THE BSD FORMULA OVER NUMBER FIELDS

JOHN E. CREMONA

Abstract. We give an explicit statement of the Birch–Swinnerton-Dyer (BSD) formula for elliptic curves over general number fields, carefully defining all the terms and showing how they may be be computed using the software packages SageMath and Magma. In particular, we highlight certain differences in the way the formula has been stated in the literature.

The Birch–Swinnerton-Dyer (BSD) formula for elliptic curves over $\mathbb{Q}$ is clearly stated and explained in many texts, but for elliptic curves over general number fields there are a number of subtleties leading to the appearance of factors which are equal to 1 in the rational case, and clear and explicit statements are harder to find. We cite three, by Tate [5], Gross [2], and T. Dokchitser [1], as well as a discussion on MathOverflow [3] highlighting the common confusions which people have. The point of this text is to clear up the potentially confusing issues.

For simplicity we concentrate on the case of elliptic curves, only briefly mentioning the adjustments needed for Abelian varieties of higher dimension g.

Traditionally, the BSD formula for an elliptic curve E defined over the number field K is written as a formula for the leading coefficient of the L-function $L(E/K, s)$ at $s = 1$, which we call the *L-value*, in terms of other quantities, including the order of the *Tate-Shafarevich group* $Ш(E/K)$. We will assume the standard conjectures, first that $L(E/K, s)$ has analytic continuation so that its behaviour at $s = 1$ is defined, and also that Ш is finite. Then we rearrange the formula so that it expresses $S = |Ш|$ in terms of quantities which may be computed.

1. Code

Where possible, we give the appropriate commands in SageMath (version 10.1) and Magma (version V2.28-3) to compute these quantities. Many of these functions have a precision parameter to control the precision of the output: the input is always exact, but the output may in principle be computed to arbitrary precision. Magma can compute all the quantities separately (with a little work) and can also evaluate the entire formula. SageMath can compute all the quantities; unless $K = \mathbb{Q}$, computing the analytic rank and L-value itself currently requires calling Magma from SageMath, or using PARI/GP library functions.

1.1. Magma.

For elliptic curves over general number fields all terms in the formula are computed and combined in the function `ConjecturalSha()`:

```
ConjecturalSha(E, [P1,...,Pr]);
```

1.2. SageMath.

For elliptic curves over $\mathbb{Q}$, we have:

```
E.sha().an_numerical()
```

Work is under way to implement a similar function for elliptic curves over number fields in future versions of SageMath.

2. Notation

K is a number field of degree $d = r_1 + 2r_2$, where r_1 and r_2 are the number of real and complex places, and discriminant D_K.

E is an elliptic curve defined over K, given by an integral Weierstrass equation or model with coefficients a_1, a_2, a_3, a_4, a_6. The model is assumed to be integral, but not minimal, as global minimal models do not always exist. The discriminant of the model is denoted Δ_E, which is an integral element of K, and $\mathfrak{d}_E$ denotes the minimal discriminant ideal of E, an integral ideal which is independent of the model. Then $(\Delta_E) = \mathfrak{u}_E^{12}\mathfrak{d}_E$ for some fractional ideal $\mathfrak{u}_E$. A global minimal model exists if and only if $\mathfrak{u}$ is principal; in general the class of $\mathfrak{u}_E$ is an invariant of E.

The Mordell-Weil group $E(K)$ is a finitely-generated abelian group of rank $r = r(E(K))$, with finite torsion subgroup $T = E(K)_{\text{tors}}$. Let $P_1, \dots, P_r \in E(K)$ be points which generate $E(K)/T$.

The L-function of E is $L(E/K, s)$. By the weak BSD conjecture, we assume that its order of vanishing at $s = 1$ (that is, the *analytic rank* of E/K) is equal to the Mordell-Weil rank r.

3. Outline formula

The formula for S has the form

$$S = \frac{(\text{field factor})(L\text{-value})}{(\text{Mordell-Weil factor})(\text{local factor})}.$$

We will define each of these factors in turn. Note that while different sources disagree on the exact definitions of the factors, the differences cancel out in the product: in particular there is a factor 2^{r_2} which is included either by multiplying the local-factor (as in [1]) and or by dividing the field-factor (as in [5]).

4. The field factor

This is

$$|D_K|^{1/2},$$

so is equal to 1 for $K = \mathbb{Q}$.

Gross includes this in with the local factors at infinite places. In some expositions, there is an extra factor of 2 in the contribution from each complex place to the local factor, and the field factor is set to $|D_K|^{1/2}/2^{r_2}$ to compensate.

