Notes on the Gross-Zagier formula: motivation and outline of proof

The idea of these notes is first to locate the Gross-Zagier formula in number theory, i.e. show why we care about it, and second to give a vague outline of the proof, both for my own benefit. Section 1 is largely based on the lecture notes of Chao Li [1], while section 2 draws heavily on Andrew Snowden's introduction [2].

1. Motivation

We start with the BSD, which needs no motivation; we'll work over $\mathbb{Q}$ for simplicity.

Conjecture (Birch–Swinnerton-Dyer)**.** *Let E be an elliptic curve over $\mathbb{Q}$, with L-function $L(E,s)$, and let $r_{\text{alg}} = \operatorname{rank} E(\mathbb{Q})$ and $r_{\text{an}} = \operatorname{ord}_{s=1} L(E,s)$ be the algebraic and analytic ranks of E. Then $r_{\text{alg}} = r_{\text{an}} =: r$ and the leading coefficient of the L-function at $s = 1$ is*

$$L^{(r)}(E,1) \sim R(E)\Omega(E)$$

where $R(E) = \det(\langle P_i, P_j\rangle)$ for $\{P_i\}$ a basis for the free part of $E(\mathbb{Q})$ and $\langle\cdot,\cdot\rangle$ the Néron-Tate pairing on $E(\mathbb{Q})$, $\Omega(E) = \int_{E(\mathbb{R})} \omega_E$ is the Néron period for ω_E the Néron differential, and $\sim$ denotes that the two sides are equal up to multiplication by some nonzero rational number.

The part of this theorem that I want to focus on is the case when $r_{\text{an}} = 1$. In particular we want to show that if $r_{\text{an}} = 1$ then $r_{\text{alg}} \geq 1$, or equivalently there is some point in $E(\mathbb{Q})$ of infinite order, and $L'(E,1) \sim R(E)\Omega(E)$. This is proved using the Gross-Zagier formula.

Let's first introduce some notation: let N be the conductor of E, and let K be an imaginary quadratic number field of discriminant $d_K < 0$ with ring of integers $\mathcal{O}_K$, with E_K the base change of E to K. Let f be the newform associated to E, $\varphi : X_0(N) \to E$ be a modular parametrization, and ω be a differential form on E such that $\varphi^*\omega = 2\pi i f(z)\,dz$.

The points of $X_0(N)(\mathbb{C})$ are cyclic N-isogenies $E' \to E''$ defined over $\mathbb{C}$, and each elliptic curve can be viewed as a quotient of $\mathbb{C}$ by a lattice, i.e. $E'(\mathbb{C}) = \mathbb{C}/\Lambda'$ and $E''(\mathbb{C}) = \mathbb{C}/\Lambda''$. Each elliptic curve has complex multiplication by $\mathcal{O}_K$ if and only if the corresponding lattice is a fractional ideal of K, and the kernel of this isogeny is Λ'/Λ'' which is then an ideal of $\mathcal{O}_K$. For this to be a cyclic N-isogeny, we need to have $\mathcal{O}_K/(\Lambda'/\Lambda'') \simeq \mathbb{Z}/N\mathbb{Z}$; since we can choose the lattices Λ', Λ'' arbitrarily we see there is a one-to-one correspondence between points $(E' \to E'') \in X_0(N)(\mathbb{C})$ with complex multiplication by $\mathcal{O}_K$ and ideals I of $\mathcal{O}_K$ such that $\mathcal{O}_K/I \simeq \mathbb{Z}/N\mathbb{Z}$.

Letting $N = p_1^{e_1} \cdots p_m^{e_m}$, by the Chinese remainder theorem we have $\mathbb{Z}/N\mathbb{Z} \simeq \bigoplus_i \mathbb{Z}/p_i^{e_i}\mathbb{Z}$, and so such ideals I correspond to tuples (I_i) such that for each i we have $\mathcal{O}_K/I_i \simeq \mathbb{Z}/p_i^{e_i}\mathbb{Z}$. This justifies our requirement on K that every prime dividing n split in K: in this case we can choose a prime $\mathfrak{p}_i$ above each p_i and it will satisfy $\mathcal{O}_K/\mathfrak{p}_i \simeq \mathbb{Z}/p_i\mathbb{Z}$ and so $\mathcal{O}_K/\mathfrak{p}_i^{e_i} \simeq \mathbb{Z}/p_i^{e_i}\mathbb{Z}$. This involves choosing one of the two primes over p_i, each e_i times, and so if $m' = \sum_i e_i$ then there are $2^{m'}$ choices of such an ideal I; and each I together with a choice of a class in $\operatorname{Cl} K$ determines the two elliptic curves, and so the total number of such points is $2^s|\operatorname{Cl} K|$.

