in the vector space $E(\mathbf{Q}) \otimes \mathbf{R} \approx \mathbf{R}^r$ with the metric defined by h. The number of points in Λ with $h \leq H$ for H large is asymptotically equal to $\frac{1}{\sqrt{R}}$ (the volume of $\mathbf{R}^r/\Lambda$) times $\frac{(\pi H)^{r/2}}{(r/2)!}$ (the volume of an r-dimensional sphere of radius $\sqrt{H}$), and the number of $P \in E(\mathbf{Q})$ with $h(P) \leq H$ is T times this. The result now follows from equation (10).

Notice that (9) implies that the zeta function

$$Z(E,s) = \sum_{P \in E(\mathbf{Q})} \frac{1}{|P|^s}$$

converges for all s with $\mathrm{Re}(s) > 0$ and satisfies

$$Z(E,s) \sim \frac{T}{\sqrt{R}} (s/\pi)^{-r/2} \qquad (s \searrow 0) \tag{11}$$

(to see this, write $Z(E,s)$ as $s \int_0^\infty \mathcal{N}(B)\, B^{-s-1}\, dB$ and use the estimate on $\mathcal{N}(B)$). Either (9) or (11) defines both r and R/T^2 in an elementary way, without reference to the group structure or the Néron-Tate height pairing on $E(\mathbf{Q})$.

Finally, we make a stab at bringing Ω into the picture. The presence of the factor $\sqrt{R}$ in (11), as well as the fact that $Z(E,s)$ has a pole of half-integral order at 0 if r is odd, suggests looking at $Z(E,s)^2$, which is the sum of $|P|^{-s}|Q|^{-s}$ over all pairs $(P,Q) \in E(\mathbf{Q})^2$. We could look instead at the subsum over pairs of points (P,Q) which are close in the real topoogy, say $|z(P-Q)/x(P-Q)| < \epsilon$ in the standard Tate-Weierstrass form $y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz^2 + a_6z^3$. These points, at least if $E(\mathbf{Q})$ is infinite, will constitute asymptotically a proportion ν of the set of all pairs, where ν is the ratio of the length of $\{(x:y:z) \in E(\mathbf{R}) : |z/x| < \epsilon\}$ to that of all of $E(\mathbf{R})$. The first of these lengths is given by an incomplete elliptic integral which is asymptotic to $2\sqrt{\epsilon}$ as $\epsilon \to 0$, while the second, a complete elliptic integral, equals Ω. Thus the two-point zeta function in question looks like $\frac{2\sqrt{\epsilon}}{\Omega} Z(E,s)^2$, and its leading terms at 0 like $2\sqrt{\epsilon}\,\frac{T^2}{\Omega R}(\pi/s)^r$. Obviously this is very vague, but at least it suggests that the natural combination of $T/\sqrt{R}$ and Ω is $T^2/\Omega R$, the quantity occurring in the Birch–Swinnerton-Dyer conjecture.

3. The local Diophantine invariants $L(E,s)$ and c

The L-function of E defined by (2), while natural from several points of view (in particular, the conjectured holomorphic continuation and functional equation), has several defects from a purely Diophantine point of view:

(i) The numbers $a(p)$ are defined in terms of counting solutions, but the coefficients $a(n)$ for n composite (say $n = p^2$ or $n = pq$) can only be defined by using the recursion given by the Euler product (2); we do not know how to compute $a(n)$ directly in terms of numbers of solutions of equations, or without knowing the prime factorization of n.

(ii) The function $L(E,s)$ ought to summarize all the local data about E which is relevant for the Birch–Swinnerton-Dyer conjecture, but fails to do so: at primes of bad reduction the Euler factor of $L(E,s)$ contains too little information and we are obliged to include the number c_p as an extra ("fudge") factor in (3).

