Provable Computation of Motivic L-functions

Robert W. Bradshaw

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2010

Program Authorized to Offer Degree: Department of Mathematics

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Robert W. Bradshaw

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

William A. Stein

Reading Committee:

William A. Stein

Ralph Greenberg

Neal I. Koblitz

Date:

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with “fair use” as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted “the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform.”

Signature______________________________

Date__________________________________

University of Washington

Abstract

Provable Computation of Motivic L-functions

Robert W. Bradshaw

Chair of the Supervisory Committee:
Professor William A. Stein
Mathematics

L-functions have been a central object of study in number theory ever since the discovery of the Riemann zeta function, and are still an area of active research. The behavior of the L-function at specific points called special values often gives algebraic information about the object to which it is attached. We give an algorithm to provably compute values and derivatives of L-functions at arbitrary points on the complex plane using their functional equations, and give several applications of this algorithm, in particular computing Heegner points and investigations of the Birch and Swinnerton-Dyer conjecture.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

ACKNOWLEDGMENTS

Firstly, I owe my deepest gratitude to my advisor, William Stein, for generously and freely sharing his time, insight, and ideas with me, and most of all introducing me to and helping me navigate many beautiful and deep topics in number theory. His seemingly boundless energy, enthusiasm, and encouragement have been a great source of inspiration for me.

It has been a pleasure to work with and learn from my peers and the faculty of here at the University of Washington, especially Ralph Greenberg, Robert Miller, Tom Boothby, and Craig Citro. I have also benefited greatly from many mathematicians from all across the globe that I've had the opportunity to collaborate with and learn from, most notably Kiran Kedlaya and John Cremona.

I am grateful for the funding I have received throughout my time here, from the department, from the NSF through several of my advisors grants, and from the wonderful women of the Seattle chapter of ARCS.

I am overflowing with gratitude to my family, and especially my wife, Camille, for their understanding even when they could not understand what I was working on, and their confidence in me even, and perhaps especially, when I lacked confidence in myself. I would not have gotten to this point without their support.

Lastly, I would like to show my gratitude to my Creator, who gave me the talents, opportunity, and aid needed for this work as well as in all other areas of my life.

DEDICATION

to my loving and patient wife, Camille

Chapter 1

INTRODUCTION

L-functions have been a central object of study in number theory ever since the discovery of the Riemann zeta function, and are still an area of active research. They come up in the ancient (and still unsolved) congruent number problem, play an integral role in many classical results such as the infinitude of primes in arithmetic progressions, and underlie the deep connection between elliptic curves and modular forms that was used to prove Fermat's last Theorem. Two of the Clay Math Institutes' millennium problems—the Rieman Hypothesis and the Birch and Swinnerton-Dyer Conjecture—are direct questions about the behavior of certain L-functions.

L-functions are complex analytic functions on the complex plane that encode rich information about algebraic objects. They are usually given by a Dirichlet series

$$L(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

where the a_n have some local, arithmetic significance, such as counting points or prime ideals. Though the series only converges for sufficiently large $\operatorname{Re} s$, it has a (possibly conjectural) analytic continuation to the entire complex plane (except for sometimes a single pole). The behavior of the L-function at specific points called special values often gives global information about the object to which it is attached. A concrete example of this in the case of elliptic curves will be given in Chapter 2. The main contribution of this thesis is a algorithm for *provably* evaluating these L-functions at arbitrary points in the complex plane, along with specific applications.

The structure of this dissertation is as follows: the rest of this chapter will be devoted to summarizing basic properties of L-functions. In Chapter 2 we investigate the Birch and Swinnerton-Dyer conjecture, which has been the primary motivating example of this work.

Chapter 3 gives the basic outline of how to compute L-functions, which we make rigorous in Chapter 4. Finally, Chapters 5 and 6 give some applications and examples.