Given such a point $x_K \in X_0(N)$ with complex multiplication by $\mathcal{O}_K$, by the theory of complex multiplication it is defined over the Hilbert class field H_K of K. By the modular parametrization, each $x_k \in X_0(N)(H_K)$ gives a point $\varphi(x_k) \in E(H_K)$, which we can trace down to a point

$$y_K := \sum_{\sigma \in \mathrm{Gal}(H_K/K)} \sigma(\varphi(x_K)) \in E(K),$$

where the sum is taken in the abelian group $E(K)$. Note that since $\mathrm{Gal}(H_K/K) \simeq \mathrm{Cl}\, K$, our point y_K is independent of the choice of class we made above in $\mathrm{Cl}\, K$ to get x_K, though it still depends on the choice of the ideal I; therefore there are $2^{m'}$ choices for y_K. We call such y_K a Heegner point on E; fix some such point.

Theorem (Gross-Zagier)**.** *We have $L(E_K, 1) = 0$ and*

$$L'(E_K, 1) = \frac{\int_{E(\mathbb{C})} \omega \wedge \overline{i\omega}}{\sqrt{-d_K}|\mathcal{O}_K^\times/\{\pm 1\}|^2} \langle y_K, y_K \rangle .$$

This has the following immediate corollary.

Corollary. *Suppose that $L'(E_K, 1) \neq 0$. Then there is a point of infinite order on E_K.*

Proof. If $L'(E_K, 1) \neq 0$, then by the Gross-Zagier formula $\langle y_K, y_K \rangle = \hat{h}(y_K) \neq 0$ where $\hat{h}$ is the canonical height; and since this can be written as $\hat{h}(y_K) = \lim_{n\to\infty} \frac{h(ny_K)}{n^2}$ for the naive height h, if y_K is torsion then $h(ny_K)$ is bounded and so $\hat{h}(y_K) = 0$, so we conclude that y_K must have infinite order. □

Lemma. *Let $E^{(K)}$ be the quadratic twist of E by K. Then $L(E_K, s) = L(E, s)L(E^{(K)}, s)$.*

Proof. Let $\chi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(K/\mathbb{Q}) \simeq \{\pm 1\}$ be the Galois character given by restriction to K, and let p be a prime. The action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $E^{(K)}(\overline{\mathbb{Q}})[p]$ is given by that on $E(\overline{\mathbb{Q}})[p]$ twisted by χ. In particular if $\chi(\mathrm{Frob}_p) = 1$ then $\det(1 - p^{-s} \mathrm{Frob}_p | E^{(K)}(\overline{\mathbb{Q}})[p]) = \det(1 - p^{-s} \mathrm{Frob}_p | E(\overline{\mathbb{Q}})[p])$ and if $\chi(\mathrm{Frob}_p) = -1$ then $\det(1 - p^{-s} \mathrm{Frob}_p | E^{(K)}(\overline{\mathbb{Q}})[p]) = \det(1 + p^{-s} \mathrm{Frob}_p | E(\overline{\mathbb{Q}})[p])$. On the other hand we have $\chi(\mathrm{Frob}_p) = 1$ if and only if p splits in K, while $\chi(\mathrm{Frob}_p) = -1$ if p is inert. Thus if p splits then the product of the local factors is just

$$\det(1 - p^{-s} \mathrm{Frob}_p | E(\overline{\mathbb{Q}})[p])^2$$

while if p is inert then it is

$$\det(1 - p^{-s} \mathrm{Frob}_p | E(\overline{\mathbb{Q}})[p]) \det(1 + p^{-s} \mathrm{Frob}_p | E(\overline{\mathbb{Q}})[p]) = \det(1 - p^{-2s} \mathrm{Frob}_p | E(\overline{\mathbb{Q}})[p]),$$

and taking the product over all p we get exactly $L(E_K, s)$. □

Corollary. *If the analytic rank r_{an} of E is 1, then $r_{\mathrm{alg}} \geq 1$.*