(iii) $L(E,s)$ conjecturally has a zero of order r at $s=1$, but a pole would be much more convenient, since the presence of a pole in the analytic continuation of a Dirichlet series makes itself felt in the asymptotics of its coefficients, while a zero is not visible in this way.

In this section we will attempt to remedy or partially remedy these defects by introducing a new Dirichlet series which is related to, but not expressible in terms of, $L(E,s)$. Actually, we shall define three such Dirichlet series, but we give only the most natural one here, and mention the others as variants at the end of the section. We denote by $N(n)$ the cardinality of the finite set $E(\mathbf{Z}/n\mathbf{Z})$ and set

$$D(E,s)=\sum_{n=1}^{\infty} N(n)\,n^{-s}.$$

The series converges absolutely for $\mathrm{Re}(s)>2$. We will prove:

PROPOSITION. *The product $D(E,s)\,L(E,s)$ extends meromorphically to the half-plane $Re(s)>\frac{5}{6}$, is holomorphic for $Re(s)\geq 1$ except for a simple pole at $s=2$, and has the value -1 at $s=1$.*

COROLLARY. *Part* **(A)** *of the Birch–Swinnerton-Dyer conjecture is equivalent to*

(A′) *$D(E,s)$ continues meromorphically to $s=1$ and has a pole there of order exactly r (for $r=0$ this means that $D(E,s)$ is holomorphic and non-zero at $s=1$.)*

If this holds, then $D(E,s)\sim -\lambda^{-1}c(s-1)^{-r}$ as $s\to 1$.

Thus $D(E,s)$ remedies the problems (i) ($N(n)$ is defined directly by counting solutions of (1) modulo n, without reference to the prime factorization of n) and (ii) (the leading term of $D(E,s)$ at $s=1$ involves the same combination of λ and c as occurs in the Birch–Swinnerton-Dyer conjecture, so $D(E,s)$ encodes all the interesting local information). It also partially solves (iii), since "rth order zero" has been replaced by "rth order pole" in going from **(A)** to **(A′)**. However, if $E(\mathbf{Q})$ is finite, then there is no pole, and even if $r>0$ the effect of the pole at $s=2$ (to say nothing of possible other poles in $\mathrm{Re}(s)>1$ if the Riemann hypothesis is false for $L(E,s)$) will dominate the asymptotics of the coefficients $N(n)$ and tend to swamp the contribution of the pole at $s=1$, making it hard to "see" the numbes r and $\lambda^{-1}c$ (the pole at $s=2$ will contribute a term Ax^2 to $\sum_{n\leq x} N(n)$, while the pole at $s=1$—if r is positive—will give a smaller order contribution $-\lambda^{-1}cx(\log x)^{r-1}/(r-1)!$). Also on the negative side, of course, is that $D(E,s)$ does not even conjecturally have an analytic continuation to all s or satisfy a functional equation.

To prove the proposition, we observe that $N(n)$ is clearly multiplicative and hence that $D(E,s)$ has an Euler product

$$D(E,s) = \prod_p D_p(p^{-s}), \qquad D_p(x) = \sum_{\nu=0}^{\infty} N(p^\nu)x^\nu.$$

We look at the Euler factors separately, starting with the case of good reduction, $p \nmid \Delta$. By definition, $N(p) = |E(\mathbf{Z}/p\mathbf{Z})| = p + 1 - a(p)$. Since the points of $E(\mathbf{Z}/p\mathbf{Z})$ are non-singular (i.e., some partial derivative of f is non-zero mod p), Hensel's lemma implies that they each lift to exactly $p^{2(\nu-1)}$ solutions of $f = 0$ in $(\mathbf{Z}/p^\nu\mathbf{Z})^3$ for $\nu \geq 1$. But the order of $(\mathbf{Z}/p^\nu\mathbf{Z})^\times$ is $p^{\nu-1}$ times that of $(\mathbf{Z}/p\mathbf{Z})^\times$, so this gives $N(p^\nu) = p^{\nu-1}N(p)$ for the number of solutions in $\mathbf{P}^2(\mathbf{Z}/p^\nu\mathbf{Z})$. Hence

$$D_p(x) = 1 + \sum_{\nu=1}^{\infty} \bigl(p + 1 - a(p)\bigr)\, p^{\nu-1}x^\nu = \frac{1 - \bigl(a(p)-1\bigr)x}{1 - px}.$$

