ON THE CONJECTURE OF BIRCH AND SWINNERTON-DYER FOR AN ELLIPTIC CURVE OF RANK 3

Joe P. Buhler, Benedict H. Gross and Don B. Zagier

Department of Mathematics
Reed College
Portland, Oregon 97202
USA

Department of Mathematics
Brown University
Providence, RI 02912
USA

Max-Planck-Institut für Mathematik
Gottfried-Claren-Str. 26
D-5300 Bonn 3
Fed. Rep. of Germany

and

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

MPI/SFB 84-27

ON THE CONJECTURE OF BIRCH AND SWINNERTON-DYER FOR AN ELLIPTIC CURVE OF RANK 3

Joe P. Buhler, Benedict H. Gross and Don B. Zagier

The object of this note is to verify the conjecture of Birch and Swinnerton-Dyer numerically (to high accuracy) for the elliptic curve

$$E: \quad y^2 = 4x^3 - 28x + 25. \tag{1}$$

The conductor of E is 5077, which is apparently the smallest conductor for a curve of rank 3 over $\mathbb{Q}$. Since previous accurate numerical verifications were done for modular curves of rank 0 or 1, and these can now be confirmed theoretically [2], [4], it seemed desirable to test a curve having a larger rank.

We assume some familiarity with the theory of elliptic curves; good references are [3] and [5].

1. The canonical height function

One of the main ingredients in the Birch-Swinnerton-Dyer formula is the regulator, i.e. the determinant of the matrix expressing the canonical height pairing on $E(\mathbb{Q})\otimes\mathbb{R}$ with respect to a $\mathbb{Z}$-basis of $E(\mathbb{Q})/E(\mathbb{Q})_{tors}$. In this section we describe how to calculate the canonical height of a point $P \in E(\mathbb{Q})$.

We first recall the definition. The global minimal model for E has the form

$$y^2 + y = x^3 - 7x + 6, \tag{2}$$

obtained by replacing y by $2y+1$ in (1) and dividing by 4. If $P \in E(\mathbb{Q})$, then the <u>naive height</u> of P is defined as

(3) $\qquad h(P) = \log \max(|a|, b)\ , \quad x(P) = \frac{a}{b}\ ,\ b > 0,\ (a,b) = 1$

(here it does not matter whether we use model (1) or (2) for E, as the x-coordinates are the same); the <u>canonical height</u> is the unique quadratic form $\hat{h}$ on $E(\mathbb{Q}) \otimes \mathbb{R}$ such that $\hat{h}(P) - h(P)$ is bounded, and the <u>canonical height pairing</u> is the associated bilinear form $\langle P,P'\rangle = \frac{1}{2}(\hat{h}(P+P') - \hat{h}(P) - \hat{h}(P'))$. The definition of $\hat{h}$ immediately implies the formula $\hat{h}(P) = \lim_{n\to\infty} n^{-2} h(nP)$, but this is not convenient for calculations. A formula which is usable is

(4) $$\hat{h}(P) = \log b + F(x(P)),$$

where b denotes the denominator of $x(P)$ as in (3) and $F(x)$ is the real-valued function defined by

(5) $$F(x) = \log|x| + \sum_{n=0}^{\infty} 4^{-n-1} \log z_n,$$

$$z_n = 1 + \frac{14}{x_n^2} - \frac{50}{x_n^3} + \frac{49}{x_n^4}, \quad x_0 = x, \quad x_{n+1} = \frac{x_n^4 + 14x_n^3 - 50x_n^2 + 49}{4x_n^4 - 28x_n^2 + 25x_n}.$$

Near $x = 0$ the first two terms in (5) become infinite, but we can combine them to obtain