Proof. Let $E^{(K)}$ be the quadratic twist of E by K. By a theorem of Waldspurger we can choose K such that $L(E^{(K)}, 1) \neq 0$, so since $r_{\mathrm{an}} = 0$ we have $\mathrm{ord}_{s=1} L(E_K, s) = \mathrm{ord}_{s=1} L(E, s) + \mathrm{ord}_{s=1} L(E^{(K)}, s) = 1 + 0 = 1$, and so $L'(E_K, 1) \neq 0$. Therefore by the above y_K is a point of infinite order on E_K. Let c be the unique nontrivial automorphism of

K fixing $\mathbb{Q}$, i.e. complex conjugation. Recall that the completed L-function $\Lambda(E,s)$ satisfies $\operatorname{ord}_{s=1}\Lambda(E,s) = \operatorname{ord}_{s=1} L(E,s)$ and $\Lambda(E,s) = \epsilon\Lambda(E, 2-s)$ for $\epsilon \in \{\pm 1\}$; in this case we have $\epsilon = -1$, since if $\epsilon = 1$ then $\Lambda(E, s-1)$ would be an even function of s and therefore would have even order at $s = 1$. Recall that

$$y_K = \sum_{\sigma \in \mathrm{Gal}(H_K/K)} \sigma(\varphi(x_K)),$$

so that

$$c(y_K) = \sum_{\sigma \in \mathrm{Gal}(H_K/K)} c(\sigma(\varphi(x_K))).$$

Complex conjugation acts on $\mathrm{Gal}(H_K/K)$ by inversion: indeed, if $\mathfrak{a}$ is a class in $\mathrm{Cl}\, K \simeq \mathrm{Gal}(H_K/K)$, then $\bar{\mathfrak{a}}\mathfrak{a} = (\mathrm{N}(\mathfrak{a}))$ is principal and so $\bar{\mathfrak{a}} = \mathfrak{a}^{-1}$, or in other words $c \circ \sigma = \sigma^{-1} c$. Therefore this is

$$\sum_{\sigma \in \mathrm{Gal}(H_K/K)} \sigma^{-1}(c(\varphi(x_K))) = \sum_{\sigma \in \mathrm{Gal}(H_K/K)} \sigma(c(\varphi(x_K)))$$

by permuting the σ. Complex conjugation acts on $\varphi(x_K)$ (up to torsion) by $-\epsilon$ since ϵ is the eigenvalue of the Atkin-Lehner operator on f, and since $\epsilon = -1$ we conclude $c(y_K) = y_K$ up to torsion and so there exists some torsion point $z \in E(K)$ such that $y_K - z \in E(\mathbb{Q})$. Since y_K has infinite order and z is torsion $y_K - z$ also has infinite order, and so $r_{\mathrm{alg}} \geq 1$. □

We can also prove the formula part of the BSD in this case (up to a rational factor). First, we need to prove that the formula holds when $r_{\mathrm{an}} = 0$; in this case, $R(E) = 1$ since the free part of $E(\mathbb{Q})$ is trivial, so this is just the statement that $L(E,1)$ is a nonzero rational multiple of $\Omega(E)$.

Theorem (Birch)**.** *When $r_{\mathrm{an}} = 0$, we have*

$$L(E,1) \sim \Omega(E).$$

Proof. Using the modular parametrization $\varphi : X_0(N) \to E$, we can pull back the Néron differential to get a rational multiple of $2\pi i f(z)\, dz$, since there is some holomorphic differential ω on E which pulls back to this form and holomorphic differentials on $E/\mathbb{Q}$ are unique up to scaling by a rational. Therefore $\Omega(E)$ is a rational multiple of the integral of $2\pi i f(z)\, dz$, which is just the L-function of f evaluated at 1; but since f corresponds to E we have $L(f,s) = L(E,s)$ and so $L(E,1) \sim \Omega(E)$. □

Corollary. *If $r_{\mathrm{an}} = 1$, then*

$$L'(E,1) \sim R(E)\Omega(E).$$

This result also requires the input of Kolyvagin's Euler system to show that $r_{\mathrm{alg}} \leq 1$, which is beyond the scope of these notes, so we assume this result.