It follows that the Euler factor of $D(E,s)/\zeta(s-1)$ at p is $1 - \bigl(a(p)-1\bigr)p^{-s}$, which for $s = 1$ takes on the value $\dfrac{N(p)}{p}$, the reciprocal of the value of the corresponding Euler factor of $L(E,s)$. Thus the Euler product of $D(E,s)L(E,s)/\zeta(s-1)$ converges (indeed, terminates) at $s = 1$. Nevertheless, this is not the right combination to look at, since it diverges for s near 1 (its pth Euler factor for $p \nmid \Delta$ equals $1 + p^{-s}(1 + \mathrm{o}(1))$ for $\mathrm{Re}(s) > \frac{1}{2}$, $s \neq 1$). Instead, we set

$$\psi_p(x) \overset{\text{def}}{=} \frac{(1-px)(1-x)}{1-px^2} D_p(x)L_p(x)$$

(where $L_p(p^{-s})$ is the pth Euler factor of $L(E,s)$) and find for $p \nmid \Delta$

$$\psi_p(x) = \frac{(1-x)\bigl(1-(a(p)-1)x\bigr)}{\bigl(1-px^2\bigr)\bigl(1-a(p)x+px^2\bigr)} = 1 + \frac{x^2(1-px)\bigl(a(p)-1+px\bigr)}{\bigl(1-px^2\bigr)\bigl(1-a(p)x+px^2\bigr)}.$$

This equals 1 at $x = p^{-1}$ and $1 + \mathrm{O}(p^{3/2}x^3)$ for $x = \mathrm{o}(p^{-1/2})$ (since $a(p) = \mathrm{O}(p^{1/2})$), so the product $\prod_{p \nmid \Delta} \psi_p(p^{-s})$ converges absolutely in $\mathrm{Re}(s) > \frac{5}{6}$ and equals 1 at $s = 1$.

We will show that $\psi_p(x)$ for $p|\Delta$ is a rational function having no poles in $|x| \leq \dfrac{1}{p}$ and satisfying $\psi_p\bigl(\frac{1}{p}\bigr) = c_p$. It follows that $\psi(s) = \prod_p \psi_p(p^{-s})$ extends to a non-zero holomorphic function in $\mathrm{Re}(s) > \dfrac{5}{6}$ with $\psi(1) = c$. This will prove the proposition, since $D(E,s)L(E,s) = \dfrac{\zeta(s)\zeta(s-1)}{\zeta(2s-1)}\,\psi(s)$ and $\dfrac{\zeta(s)\zeta(s-1)}{\zeta(2s-1)}$ is a meromorphic function with a simple pole at $s = 2$ and no other pole in $\mathrm{Re}(s) \geq 1$, equal to -1 at $s = 1$. So let p be a prime of bad reduction. If $E^0(\mathbf{Z}/p^\nu\mathbf{Z})$ denotes the set mapping to the non-singular part of