(6) $$F(x) = \frac{1}{4}\log(x^4 + 14x^2 - 50x + 49) + \sum_{n=1}^{\infty} 4^{-n-1} \log z_n,$$

a formula which now makes sense for all x. Note that the formula relating x_{n+1} to x_n is the formula relating $x(2P)$ to $x(P)$ for $P \in E$, so that $x_n = x(2^n P)$. In particular, $x_n \geq e_3 = 1.946\ldots$ for $n \geq 1$, where $e_1 < e_2 < e_3$ denote the roots of the polynomial $4x^3 - 28x + 25$, so z_n lies between 1 and 1.328... and $\log z_n$ between 0 and 0.284... . Therefore the series in (5) or (6) converges very rapidly and we can calculate $\hat{h}(P)$ to any desired degree of accuracy.

Formula (4) is the specialization to our case of a general recipe of Tate [6] for computing heights; indeed, $F(x(P))$ is Tate's formula for the

infinite component of $\langle P,P\rangle$ while $\nu_p(b)\log p$ (p prime) gives the p-component of the canonical height (even for the prime $p=5077$ of bad reduction, since the fibre of the Néron model at p is irreducible). However, Tate's result, although quoted in the literature, has not yet been published, so we give a direct proof of (4) in our case. By virtue of the definition, it will suffice to show that the expression on the right-hand side of (4) differs by a bounded amount from $h(P)$ and is multiplied by 4 if P is replaced by $2P$. By the formula already cited, replacing P by $2P$ replaces $x(P)=a/b$ by $x(2P)=a^*/b^*$, where

$$a^* = a^4+14a^2b^2-50ab^3+49b^4 , \quad b^* = 4a^4-28a^2b^2+25ab^3 .$$

We claim that b^* is the exact denominator of $x(2P)$. Indeed, an elementary calculation with g.c.d.'s shows that $(a^*,b^*)=1$ for any integers a, b with $(a,b)=1$ unless $a\equiv 92b \pmod{5077}$, in which case $5077|(a^*,b^*)$. But this cannot happen here since $4x^3-28x+25 = 4(x-92)^2(x+184)+5077(20x-1227)$ would be divisible by 5077 but not by 5077^2 if x were $\equiv 92 \pmod{5077}$ and hence could not be a square. (This is an elementary restatement of the fact that the Néron model at 5077 has only one component.) On the other hand, replacing P by $2P$ replaces x_n, z_n by x_{n+1}, z_{n+1} in (5), so

$$\begin{aligned} F(x(2P)) &= \log|x(2P)| + \sum_{n=0}^{\infty} 4^{-n-1}\log z_{n+1} \\ &= \log|x(2P)| + 4\,(F(x)-\log|x|-4^{-1}\log z_1) \\ &\qquad\qquad (\,x = x(P)\,) \\ &= 4\,F(x) - \log(4x^4-28x^3+25) \\ &= 4\,(F(x(P))+\log b) - \log b^* , \end{aligned}$$

proving the first assertion. As to the difference of h and $\hat{h}$,

we can write (3) as $h(P) = \log b + \log \max\left(\frac{|a|}{b}, 1\right)$, so

$$\hat{h}(P) - h(P) = F(x) - \log \max(|x|, 1) \qquad (x = x(P)).$$

If $x \geq e_3 = 1.94\ldots$ is in the right-hand component of $E(\mathbb{R})$, then the same is true for all x_n $(n \geq 0)$, so $1 \leq z_n \leq 1.328\ldots$ for all n in (5) and therefore

$$0 \leq F(x) - \log x \leq \sum_{n=0}^{\infty} 4^{-n-1} \log(1.328\ldots) = 0.0947\ldots .$$

The other component $e_1 \leq x \leq e_2$ of $E(\mathbb{R})$ is compact and we easily find the minimum and maximum of $F(x) - \log \max(|x|, 1)$ there to be $0.4006\ldots$ and $1.205\ldots$ (obtained for $x = e_1$ and $x = -1$, respectively; see Figure 1). Hence in all cases we have

$$(7) \qquad h(P) \leq \hat{h}(P) \leq h(P) + 1.205\ldots .$$

This completes the proof of (3). We remark that the difference between the naive and canonical heights on elliptic curves has been studied by several authors (cf. [7] and the literature cited there) but that the inequality (7) is much sharper than the one obtained by specializing their results, suggesting that some improvements in the general case may still be possible.