Proof. Differentiating the equation $L(E_K, s) = L(E,s)L(E^{(K)}, s)$ and evaluating at 1, we get $L'(E_K, 1) = L'(E,1)L(E^{(K)}, 1)$ since $r_{\mathrm{an}} = 1$ and so $L(E,1) = 0$. Choosing K as above so that $L(E^{(K)}, 1) \neq 0$, applying the above we have $L(E^{(K)}, 1) \sim \Omega(E^{(K)})$. Suppose that E

has defining equation $y^2 = x^3 + ax + b$ (as we may, since we are in characteristic 0); then the quadratic twist has defining equation $d_K y^2 = x^3 + ax + b$, and the corresponding Néron differentials are

$$\omega_E = \frac{dx}{2y}, \qquad \omega_{E^{(K)}} = \frac{dx}{2y\sqrt{|d_K|}}$$

so that since $\omega \sim \omega_E$ we have $\int_{E(\mathbb{C})} \omega \wedge \overline{i\omega} \sim \Omega(E)\Omega(E^{(K)})\sqrt{-d_K}$. Therefore by the Gross-Zagier formula this gives

$$L(E,1)\Omega(E^{(K)}) \sim \frac{\int_{E(\mathbb{C})} \omega \wedge \overline{i\omega}}{\sqrt{-d_K}} \langle y_K, y_K \rangle \sim \Omega(E)\Omega(E^{(K)}) \langle y_K, y_K \rangle .$$

Since $L(E^{(K)},1) \sim \Omega(E^{(K)})$ is nonzero, the result follows if $r_{\text{alg}} = 1$, since then y_K is a generator for $E(\mathbb{Q})/E(\mathbb{Q})_{\text{tor}}$ (since $\langle y_K, y_K \rangle \neq 0$, since $L'(E_K,1) = L'(E,1)L(E^{(K)},1) \neq 0$ by assumption) and so $R(E) = \langle y_K, y_K \rangle$. From above we know that $r_{\text{alg}} \geq 1$; by our assumption that $r_{\text{alg}} \leq 1$ we can conclude. □

2. Proof sketch

We want to rewrite both sides of the Gross-Zagier formula in terms of modular forms. We can first observe that

$$\int_{E(\mathbb{C})} \omega \wedge \overline{i\omega} = \frac{1}{\deg \varphi} \int_{X_0(N)(\mathbb{C})} (\varphi^*\omega) \wedge \overline{i(\varphi^*\omega)} = \frac{1}{\deg \varphi} \int_{X_0(N)(\mathbb{C})} 4\pi^2 f(z)\overline{f}(z)\, dz \wedge d\bar{z}$$

is the (appropriately normalized) Petersson inner product $\frac{1}{\deg \varphi}(f,f)$. Thus if we define two functions on the set of eigenforms f of level N by

$$\mu(f) = L'(E_K, s), \qquad \nu(f) = \frac{1}{\deg \varphi \sqrt{-d_K} |\mathcal{O}_K^\times / \{\pm 1\}|^2} (f,f) \langle y_K, y_K \rangle$$

where E is the elliptic curve over $\mathbb{Q}$ associated to f and y_K is a Heegner point on E, then the Gross-Zagier formula is equivalent to

$$\mu(f) = \nu(f).$$

We can extend these by linearity to the space of newforms of level N, since the eigenforms form a basis. Since the Petersson product $(\cdot,\cdot)$ is nondegenerate, any linear function on this space is represented by some cusp form, and so we can find cusp forms F and G, unique up to oldforms, such that

$$\mu(f) = (f, F), \qquad \nu(f) = (f, G)$$

for every f. Thus it suffices to show that $F = G$ up to oldforms, i.e. if $F = \sum_n \alpha_n q^n$ and $G = \sum_n \beta_n q^n$ then for every n coprime to N we have $\alpha_n = \beta_n$. To prove the Gross-Zagier formula, the idea is then to compute the Fourier coefficients of F and G and check that they are equal in this case.