$E(\mathbf{Z}/p\mathbf{Z})$, then $|E^0(\mathbf{Z}/p\mathbf{Z})| = |E(\mathbf{Z}/p\mathbf{Z})| - 1 = p - a(p)$ and $|E^0(\mathbf{Z}/p^\nu\mathbf{Z})| = p^{\nu-1}|E(\mathbf{Z}/p\mathbf{Z})|$ by Hensel's lemma as before. But $|E(\mathbf{Z}/p^\nu\mathbf{Z})| \sim c_p|E^0(\mathbf{Z}/p^\nu\mathbf{Z})|$ for $\nu \to \infty$ by the definition of c_p, so $N(p^\nu) = c_p\big(p - a(p)\big)p^{\nu-1} + \mathrm{o}(p^\nu)$. This shows that $D_p(x)$, which is a rational function by general principles, has a simple pole of principal part $c_p\dfrac{1 - a(p)/p}{1 - px}$ at $x = \dfrac{1}{p}$ and no other poles in $|x| \le \dfrac{1}{p}$. Since $L_p(x) = \dfrac{1}{1 - a(p)x}$ for $p|\Delta$,the assertions about $\psi_p(x)$ follow. This completes the proof of the proposition.

Remark. The function $D(E,s)$ cannot be expressed as a quotient of finite products of functions $\zeta(ns - m)$ and $L(E, ns - m)$ ($n \in \mathbf{N}$, $m \in \mathbf{Z}$)), because the rational function $1 - (a-1)x$ cannot be factored into functions of the form $1 - p^m x^n$ and $1 - ap^m x^n + p^{2m+1}x^{2n}$. On the other hand, we could contine the partial factorization further and get the meromorphic continuation of $D(E,s)$ (assuming that of $L(E,s)$) into a bigger half-plane. For instance, the identity

$$\frac{(1-x)(1-px)}{(1-px^2)(1-a(p)x+px^2)(1-pa(p)x^3+p^3x^6)}\,D_p(x)$$
$$= 1 + \frac{a(p)x^2(1-p^4x^7) - x^2(1-p^5x^8) + (p^2 - pa(p)^2)x^4(1-px^2)}{(1-px^2)(1-a(p)x+px^2)(1-pa(p)x^3+p^3x^6)} \qquad (p\nmid\Delta)$$

gives the meromorphic continuation of $D(E,s)$ to the half-plane $\Re(s) > \dfrac{3}{4}$ (assuming that $L(E,s)$ is known to be meromorphic), because the expression on the right is $1 + \mathrm{O}(p^{-1-2\epsilon})$ for $x = \mathrm{O}(p^{-3/4-\epsilon})$. However, it seems unlikely that $D(E,s)$ continues meromorphically to the whole plane, and anyway we are mainly interested in the point $s = 1$.

Finally, we define the two variants of $D(E,s)$ mentioned at the beginning of the section. The definition of $N(n)$ can be written

$$N(n) = |\{(x,y,z) \in (\mathbf{Z}/n\mathbf{Z})^3_0 \; : \; f(x,y,z) \equiv 0 \pmod{n}\}/(\mathbf{Z}/n\mathbf{Z})^\times|$$

where $(\mathbf{Z}/n\mathbf{Z})^3_0$ denotes the set of triples $(x,y,z) \in (\mathbf{Z}/n\mathbf{Z})^3$ with $\gcd(x,y,z,n) = 1$. The new Dirichlet series are $D_j(E,s) = \sum N_j(n)\,n^{-s}$ ($j = 1$, 2), where

$$N_1(n) = |\{(x,y,z) \in (\mathbf{Z}/n\mathbf{Z})^3_0 \; : \; f(x,y,z) \equiv 0 \pmod{n}\}|,$$
$$N_2(n) = |\{(x,y,z) \in (\mathbf{Z}/n\mathbf{Z})^3 \; : \; f(x,y,z) \equiv 0 \pmod{n}\}|,$$

i.e., we count all coprime, or all, solutions modulo n, rather than only non-proportional ones. Since $(\mathbf{Z}/n\mathbf{Z})^\times$ acts freely on $(\mathbf{Z}/n\mathbf{Z})^3_0$, $N_1(n) = N(n)\varphi(n)$, where $\varphi(n) = |(\mathbf{Z}/n\mathbf{Z})^\times|$ is Euler's function. Hence for $p \nmid \Delta$ we have