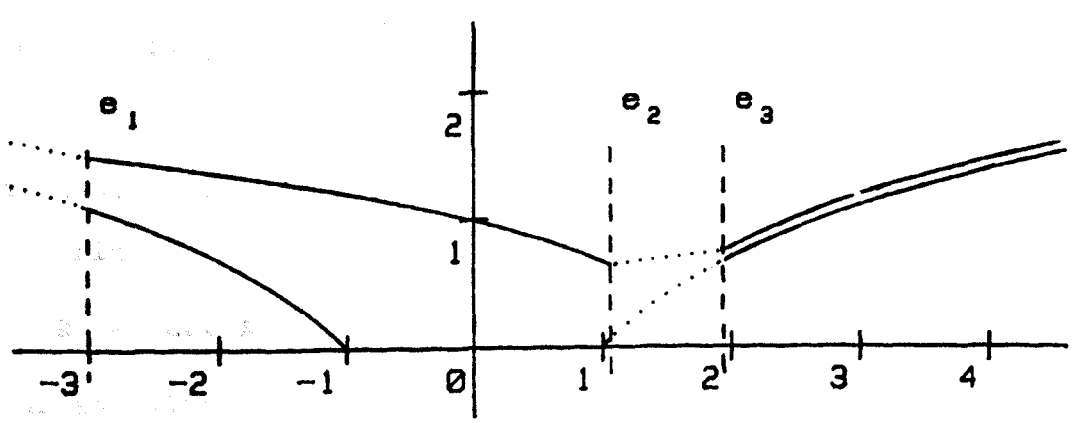

Figure 1. The functions $F(x)$ and $\log \max(|x|, 1)$

2. The Mordell-Weil group and the regulator

Let N_p $(p \neq 5077)$ denote the cardinality of $E(\mathbb{Z}/p\mathbb{Z})$, i.e. 1 plus the number of solutions of (2) in integers modulo p. Then $|E(\mathbb{Q})_{tors}|$ must divide N_p for all p; since $N_2 = 5$ and $N_3 = 7$ it follows that $E(\mathbb{Q})$ is free abelian. We claim that it is of rank 3, generated by the three points

$$P_0 = (0,2)\ , \qquad P_1 = (1,0)\ , \qquad P_2 = (2,0)\ .$$

It follows from equation (7) that these are the only points with canonical height less than 1, since $h(P) \le \hat{h}(P) \le 1$ implies $|a| \le e$, $b \le e$ and hence (since b is always a square) $b=1$, $a \in \{-2,-1,0,1,2\}$; of these five candidates, only $a=0,1,2$ lead to points with $\hat{h}(P) < 1$. On the other hand, one sees by a 2-descent that P_0, P_1, P_2 generate $E(\mathbb{Q})/2E(\mathbb{Q})$, which is of rank 3 over $\mathbb{Z}/2\mathbb{Z}$. These two facts and the fact that $E(\mathbb{Q})$ is torsion-free imply by the usual proof of the Mordell-Weil theorem (cf. any text on elliptic curves) that $E(\mathbb{Q}) = \mathbb{Z}P_0 + \mathbb{Z}P_1 + \mathbb{Z}P_2$ as claimed. Using the algorithm of §1 we can calculate the entries of the matrix

$$A = (\langle P_i, P_j \rangle)_{0 \le i,j \le 2} = \begin{pmatrix} .9909\ldots & -.2365\ldots & -.2764\ldots \\ -.2365\ldots & .6682\ldots & .0333\ldots \\ -.2764\ldots & .0333\ldots & .7670\ldots \end{pmatrix}$$

