THE BSD FORMULA OVER NUMBER FIELDS

JOHN E. CREMONA

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, Gross, and T. Dokchitser) below, as well as a discussion on MathOverflow 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 is written as a formula for the leading coefficient of the L-series $L(E, s)$ at $s = 1$, which we call the *L-value*, in terms of other quantities, including the order of the *Tate-Shafarevich group* Ш. We always assume the standard conjectures, first that $L(E, s)$ 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 $S = |\text{Ш}|$ in terms of quantities which may be computed.

1. CODE

Where possible, we give the appropriate commands in SageMath and Magma 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 except for the L-value itself (unless $K = \mathbb{Q}$).

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

$$\texttt{ConjecturalSha(E, [P1,...,Pr]);}$$

1.2. **SageMath.** For elliptic curves over $\mathbb{Q}$ only, we have:

$$\texttt{E.sha().an\_numerical()}$$

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 discrimninant 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. We do not assume that the model we have is minimal; global minimal models do not always exist anyway. We denote by $\Delta(E)$ the discriminant of the model, which is an integral element of K, and by $\mathfrak{d}(E)$ the minimal discriminant ideal of E, an integral ideal which is independent of

the model. Then $(\Delta(E))/\mathfrak{d}(E) = \mathfrak{u}^{12}$ for some fractional ideal $\mathfrak{u}$. (A global minimal model exists if and only if $\mathfrak{u}$ is principal; in general the class of $\mathfrak{u}$ 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 $T = E(K)_{tors}$. Let $P_1, \ldots, P_r \in E(K)$ be points which generate $E(K)/T$.

3. OUTLINE FORMULA

The formula for S has the form

$$S = \frac{(\text{Field factor})(L\text{-value})}{(\text{Mordell-Weil factor})(\text{Local factor})}$$

4. FIELD FACTOR

This is simply

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

so is equal to 1 for $K = \mathbb{Q}$). Gross includes this in with the local factors at infinite places. 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 $K = \mathbb{Q}$ only:

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

Otherwise on a system with Magma installed one can resort to

```
RR(magma(E).LSeries().Evaluate(1, Derivative=r)) / factorial(r)
```

6. Mordell-Weil factor

This is

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

where $R()$ is the regulator, i.e. the determinant of the height pairing matrix. This needs to be normalised correctly!

(1) The (i,j)-entry of the height pairing matrix 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 of the P_i;

(2) The canonical height should be defined with respect to the x-coordinate (so that $x(P) - \hat{h}(P)$ is bounded independently of P). There is another convention (used in some of Silverman's papers) which is half of this, as it is defined with respect to the divisor (O) instead of $2(O)$. In the Math-Overflow post, Silverman explains why it makes sense to use $2(O)$ in the context of BSD.

(3) The canonical height should **not** be normalized (by dividing by the degree d) to be invariant under base-change. Note that both Magma and SageMath by default normalize the height. SageMath has an option to not normalize for heights, but not (as of version 9.1) for the regulator. Hence in both cases 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:**

```
E.regulator_of_points([P1,...,Pr]) * d^r / E.torsion_order()^2
```

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

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

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

NB If the 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.

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

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

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

7. Local factors

The complete local factor is a product of individual local factors, one for each place of K. In our expresson the local factors at infinite places will be defined in a way which depends on the model, so the complete local factor also includes a correction factor making the expression independent of the model.

7.1. At finite places.
At a finite place v the local factor is c_v, the Tamagawa number at v. This is a positive integer, the index $c_v = [E(K_v) : E(K_v)^0]$, i.e. the number of connected components of $E(K_v)$. At a place of good reduction we have $c_v = 1$, so this is in effect a finite product.

7.2. Magma code:

$$\&*\texttt{TamagawaNumbers(E);}$$

gives the product of the Tamagawa numbers c_v.