For Abelian varieties of dimension g the factor is $|D_K|^{g/2}$.

4.1. **Magma code:**

```
Sqrt(Abs(Discriminant(K)));
```

4.2. **SageMath code:**

```
RR(K.discriminant().abs()).sqrt()
```

5. The L-value

This is

$$L^{(r)}(E,1)/r!.$$

Note that in general it is only conjectural that the L-function of E has analytic continuation as far as $s=1$, so both the fact that this quantity is well-defined, and the analytic properties assumed in the algorithm to evaluate its derivative at $s=1$, are conjectural.

5.1. **Magma code:**

```
Evaluate(LSeries(E),1: Derivative:=r) / Factorial(r);
```

5.2. **SageMath code:** For $K=\mathbb{Q}$:

```
E.lseries().dokchitser().derivative(1,r) / factorial(r)
```

For general K (including $K=\mathbb{Q}$):

```
RR(pari(E).lfun(f'1+x+O(x^{r+1})').polcoef(r))
```

6. The Mordell-Weil factor

This is

$$\frac{R_{NT}(P_1,\ldots,P_r)}{|T|^2},$$

where R_{NT} is the Néron-Tate regulator, i.e. the determinant of the Néron-Tate height pairing matrix of the generators, whose (i,j)-entry is

$$\langle P_i,P_j\rangle = \frac{1}{2}(\hat{h}(P_i+P_j)-\hat{h}(P_i)-\hat{h}(P_j)),$$

so in particular the diagonal entries are the canonical heights $\hat{h}(P_i)$ of the P_i. Here the canonical height $\hat{h}$ needs to be normalised correctly:

(1) It should be defined with respect to the x-coordinate, or equivalently with respect to the divisor $2(O)$. There is another convention (used in some of Silverman's papers, for example [4]) which is half of this, being defined with respect to the divisor (O) instead of $2(O)$. In the MathOverflow post [3], Silverman explains why it makes sense to use $2(O)$ in the context of BSD.
(2) It should **not** be normalized (by dividing by the degree d) to be invariant under base-change. This non-normalised canonical height is known as the *Néron-Tate height*. Note that both Magma and SageMath normalize the height by default. SageMath has an option to not normalize canonical heights or the regulator, but with Magma the regulator needs to be adjusted by multiplying by d^r.

6.1. **Magma code:**

```
Regulator(P1,...,Pr) * d^r / Order(TorsionSubgroup(E))^2;
```

6.2. **SageMath code:** Here the regulator function has a parameter to control the normalisation, which by default is set to True, so we explicitly set it to False for our purposes:

```
E.regulator_of_points([P1,...,Pr], normalised=False) / E.torsion_order()^2
```

6.3. **Saturation.** The expressions given here will only give the correct regulator if the points $P_1, \dots, P_r$ are *saturated*, i.e. generate all of $E(K)$ modulo torsion, so that $[E(K) : \langle P_1, \dots, P_r\rangle] = |T|$. An alternative expression is sometimes seen which avoids this condition:

$$\frac{R(P_1, \dots, P_r)}{[E(K) : \langle P_1, \dots, P_r\rangle]^2};$$

this is a neat trick, but not very useful in practice for computations.

Note that if we evaluate the formula with points are not saturated, but only generate a subgroup (modulo torsion) of index n, then this factor, and hence the overall expression for S, will be $1/n^2$ times the correct value. This will often lead to a non-integral value of S, which can be detected.

6.4. **Higher genus.** For Abelian varieties A of general dimension g the formula involves both A and its dual A', and points $P_1', \dots, P_r'$ generating $A'(K)$ modulo torsion. (Recall that A and A' are isogenous so have the same rank r.) The factor is then

$$\frac{\det(\langle P_i, P_j'\rangle)}{\#A(K)_{\text{tors}} \cdot \#A'(K)_{\text{tors}}}$$

where now $\langle \cdot, \cdot \rangle$ is the height pairing between A and A'.

7. The local factor

The complete local factor is a product of individual local factors, one for each place of K. We will define the local factors at infinite places in a way which depends on the model, but the complete local factor also includes a correction factor making the complete expression independent of the model. The correction factor is 1 for a global minimal model, but these do not always exist.

7.1. **Finite places.** At a finite place v the local factor is c_v, the Tamagawa number at v, which is defined to be the index $c_v = [E(K_v) : E(K_v)^0]$, i.e. the number of connected components of $E(K_v)$, and is a positive integer. At a place of good reduction we have $c_v = 1$, so the product of these factors over all finite places is in effect a finite product.