to any desired accuracy. The regulator is the determinant of this matrix:

$$(8) \qquad R = \det A = .41714355875838396981711954461 8093\ldots \quad .$$

As an illustration, we have given the representations of P as $n_0P_0 + n_1P_1 + n_2P_2$ and the naive and canonical heights of P for 18 integral points $P \in E(\mathbb{Q})$ in Table 1; the canonical heights can be

computed either by the algorithm of §1 or as $(n_0 n_1 n_2)\, A\, (n_0 n_1 n_2)^t$. One has of course also the negatives $-P = (x,-y-1) = -n_0P_0-n_1P_1-n_2P_2$ with the same heights. The large number of 36 integral points seems to be typical of curves with a high rank relative to their conductor.

x	y	n_0	n_1	n_2	$\hat{h}(P)$	h(P)
-3	0	0	-1	-1	1.50192454	1.09861229
-2	3	0	-1	1	1.36857251	.69314718
-1	3	-1	0	-1	1.20508110	0.00000000
0	2	1	0	0	.99090633	0.00000000
1	0	0	1	0	.66820517	0.00000000
2	0	0	0	1	.76704336	.69314718
3	3	1	1	0	1.18592770	1.09861229
4	6	-1	-1	-1	1.46677848	1.38629436
8	21	1	-1	0	2.13229530	2.07944154
11	35	-1	-1	1	2.43916362	2.39789527
14	51	0	2	0	2.67282066	2.63905733
21	95	0	0	-2	3.06817342	3.04452244
37	224	-2	0	-1	3.62493152	3.61091791
52	374	1	-1	2	3.96137952	3.95124372
93	896	2	2	1	4.53836901	4.53259949
342	6324	-2	0	1	5.83640586	5.83481074
406	8180	0	2	2	6.00769815	6.00635316
816	23309	1	3	-1	6.70508531	6.70441435