First, we look at F: suppose that $f = \sum_n a_n q^n$ is an eigenform, corresponding to the elliptic curve E. We have

$$L(E_K, s) = \prod_{\mathfrak{p}} \frac{1}{1 - a_{\mathrm{N}(\mathfrak{p})}\,\mathrm{N}(\mathfrak{p})^{-s} + \mathrm{N}(\mathfrak{p})^{1-2s}} = \sum_{\mathfrak{n}} a_{\mathrm{N}(\mathfrak{n})}\,\mathrm{N}(\mathfrak{n})^{-s}$$

where $\mathfrak{p}$ ranges over the primes of K, $\mathfrak{n}$ ranges over the nonzero ideals of $\mathcal{O}_K$, and N is the norm function from ideals of $\mathcal{O}_K$ to $\mathbb{Z}_{\geq 0}$. Letting $c_K(n)$ be the number of ideals of $\mathcal{O}_K$ with norm n, this is

$$\sum_n c_K(n) a_n n^{-s} = \int_0^\infty y^{s-1} \sum_n c_K(n) a_n e^{-2\pi n y}\,dy$$

by the usual Mellin transform argument. To evaluate the inner sum, set $\theta = \sum_n c_K(n) q^n$; then we have

$$f\bar{\theta} = \sum_{m,n} a_m c_K(n) q^m \bar{q}^n$$

and so evaluating at $z = x + iy$, so that $q = e^{2\pi i z} = e^{2\pi i x - 2\pi y}$, and integrating with respect to x gives

$$\begin{aligned}
\int_0^1 f(x+iy)\overline{\theta(x+iy)}\,dx &= \int_0^1 \sum_{m,n} a_m c_K(n) e^{2\pi i x(m-n)} e^{-2\pi y(m+n)}\,dx \\
&= \sum_{m,n} a_m c_K(n) e^{-2\pi(m+n)y} \int_0^1 e^{2\pi i(m-n)x}\,dx \\
&= \sum_n a_n c_K(n) e^{-4\pi n y}.
\end{aligned}$$

(Snowden [2] has $e^{-2\pi n y}$ instead of $e^{-4\pi n y}$; I'm not sure of the source of the discrepancy.)

Let Γ_∞ be the subgroup of $\Gamma_0(N)$ generated by $\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$. If $E_s(z)$ is the Eisenstein series

$$E_s(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \mathrm{Im}(\gamma \cdot z)^{s-1},$$

then we have

$$f(z)\overline{\theta(z) E_{\bar{s}}(z)} = f(z)\overline{\theta(z)} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} (\overline{\gamma \cdot z})^{s-1}$$

and so integrating over $\Gamma_0(N)\backslash H$ where H is the upper half-plane gives

$$\int_{\Gamma_\infty \backslash H} f(z)\overline{\theta(z)} y^{s-1}\,dz = \int_0^\infty \int_0^1 f(x+iy)\overline{\theta(x+iy)} y^{s-1}\,dx\,dy$$

since $\Gamma_\infty\backslash H$ is just the rectangle $[0,1) \times (0,\infty)$. By the above this is precisely

$$L(E_K, s) = \int_{\Gamma_0(N)\backslash H} f(z)\overline{\theta(z) E_{\bar{s}}(z)}\,dz.$$

Differentiating with respect to s and evaluating at $s = 1$ gives

$$L'(E_K, 1) = \int_{\Gamma_0(N)\backslash H} f(z)\overline{\theta(z)} \frac{d}{ds}\bigg|_{s=1} \overline{E_{\bar{s}}(z)}\, dz,$$

and so defining $\tilde{E}(z)$ to be the holomorphic part of $\frac{d}{ds}\big|_{s=1} E_{\bar{s}}(z)$ we have

$$L'(E_K, 1) = (f, \theta\tilde{E})$$

and so $F = \theta\tilde{E}$. We can then compute the Fourier coefficients of $\tilde{E}$ from those of E_s and then of $\theta\tilde{E}$ from those of θ and $\tilde{E}$ explicitly.