$$\begin{aligned} D_{1,p}(x) \overset{\text{def}}{=} \sum_{\nu=0}^{\infty} N_1(p^\nu)\,x^\nu \; &= \; 1 + \sum_{\nu=1}^{\infty}\big(p + 1 - a(p)\big)(p-1)\,p^{2\nu-2}\,x^\nu \\ &= \frac{1 \,-\, \big(a(p)(p-1)+1\big)x}{1 \,-\, p^2x} \\ &= \frac{\big(1 - p^3x^2\big)\big(1 - a(p)px + p^3x^2\big)}{1 \,-\, p^2x}\psi_{1,p}(x) \end{aligned}$$

with

$$\psi_{1,p}(x) = 1 + x\frac{a(p)(1-p^4x^2) - (1-p^6x^3)}{(1-p^3x^2)(1-a(p)px+p^3x^2)},$$

which is equal to 1 at $x = p^{-2}$ and to $1 + \mathrm{O}(p^{9/2}x^3)$ for $x = \mathrm{o}(p^{-3/2})$. An argument like the one before shows that $\psi_{1,p}(x) = (1-p^2x)L_p(px)D_{1,p}(x)/(1-p^3x^2)$ for $p|\Delta$ is a rational function which is holomorphic in $|x| \le p^{-2}$ and equal to c_p at $x = p^{-2}$. Thus

$$D_1(s) = \frac{\zeta(s-2)}{\zeta(2s-3)L(E,s-1)}\prod_p \psi_{1,p}(p^{-s})$$

has a meromorphic continuation to $\mathrm{Re}(s) > \frac{11}{6}$ with a simple pole at $s = 3$ and a leading term $-\lambda^{-1}c(s-2)^{1-r}$ at $s = 2$. This is even worse than for $D(s)$ because now we get an actual pole only if the rank r is at least 2. For $D_2(s)$ the calculations are messier since $(\mathbf{Z}/n\mathbf{Z})^\times$ does not act freely on $(\mathbf{Z}/n\mathbf{Z})^3$ and we have to take into account the various isotropy groups. We find for $p \nmid \Delta$

$$N_2(p^\nu) = 1 + \sum_{\frac{\nu}{3} \le \mu < \nu} p^{3(\nu-\mu-1)}(p^3-1) + \sum_{0 \le \mu < \frac{\nu}{3}} p^{3\nu-2}(p-1)(p+1-a(p))$$

(to get this, count the number of solutions with $\gcd(x,y,z,p^\nu) = p^\mu$, $0 \le \mu \le \nu$),

$$\begin{aligned} D_{2,p}(x) &= \frac{1+x+p^3x^2}{1-p^6x^3} + \frac{(p+1-a(p))(p-1)x}{(1-p^2x)(1-p^6x^3)} \\ &= \frac{(1-p^5x^3)(1-p\,a(p)x+p^3x^2)}{(1-p^2x)(1-p^6x^3)}\left(1 + x\frac{(a(p)-p^2x)(1-p^6x^3)}{(1-p^5x^3)(1-pa(p)x+p^3x^2)}\right), \end{aligned}$$

and hence (treating the primes $p|\Delta$ as before) finally

$$D_2(s) = \frac{\zeta(s-2)\zeta(3s-6)}{\zeta(3s-2)L(E,s-1)}\psi_2(s)$$

with $\psi_2(s)$ holomorphic and non-zero in $\mathrm{Re}(s) > \frac{15}{8}$ and $\psi_2(2) = c$. Thus $D_2(s)$ has a meromorphic continuation to $\mathrm{Re}(s) > \frac{15}{8}$ with a simple pole at $s = 3$ and leading term $\frac{3c}{4\lambda}(s-2)^{1-r}$ at $s = 2$.