Table 1. Integral points on E

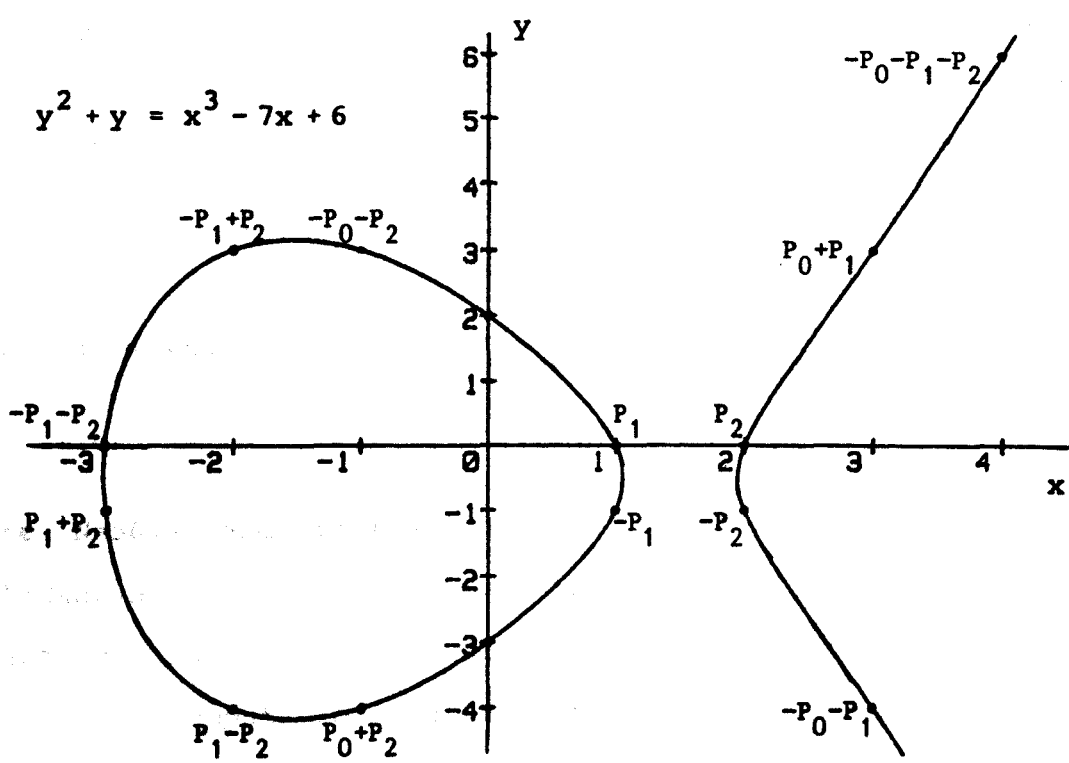

Figure 2. Integral points on E

3. The period

The group $E(\mathbb{R})$ has two connected components. Let $\omega = \frac{dx}{2y+1}$ be a Néron differential on E over $\mathbb{Z}$, and $|\omega|$ the associated measure on $E(\mathbb{R})$. Define the period Ω by

$$\Omega = \int_{E(\mathbb{R})} |\omega| = 2\int_{E(\mathbb{R})^0} |\omega|.$$

If we write (0.1) in the form $y^2 = 4(x-e_1)(x-e_2)(x-e_3)$ with $e_1 < e_2 < e_3$, we may calculate this period integral using Gauss' arithmetic-geometric mean. This is defined on two positive real arguments x and y by $M(x,y) = \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$, where $x_0 = x$, $y_0 = y$, $x_{n+1} = \frac{x_n+y_n}{2}$, $y_{n+1} = \sqrt{x_n y_n}$. We find (Gauss):

$$\Omega = 4\int_{e_3}^{\infty} \frac{dx}{y} = \frac{2\pi}{M(\sqrt{e_3-e_1},\sqrt{e_3-e_2})} = \frac{2\pi}{M(2.22689\ldots,0.938503\ldots)} \tag{9}$$

$$= 4.151687983086933049884175683507286\ldots$$

4. The L-series

The L-series for E over $\mathbb{Q}$ is given by an Euler product which converges in the right half plane $\mathrm{Re}(s) > 3/2$:

$$L(E,s) = (1+5077^{-s})^{-1} \prod_{p\neq 5077} (1-a_p p^{-s}+p^{1-2s})^{-1} = \sum_{n=1}^{\infty} a_n n^{-s}$$

where a_p $(p\neq 5077)$ equals $p+1-N_p$ with N_p as in §2. We have

$$\Lambda(s) := N^{s/2}(2\pi)^{-s}\Gamma(s)L(E,s) = \int_0^{\infty} f\left(\frac{iy}{\sqrt{N}}\right) y^{s-1}\, dy ,$$

where $N=5077$ and $f(\tau) = \sum_{n=1}^{\infty} a_n e^{2\pi i n\tau}$ $(\tau\in\mathbb{C},\ \mathrm{Im}(\tau)>0)$. The Weil-Taniyama conjecture asserts that $f(\tau)$ is a cusp form of weight 2 and level N = conductor of E. This could be checked by a finite computation, but we have not carried it out (the space

of cusp forms of weight 2 and level 5077 has dimension over 400) and will simply assume its truth (thus this note could more properly be described as a simultaneous numerical verification of the Birch-Swinnerton-Dyer and Weil-Taniyama conjectures). Then $f(\tau)$ satisfies the functional equation $f(-1/N\tau) = N\tau^2 f(\tau)$ and the analytic continuation and functional equation of $L(E,s)$ follow:

$$(10) \qquad \Lambda(s) = \int_1^\infty f\left(\frac{iy}{\sqrt{N}}\right)\left(y^{s-1} - y^{1-s}\right) dy = -\Lambda(2-s) .$$