1.1 Euler products and functional equations

If the Dirichlet coefficients of an L-function are multiplicative, than the series has an Euler product factorization

$$L(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} = \prod_{p \text{ prime}} \left(1 + \frac{a_p}{p^s} + \frac{a_{p^2}}{p^{2s}} + \cdots\right).$$

For example, given a Dirichlet character χ, the Dirichlet L-function $L(\chi, s)$ factors as

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - \chi(p)p^{-s}}.$$

Though this may seem like a strict requirement, as we will see in the next section, most of the L-functions we are interested in studying naturally arise as Euler products. These Euler products are what allows one to study the L functions (and the objects to which they are attached) locally as well as globally.

L-functions are also expected to satisfy a functional equation, relating values on a right half plane with those on a left. Specifically, one defines the completed L-function $\Lambda(s) = L_\infty(s)L(s)$, where $L_\infty(s)$ typically consists of gamma and exponential factors and can be seen as the factor of the Euler product corresponding to the infinite prime. Then one has the relation

$$\Lambda(s) = \epsilon\Lambda(w - s)$$

for some real w (usually an integer) and complex (usually a root of unity) ϵ. (In general, $\Lambda(s)$ may be related to a dual L-function's $\hat{\Lambda}$ rather than to itself.) This gives the L-function a kind of a symmetry about the line $\operatorname{Re} s = \frac{w}{2}$, which is called the critical line. (Some people always normalize the L-function such that $w = 1$, which we will not require in this thesis.) As a specific example, the completed zeta function $\Lambda(s) = \pi^{s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$ satisfies the functional equation

$$\Lambda(s) = \Lambda(1-s)$$

and has the critial line $\operatorname{Re} s = \frac{1}{2}$. The well-know Riemann hypothesis is that all the non-trivial zeros of $\zeta(s)$ lie on this line—the generalized Riemann hypothesis is the conjecture that similar behaviors are also expected of more general L-functions.

1.2 Langlands Program

The development of class field theory gave rise to generalizations of the Dirichlet L-functions useful for studying finite extensions of number fields, and starts to introduce some ideas from representation theory. Characters are no more than one-dimensional representations, and indeed one-dimensional representations of finite, abelian extensions of $\mathbb{Q}$ are in correspondence with Dirichlet L-functions (Kronecker-Weber theorem). Hecke L-functions arise when one considering finite abelian extensions of any number field (specializing to Dirichlet L-functions when the ground field is $\mathbb{Q}$). Artin L-functions take things yet further, allowing higher-dimensional representations of non-abelian extensions. Given a finite Galois extension of number fields K/F with Galois group G, and a representation on a finite dimensional vector space $\rho : G \to V$, the Artin L-function is defined as an Euler product

$$L(\rho, s) = \prod_{\mathfrak{p} \text{ prime}} \frac{1}{\det(I - \rho(\mathrm{Frob}_{\mathfrak{p}})N(\mathfrak{p})^{-s})}.$$

(At ramified $\mathfrak{p}$, the representation is restricted to the part of V fixed by the inertia group.)

Meanwhile, L-functions were being attached to other objects, such as algebraic curves and automorphic forms. The common thread in these constructions is that they all involved certain Galois groups acting on specific vector spaces, using the same construction of setting the Euler products to be the characteristic polynomials of the action of Frobenious on these spaces.

Seeking to unite all these diverse viewpoints in a single generalization, Langlands proposed that all these L-functions can be viewed as arising from automorphic representations, which are representations into reductive groups over rings of adeles. He went on to conjecture specific forms for the Euler factorizations, functional equations, and functorial properties of these L-functions, which should contain as special cases many of the other L-functions studied thus far. These conjectures are still open problems in full generality and actively being researched, see for example [2] and [16] for an overview of the subject.

Chapter 2

THE BIRCH AND SWINNERTON-DYER CONJECTURE

Let E be an elliptic curve defined over $\mathbb{Q}$. By the Mordell-Weil theorem, $E(\mathbb{Q})$ is a finitely generated abelian group, i.e., $E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus T$ for some torsion subgroup T and a non-negative rank r.