4. The Tate-Shafarevich group

Since the conjecturally finite order of the Tate-Shafarevich group Ш of $E/\mathbf{Q}$ enters into the Birch–Swinnerton-Dyer formula, we would like to relate it to Diophantine properties of E, i.e., to find a connection between the groups $E(\mathbf{Q})$ and Ш. The standard descent sequence

$$0 \longrightarrow E(\mathbf{Q})/mE(\mathbf{Q}) \longrightarrow \mathrm{Sel}_m \longrightarrow \text{Ш}[m] \longrightarrow 0 \qquad (12)$$

relates the m-torsion in Ш to the cokernel of $E(\mathbf{Q}) \xrightarrow{\cdot m} E(\mathbf{Q})$ for each natural number m, and it is reasonable to ask whether this sequence "lifts" to an exact sequence independent of m. One could try to construct a sequence $0 \to E(\mathbf{Q}) \to \text{Sel} \to Ш \to 0$ for some group "Sel", but a little thought show that this cannot be done in a natural way and anyway would not induce a sequence as in (12). Instead, we should look for a *four-term* sequence

$$0 \longrightarrow E(\mathbf{Q}) \longrightarrow \mathcal{E} \longrightarrow \mathcal{S} \longrightarrow Ш \longrightarrow 0. \tag{13}$$

This is motivated by two considerations:

1. An exact sequence of abelian groups

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{f} C \xrightarrow{p} D \longrightarrow 0$$

induces a short exact sequence

$$0 \longrightarrow A/mA \longrightarrow S_m \longrightarrow D[m] \longrightarrow 0$$

for every natural number m, where

$$S_m = \frac{\{(b,c) \in B \times C \mid f(b) = mc\}}{\{(mb, f(b)) \mid b \in B\}} \tag{14}$$

and the maps $A/mA \to S_m$ and $S_m \to D[m]$ are induced by $a \mapsto (i(a), 0)$ and $(b,c) \mapsto p(c) + mD$, respectively. (The proof is an easy diagram chase.)

2. The analogy between the Birch–Swinnerton-Dyer conjecture for elliptic curves and the Dirichlet class number formula for number fields K makes $E(\mathbf{Q})$ correspond to the unit group U_K and Ш to the class group $\mathcal{C}_K$; these are related by a four-term exact sequence

$$0 \longrightarrow U_K \longrightarrow K^\times \xrightarrow{f} \mathcal{I}_K \longrightarrow \mathcal{C}_K \longrightarrow 0,$$

where $\mathcal{I}_K$ is the group of fractional ideals of K and f the map associating to each number in K the ideal it generates.

To look for an exact sequence as in (13) we must first choose a good definition of Ш. The cohomological definition of Ш can be translated in a well-known way into a definition closer to the Diophantine properties of $E/\mathbf{Q}$: Ш is the set of isomorphism classes of curves C of genus 1, defined over $\mathbf{Q}$, having a point over $\mathbf{Q}_v$ for every place v of $\mathbf{Q}$, and equipped with an action of E making them into principal homogeneous spaces over E (or equivalently, with an isomorphism defined over $\mathbf{Q}$ between the Jacobian of C and E). The isomorphisms defining the classes are required to be defined over $\mathbf{Q}$ and compatible with the E-actions (or with the isomorphisms Jac(C) $\xrightarrow{\sim} E$), but not with the choices of $\mathbf{Q}_v$-rational points. The neutral element is the class of E and the sum of classes $[C']$ and $[C'']$ is the class of the curve $C = (C' \times C'')/E$, where E acts on $C' \times C''$ by $(c', c'') + e = (c' + e, c'' - e)$ and on C by $[c', c''] + e = [c' + e, c'']$.