In particular, the order of $L(E,s)$ at $s=1$ is odd and the rth derivative ($r \geq 1$ odd) is given by

$$(11) \qquad \begin{aligned} \Lambda^{(r)}(1) &= 2\int_1^\infty f\left(\frac{iy}{\sqrt{N}}\right)(\log y)^r\, dy \\ &= 2\sum_{n=1}^\infty a_n \int_1^\infty e^{-2\pi ny/\sqrt{N}} (\log y)^r\, dy. \end{aligned}$$

If $\Lambda(s)$ vanishes to order $\geq r$ at $s = 1$, then integrating (11) once by parts gives

$$(12) \qquad L^{(r)}(1) = \frac{2\pi}{\sqrt{N}}\Lambda^{(r)}(1) = 2r!\sum_{n=1}^\infty \frac{a_n}{n} G_r\left(\frac{2\pi n}{\sqrt{N}}\right),$$

where

$$G_r(x) = \frac{1}{(r-1)!}\int_1^\infty e^{-xy}(\log y)^{r-1}\frac{dy}{y} \qquad (r \geq 1) .$$

The series (12) is rapidly convergent, because $G_r(x) \sim x^{-r}e^{-x}$ as $x \to \infty$, so it can be used to compute $L^{(r)}(1)$ if we have a good algorithm to compute $G_r(x)$.

The function $G_1(x)$ is the familiar exponential integral $\int_1^\infty e^{-xy}\frac{dy}{y}$, which can be calculated for small x ($x<3$) by the power series

$$G_1(x) = \log\frac{1}{x} - \gamma + \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n\cdot n!}x^n \qquad (\gamma = \text{Euler's constant})$$

and for large x ($x>2$) by the continued fraction expansion

$$G_1(x) = \cfrac{e^{-x}}{x + \cfrac{1}{1 + \cfrac{1}{x + \cfrac{2}{1 + \cfrac{2}{x + \cfrac{3}{1 + \cdots}}}}}}$$

Taking 250 terms of the series in (12) gives $L'(1) \approx 0$ to 13 decimal places. But this implies that $L'(1) = 0$ exactly, since the main result of [2] implies that $L'(1)$ is a simple multiple of the height of some rational point on E ("Heegner point") and, as we have seen, E contains no rational points of very small non-zero height. Since $L(s)$ has odd order, we have $\mathrm{ord}_{s=1} L(s) \geq 3$.

In general, the functions $G_r(x)$ satisfy $G_0(x) = e^{-x}$, $G_r'(x) = -\frac{1}{x} G_{r-1}(x)$, so

$$G_r(x) = P_r(\log \tfrac{1}{x}) + \sum_{n=1}^{\infty} \frac{(-1)^{n-r}}{n^r n!} x^n$$

for some polynomial P_r of degree r. To determine P_r, we use the integral representation:

$$(13) \qquad G_r(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(s)}{s^r} x^{-s}\, ds \qquad \text{any } c > 0.$$

(To prove (13), we observe that the right hand side satisfies the same recursive differential equations as $G_r(x)$ and tends to zero as $x \to \infty$.) Shift the path of integration in (13) to the left; then the residue at $s = -n$ gives the term $(-1)^{n-r}x^n/n^r n!$ and the residue at $s = 0$ gives $P_r(\log \frac{1}{x})$. Hence

$$P_r(t) = \sum_{m=0}^{r} \gamma_{r-m} \frac{t^m}{m!} \quad \text{where} \quad \Gamma(1+s) = \sum_{n=0}^{\infty} \gamma_n s^n.$$

Since by Euler-Maclaurin

$$\log \Gamma(1+s) = -\gamma s + \sum_{n=2}^{\infty} \frac{(-1)^n}{n} \zeta(n) s^n ,$$

we find, for $r = 3$, the expansion

$$G_3(x) = \frac{1}{6}(\log \frac{1}{x} - \gamma)^3 + \frac{\pi^2}{12}(\log \frac{1}{x} - \gamma) - \frac{\zeta(3)}{3}$$

$$+ \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n^3 n!}$$

which converges for all x. Using this we find the value

$$\lim_{s \to 1} \frac{L(E,s)}{(s-1)^3} = 2 \sum_{n=1}^{\infty} \frac{a_n}{n} G_3\left(\frac{2\pi n}{\sqrt{5077}}\right)$$