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

$$\texttt{E.tamagawa\_product\_bsd()}$$

gives the product of all Tamagawa numbers c_v times the correction factor $N(\mathfrak{u})$ (defined below) to allow for the model not being minimal, while `E.tamagawa_product()` returns just the product of the Tamagawa numbers (over $\mathbb{Q}$ only).

7.4. At infinite places.
At each infinite place 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 it. Let $\omega_E = dx/(2y + a_1 x + a_3)$ be the usual invariant differential associated to a Weierstrass model for E.

7.5. At a real place.
If v is a real place 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 $\Omega_v = 2x$.

7.6. At a complex place.
If v is a complex place define $\Omega_v(E)$ to be **twice** 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).

If we were to define Ω_v for complex v to be simply the integral of ω_E over $E(\mathbb{C})$, then the BSD formula would need an additional factor of 2^{r_2}.

7.7. Complete local factor.

$$\frac{1}{\prod_{v \nmid \infty} c_v \prod_{v|\infty} \Omega_v N(\mathfrak{u})}.$$

7.8. **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, $(\Delta(E)) = \mathfrak{u}^{12}\mathfrak{d}(E)$ for some fractional ideal $\mathfrak{u}$ whose valuation at each prime $\mathfrak{p}$ is $1/12$ of the difference between the valuation at $\mathfrak{p}$ of $\Delta(E)$ and that of the local minimal model at $\mathfrak{p}$.

The total contribution from the infinite places is now

$$N(\mathfrak{u}) \prod_{v|\infty} \Omega_v = \left| \frac{N(\Delta(E))}{N(\mathfrak{d}(E))} \right|^{1/12} \prod_{v|\infty} \Omega_v,$$

where $N(\cdot)$ denotes the norm of a fractional ideal.

7.9. **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}, but the individual Ω_v values are not accessible to the user.

7.10. **SageMath code:**

```
[EK.period_lattice(v).omega() for v in K.places()]
```

gives a list of the Ω_v for all infinite places, real and complex, including the factor of 2 for real places where the discriminant is positive, but **not** the factor of 2 at complex places. So to obtain $\prod_{v|\infty} \Omega_v$ one needs the following:

```
prod([EK.period_lattice(v).omega() for v in K.places()]) * 2**(K.signature()[1])
```

8. THE FULL FORMULA

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

9. COMPARISONS

9.1. **Tate.** 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 $|\mu| = |d_K|^{1/2}/2^{r_2}$. So this accounts for both our "field factor" $|d_K|^{1/2}$ and also the doubling of the complex periods.

9.2. **Gross.** Gross states the formula in the form L-value$= M(A)R(A)h(A)$ where $h(A) = |\text{Ш}|$. 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 factors match our Mordell-Weil factor and Local Factors respectively:

$R(E) = R(P_1,\ldots,R_r)/|T|^2$ for elliptic curves E, adjusted as before for general dimension.

$M(A) = M_\infty(A)M_f(A)$, where (in our notation)

$$M_\infty(A) = \prod_{v|\infty} \Omega_v \cdot N(\mathfrak{u}) \cdot |d_K|^{-1/2},$$

and $M_f(A)$ is the Tamagawa product $\prod_{v\nmid\infty} c_v$.

9.3. **Dokchitser.** On pages 3-5 he defines the regulator and L-value as we have. He says that the regulator is the determinant of the **Néron-Tate height-pairing** and the prefix "Néron-Tate" indicates that the height is relative to K and not the normalized absolute one.

He 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

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

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 $N(\mathfrak{u})$ 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.

Note that $C_{E/K}$ as defined here is independent of the model of E.

10. References

1. J. Tate: On the conjectures of Birch and Swinnerton-Dyer and a geometric Analog. Séminaire N. Bourbaki, 1966, exp. no 306, p. 415-440.
2. B. Gross
3. T. Dokchitser: Notes on the parity conjecture [https://arxiv.org/abs/1009.5389]
4. MathOverFlow [https://mathoverflow.net/questions/139575/bsd-conjecture-for-x-017]