In fact, the above is not quite right: it looks like we should replace θ by the signed sum of the θ_σ for each class $\sigma \in \mathrm{Cl}\, K$, in which we restrict to the ideals in that class. Then in the end F will be a signed sum of the $G_\sigma = \theta_\sigma\tilde{E}$.

Next, look at G: fix a basis $\{f_i\}$ of eigenforms, and let y_i be the image of the Heegner point x_K on the elliptic curve E_i corresponding to f_i. Then for any eigenform $f = \sum_i \frac{(f,f_i)}{(f_i,f_i)} f_i$ we have

$$(f, G) = \nu(f) = \sum_i \nu(f_i)\frac{(f, f_i)}{(f_i, f_i)};$$

forgetting about the various constant factors, we have $\nu(f_i) = (f_i, f_i)\langle y_i, y_i\rangle = (f_i, f_i)\langle y_i, y_i\rangle$, and so this is

$$(f, G) = \nu(f) = \sum_i \langle y_i, y_i\rangle (f, f_i)$$

and so

$$G = \sum_i \langle y_i, y_i\rangle f_i.$$

Therefore if $G = \sum_n \beta_n q^n$ and $f_i = \sum_n a_n^i q^n$ then

$$\beta_n = \sum_i \langle y_i, y_i\rangle a_n^i.$$

Let T_n be a Hecke operator. Since each f_i is an eigenform, we have $T_n f_i = a_n^i f_i$. The action of the Hecke algebra on the space of modular forms, which correspond to holomorphic differentials on $X_0(N)$, yields an action on the Jacobian $J_0(N) \simeq \bigoplus_i E_i$; again since the f_i are eigenforms, this decomposition again makes the action of the Hecke algebra diagonal, so that T_n acts on each E_i by multiplication by a_n^i. Therefore in particular $T_n y_i = a_n^i y_i$ and so by bilinearity $\langle y_i, T_n y_i\rangle = \langle y_i, y_i\rangle a_n^i$, and so

$$\beta_n = \sum_i \langle y_i, T_n y_i\rangle .$$

We can view the y_i as coming from x_K by embedding $X_0(N) \hookrightarrow J_0(N) \simeq \bigoplus_i E_i$ via the Abel-Jacobi map and taking y_i to be the projection to the ith factor; thus by orthonormality

$$\beta_n = \sum_i \langle y_i, T_n y_i\rangle = \left\langle \sum_i y_i, T_n \sum_i y_i \right\rangle = \langle x_K, T_n x_K\rangle ,$$

the Néron-Tate pairing on $J_0(N)$; under the Abel-Jacobi map this $(x_K) - (\infty)$, and so we can also regard this as the Néron-Tate pairing $\langle (x_K) - (\infty), T_n((x_K) - (\infty)) \rangle$ on $X_0(N)$. The global height pairing decomposes as a sum of local height pairings, and so it suffices to compute these, of which there are two classes, archimedean and nonarchimedean. In the archimedean case, the pairing is given by the solution to a certain differential equation, which can be solved explicitly and used to write down the local pairing in terms of counting ideals of $\mathcal{O}_K$ satisfying a resulting condition; in the nonarchimedean case, the pairing can be restated in terms of intersections of certain divisors on $X_0(N)$, which can be reduced to a question of endomorphisms of the elliptic curves corresponding to x_K at supersingular primes. These are quaternion algebras, and we can work out the formulas explicitly in them.

Again, it should be noted that rather than working with $\langle x_K, T_n x_K \rangle$ we should rather twist the second factor by some automorphism $\sigma \in \mathrm{Gal}(H_K/K)$ and sum; we then match the components with the G_σ for $\sigma \in \mathrm{Cl}\, K$ by the isomorphism $\mathrm{Gal}(H_K/K) \simeq \mathrm{Cl}\, K$.

References

[1] Chao Li. Arithmetic of L-functions: lecture notes (taken by Pak-Hin Lee). `http://www.math.columbia.edu/~phlee/CourseNotes/L-functions.pdf`, 2018.

[2] Andrew Snowden. Gross-Zagier reading seminar: introduction. `http://www-personal.umich.edu/~asnowden/notes/gz/L01.pdf`, 2014.