7.1.1. *Magma code:*

```
&*TamagawaNumbers(E);
```

gives the product of the Tamagawa numbers c_v.

7.1.2. *SageMath code:* Over all number fields (including $\mathbb{Q}$)

```
E.tamagawa_product_bsd()
```

gives the product of all Tamagawa numbers c_v multiplied by a correction factor (defined below) to allow for the model not being minimal, while `E.tamagawa_product()` returns just the product of the Tamagawa numbers.

7.2. **Infinite places.** At each infinite place v, the local factor is a suitably normalized period Ω_v, and the total contribution from all infinite places is $\prod_{v|\infty} \Omega_v$. Theoretically the Ω_v are defined in terms of a Néron differential. We first define the periods for a fixed model of E and then explain how to normalize them. Let $\omega_E = dx/(2y + a_1x + a_3)$ be the usual invariant differential associated to a Weierstrass model for E.

7.2.1. *Real places.* If v is a real place we define $\Omega_v(E)$ to be the integral of ω_E over $E(\mathbb{R})$. This is equal to the least positive real period multiplied by the number of real components at the place v, which is 1 when $v(\Delta_E) < 0$ and 2 when $v(\Delta_E) > 0$.

When $v(\Delta_E) > 0$ the period lattice of E at v has generating periods x, yi with $x, y > 0$, and $\Omega_v = 2x$. When $v(\Delta_E) < 0$ the period lattice of E at v has generating periods $2x$, $x + yi$ with $x, y > 0$, and $\Omega_v = 2x$.

7.2.2. *Complex places.* If v is a complex place we define $\Omega_v(E)$ to be the integral of $\omega_E \wedge \overline{\omega_E}$ over $E(\mathbb{C})$. If the period lattice of E at v has generating periods w_1, w_2 with $\Im(w_2/w_1) > 0$, then $\Omega_v = 2\Im(\overline{w_1}w_2)$, which is positive: it is double the covolume of the period lattice.

In [1], Ω_v for complex v is defined to be double the definition here, which appears to be an error. Note that if we instead defined Ω_v to be the lattice covolume, we would need to compensate by dividing the field factor by 2^{r_2}.

7.2.3. *Adjustment for non-minimal models.* For a global minimal model, the minimal discriminant ideal is simply the principal ideal $\mathfrak{d}_E = (\Delta_E)$. In general, $(\Delta_E) = \mathfrak{u}_E^{12}\mathfrak{d}_E$ for some fractional ideal $\mathfrak{u}_E$ whose valuation at each prime $\mathfrak{p}$ is $1/12$ of the difference between the valuation at $\mathfrak{p}$ of Δ_E and that of the local minimal model at $\mathfrak{p}$. Hence

$$N(\mathfrak{u}_E) = \left|\frac{N(\Delta_E)}{N(\mathfrak{d}_E}\right|^{1/12}.$$

The product $\prod_{v|\infty} \Omega_v N(\mathfrak{u}_E)$ is independent of the model: scaling the model for E by $\lambda \in K^*$ multiplies $\mathfrak{u}_E$ by the principal ideal (λ) and hence multiples $N(\mathfrak{u}_E)$ by $|N(\lambda)|$; in compensation, each Ω_v is divided by $|\lambda|_v$, and the product of these is also $|N(\lambda)|$.

7.3. **Complete local factor.** The complete local factor, including a contribution from all finite and infinite places and independent of the model, is

$$\prod_{v\nmid\infty} c_v \prod_{v|\infty} \Omega_v N(\mathfrak{u}_E) = \prod_{v\nmid\infty} c_v \prod_{v|\infty} \Omega_v \left|\frac{N(\Delta_E)}{N(\mathfrak{d}_E}\right|^{1/12}.$$

7.4. **Magma code:**

```
Periods(E);
```

gives a basis for the period lattice with respect to each infinite place, starting with the real places. In evaluating `ConjecturalSha()` the product of the actual real and complex periods Ω_v are computed from these, including factors of 2 for the real places where the discriminant is positive and a factor 2^{r_2} from the Dokchitser normalisation of the complex periods, but the individual Ω_v values are not easily accessible to the user.