(14)

$$\approx 1.7318499001193006897919750851$$

using the terms for $n \leq 600$ (the error made in breaking off the series here can be estimated using (12) and the formulas $G_3(x) \sim x^{-3} e^{-x}$ and $|a_n| \leq d(n)/\sqrt{n}$, where $d(n)$ is the number of divisors of n).

The results of the computations described in this section are summarized in Table 2.

n	a_n	$G_1(\frac{2\pi n}{\sqrt{N}})$	$G_3(\frac{2\pi n}{\sqrt{N}})$	$2\sum_1^n \frac{a_m}{m} G_1(\frac{2\pi m}{\sqrt{N}})$	$2\sum_1^n \frac{a_m}{m} G_3(\frac{2\pi m}{\sqrt{N}})$
1	1	1.93741992	2.26675143	3.87483985	4.53350286
2	-2	1.32687953	.98498602	1.22108079	2.56353082
3	-3	1.00056041	.54955613	-.78004003	1.46441856
4	2	.78875755	.34359041	.00871752	1.80800897
5	-4	.63840821	.22972608	-1.01273562	1.44044725
6	6	.52596620	.16064962	.03919678	1.76174648
7	-4	.43894007	.11604939	-.46244901	1.62911861
8	0	.36992797	.08592813	-.46244901	1.62911861
9	6	.31419941	.06487957	-.04351647	1.71562470
10	8	.26856035	.04977090	.38618010	1.79525814
50	-22	.00231086	.00005681	-.00236637	1.73179489
100	22	.00001521	.00000013	.00001335	1.73185001
250	48	.00000000	.00000000	.00000000	1.73184990

Table 2. Computation of $L'(1)$ and $L'''(1)$

The Conjecture

The conjecture of Birch and Swinnerton-Dyer predicts that $\mathrm{ord}_{s=1} L(E,s) = \mathrm{rank}(E) = 3$ and that

$$\lim_{s\to 1} \frac{L(E,s)}{(s-1)^3} = \Omega \cdot R \cdot \mathrm{Card}\,(Ш)$$

where $Ш$ is the (conjecturally finite) Tate-Shafarevich group of E over $\mathbb{Q}$. Equations (8) and (9) give

$$\Omega \cdot R = 1.7318499001193006897919750850601 54\ldots$$

which agrees with the right-hand side of (14) within the accuracy of our computations in §3. This strongly suggests that the conjecture is true and that $Ш = (1)$. We have checked, via a 2-descent (cf. [4]) that the 2-primary component of $Ш$ is trivial.

Bibliography

[1] Brumer, A. and Kramer, K., The rank of elliptic curves. Duke Math. J. 44 (1977), 715-743.

[2] Gross, B. and Zagier, D., Points de Heegner et dérivées de fonctions L. C.R. Acad. Sc. Paris 297 (1983), 85-87.

[3] Manin, Y. I., Cyclotomic fields and modular curves. Uspehi Mat. Nauk 26 (1971), 7-71 = Russian Math. Surveys 26 (1971), 7-78.

[4] Mazur, B. and Swinnerton-Dyer, H.P.F., Arithmetic of Weil curves. Inv. math. 25 (1974), 1-61.

[5] Tate, J., The arithmetic of elliptic curves. Inv. math. 23 (1974), 179-206.

[6] Tate, J., Letter to Serre. Oct. 1, 1979.

[7] Zimmer, H.G., On the difference of the Weil height and the Néron-Tate height. Math. Z. 147 (1976), 35-51.