Now let $\mathcal{E} = E(\overline{\mathbf{Q}})$, the set of $\overline{\mathbf{Q}}$-rational points on E, and $\mathcal{S}$ be the set of equivalence classes of pairs (C, P) with C as above and P a point of $C(\overline{\mathbf{Q}})$, the equivalence being given by $[C,P]=[C',P']$ if there is an E-equivariant isomorphism $C\to C'$ defined over $\mathbf{Q}$ and mapping P to P'. Both sets have natural group laws and there are obvious homomorphisms $E(\mathbf{Q})\hookrightarrow\mathcal{E}$, $\mathcal{S}\twoheadrightarrow\text{Ш}$ and $\mathcal{E}\to\mathcal{S}$, the last sending $P\in E(\overline{\mathbf{Q}})$ to the class of (E,P). It is not hard to see that with these definitions the sequence (13) is exact.

This gives us one way to realize (13), but it is not completely satisfactory because the introduction of $\overline{\mathbf{Q}}$, suggested by the original definition of Ш in terms of Galois cohomology, takes us further away from the Diophantine properties of E. However, we could have taken instead of $\overline{\mathbf{Q}}$ any ring R satisfying

(i) R has a unit element 1 and the map $\mathbf{Z}\to\mathbf{Z}\cdot 1_R\subset R$ is injective, and

(ii) every curve C as in the definition of Ш has an R-rational point.

Then taking $\mathcal{E}=E(R)$ and $\mathcal{S}$ to be the set of isomorphism classes of (C,P) with $P\in C(R)$, we get a four-term exact sequnce (13) as before; moreover, the group S_m defined by (14) (with $B=\mathcal{E}$, $C=\mathcal{S}$) is independent up to canonical isomorphism of the choice of R and is isomorphic to Sel_m. Possible choices for R (besides $\overline{\mathbf{Q}}$) would be

—the real numbers $\mathbf{R}$,

—the complex numbers $\mathbf{C}$,

—the p-adic integers $\mathbf{Z}_p$ (or equivalently, the p-adic numbers $\mathbf{Q}_p$) for any prime p,

—the group $\hat{\mathbf{Z}}=\varprojlim(\mathbf{Z}/n\mathbf{Z})=\prod_p\mathbf{Z}_p$ (or equivalently, the finite adeles $\mathbb{A}^f=\hat{\mathbf{Z}}\otimes\mathbf{Q}$),

—any product of these, e.g. the full adele ring $\mathbb{A}=\mathbb{A}^f\times\mathbf{R}$.

Choosing R to be $\hat{\mathbf{Z}}$ or $\mathbb{A}$ seems to be the best choice for our purpose of relating Ш to the local and global Diophantine properties of E.

5. Summary

We can summarize the contents of the paper as folows: Let

$$Z(E,s)=\sum_{\substack{(x,y,z)\in\mathbf{Z}^3/\{\pm1\}\\ \gcd(x,y,z)=1\\ f(x,y,z)=0}}\frac{1}{(x^2+y^2+z^2)^{s/2}}\qquad(\mathrm{Re}(s)>0),$$

$$D(E,s)=\sum_{n=1}^{\infty}\Bigg(\sum_{\substack{(x,y,z)\in(\mathbf{Z}/n\mathbf{Z})^3/(\mathbf{Z}/n\mathbf{Z})^\times\\ \gcd(x,y,z,n)=1\\ f(x,y,z)\equiv0\pmod n}}1\Bigg)\frac{1}{n^s}\qquad(\mathrm{Re}(s)>2),$$

and define numbers $\kappa\in\mathbf{R}_{>0}$ and $r\in\mathbf{Z}_{\geq0}$ by

$$\Omega^{-1}Z(E,s)^2\sim\kappa(\pi/s)^r\qquad(s\to0).$$

Then the Birch–Swinnerton-Dyer conjecture is equivalent to the statement that $D(E,s)$ continues meromorphically to $s=1$ and satisfies

$$D(E,s)\sim-\frac{\kappa\,|\text{Ш}|}{(s-1)^r}\qquad(s\to1),$$

and the number $|\text{Ш}|$ occurring in this formula has an interpretation in terms of points on E and on principal homogeneous spaces over E over the ring $\varprojlim(\mathbf{Z}/n\mathbf{Z})$.