7.5. **SageMath code:** For each infinite place,

```
E.period_lattice(v).omega()
```

gives the value of Ω_v, including the factor of 2 for real places where the discriminant is positive. Hence the complete local factor is given by

```
prod([EK.period_lattice(v).omega() for v in K.places()])
                                  * E.tamagawa_product_bsd()
```

8. The full formula

The analytic order of Ш is given by the formula

$$S = |D_K|^{1/2} \cdot \frac{L^{(r)}(E,1)}{r!} \cdot \frac{|T|^2}{R(P_1,\ldots,P_r)} \cdot \frac{1}{N(\mathfrak{u}_E)\prod_{v\nmid\infty} c_v \prod_{v\mid\infty}\Omega_v}.$$

9. Comparisons

9.1. **Tate.** In [5], Tate has a factor $|\mu|^g$ where g (which is d in Tate's notation) is the dimension of the variety and he states that $|\mu| = |D_K|^{1/2}/2^{r_2}$.

The difference between this and our field factor can be accounted for by doubling the contribution at the complex places, the normalization for which Tate does not make explicit.

9.2. **Gross.** In [2], Gross states the formula for an abelian variety A in the form

$$\text{L-value} = M(A)R(A)h(A),$$

where $h(A) = |Ш|$. In the analogy with the analytic class number formula, Ш plays the role of the class group whose order is usually denoted h. The other two factors match our Mordell-Weil factor, and the product of the field and local factors respectively. In the case of an elliptic curve $A = E$, in our notation these are

- $R(E) = R_{NT}(P_1,\ldots,P_r)/|T|^2$; and
- $M(E) = M_\infty(E)M_f(E)$, where

$$M_\infty(E) = \prod_{v\mid\infty}\Omega_v \cdot N(\mathfrak{u}_E)\cdot|D_K|^{-1/2},$$

and

$$M_f(E) = \prod_{v\nmid\infty} c_v.$$

9.3. **Dokchitser.** In [1, pp. 3–5], Dokchitser defines the regulator and L-value as we have. He states that the regulator is the determinant of the Néron-Tate height-pairing, where (as above) the prefix "Néron-Tate" indicates that the height is relative to K and not the normalized absolute one.

Dokchitser combines all the local factors together into a quantity called $C_{E/K}$: Writing ω for any invariant differential on E (with arbitrary scaling), and ω_v^o the local Néron differential at a finite place v, he defines

$$C_{E/K} = \prod_{v\nmid\infty} c_v\left|\frac{\omega}{\omega_v^o}\right|_v \cdot \prod_{v\mid\infty,\text{real}}\int_{E(K_v)}|\omega| \cdot \prod_{v\mid\infty,\text{complex}} 2\int_{E(K_v)}\omega_E\wedge\bar{\omega}_E,$$

with c_v the local Tamagawa number at v and $|\cdot|_v$ the normalised absolute value on K_v. If ω is the differential associated to a global minimal model then $\left|\frac{\omega}{\omega_v^o}\right|_v = 1$ for all finite v and this reduces to our unadjusted local factor. In general the factor $N(\mathfrak{u}_E)$ in our formula accounts for the difference in scaling between the differential ω_E and the local Néron differential ω_v^o at every place.

We see that $C_{E/K}$ is independent of the model of E, and is 2^{r_2} times our local factor. Otherwise, his formula is identical to ours.

10. Origin of this document and acknowledgements

This document started life as a collaborative Markdown document on `hackmd.io`, where contributions were also made by Raymond von Bommel. From there it migrated to a SageMath Jupyter notebook, which included worked examples in both SageMath and Magma, before this version. The source for this document, as well as the SageMath Jupyter notebook, are available at the public `github` repository `https://github.com/JohnCremona/BSD`.

Helpful contributions have been gratefully received from Tim Dokchitser and Michael Stoll.

References

[1] T. Dokchitser: *Notes on the parity conjecture* `https://arxiv.org/abs/1009.5389`.

[2] B. H. Gross, *On the conjecture of Birch and Swinnerton-Dyer for elliptic curves with complex multiplication.* In Number theory related to Fermat's last theorem (Cambridge, Mass., 1981), 219–236, Progr. Math., 26, Birkh¨auser. Boston, Mass., 1982.

[3] MathOverFlow `https://mathoverflow.net/questions/139575/bsd-conjecture-for-x-017`.

[4] J.H. Silverman, *Computing heights on elliptic curves*, Math. Comp. **51** (1988), 339-358.

[5] J. Tate: *On the conjectures of Birch and Swinnerton-Dyer and a geometric Analog.* Séminaire N. Bourbaki, 1966, exp. no 306, p. 415-440.