The elliptic curve has an associated L-function which can be defined as follows. For a prime p of good reduction, the points on the reduced curve $\tilde{E}(\mathbb{F}_p)$ forms a group, and we let $a_p = p + 1 - \#\tilde{E}(\mathbb{F}_p)$. For primes of bad reduction (i.e. those primes where $\tilde{E}(\mathbb{F}_p)$ is singular) we let a_p be 0, 1, or -1 according to whether E has additive, split multiplicative, or non-split multiplicative reduction at p. Let Δ_E be the minimal discriminant of E, which is divisible by exactly the primes of bad reduction. We can then give an L-function

$$L(E,s) = \prod_{p \mid \Delta_E} \frac{1}{1 - a_p p^{-s}} \prod_{p \nmid \Delta_E} \frac{1}{1 - a_p p^{-s} + p \cdot p^{-2s}}.$$

In the spirit of the previous chapter, we note the automorphic representation theoretic interpretation of these Euler factors: the a_p are in fact the traces of Frobenius acting on the Tate module $\varprojlim E[\ell^n]$ for a choice of prime $\ell \neq p$, hence the Euler factors come from the characteristic polynomials of this action. It is a deep theorem about the modularity of elliptic curves that this L-function extends to the entire complex plane and has a functional equation [5]. The completed L-function is $\Lambda(E,s) = \left(\frac{2\pi}{\sqrt{N}}\right)^{-s} \Gamma(s) L(E,s)$ which satisfies

$$\Lambda(E,s) = \pm\Lambda(E, 2-s)$$

where N is the conductor of the elliptic curve and the choice of sign is dependent on the specific elliptic curve.

2.1 The conjecture

Inspired by explicit computations of quantities related to L-functions of elliptic curves, Birch and Swinnerton-Dyer made the following conjecture:

Conjecture 2.1.1 (Birch and Swinnerton-Dyer)**.** *The order vanishing of $L(E,s)$ at $s=1$ is equal to the rank r of $E(\mathbb{Q})$.*

Over 40 years later, the general proof of this conjecture remains an open problem. As numerical evidence in favor of the conjecture grew, Birch and Swinnerton-Dyer were able to refine this conjecture to a statement about the leading coefficient of the Taylor series of $L(E,s)$ at $s=1$. Before introducing this refinement, we need to present some more invariants related to an elliptic curve over $\mathbb{Q}$.

2.2 The BSD formula

As above, let E be an elliptic curve defined over $\mathbb{Q}$. The rank r of E and its torsion subgroup have already been defined. Let ω be the invariant differential on E, and define the $\Omega_E = \int_{E(\mathbb{R})} \omega$, which is either the real period of E or twice the real period (according to whether $E(\mathbb{R})$ has one or two components). Let $\hat{h}$ be the Néron-Tate canonical height on E. The regulator Reg_E is defined to be the determinant of the height pairing matrix of a basis of $E(\mathbb{Q})/E(\mathbb{Q})_{tor}$. (In the case that $E(\mathbb{Q})$ has rank 0, the convention is to take the regulator to be 1.) For each prime p, the Tamagawa number c_p is the the size of the rational part of the component group $\Phi_{E,p}(\mathbb{F}_p)$, i.e., the number of connected components of the rational part of the special fiber of the Néron model of $\mathcal{E}$ of E over p. These component groups are trivial for primes of good reduction. Finally, we need to define the Shafarevich-Tate group $\text{Ш}(E)$. It is defined as the kernel of the global to local map on cohomology $\text{Ш}(E) = \ker\Big(H^1(\mathbb{Q},E) \to \prod_{p\leq\infty} H^1(\mathbb{Q}_p,E)\Big)$.

We are now ready to state the conjectural BSD formula.

Conjecture 2.2.1 (Birch and Swinnerton-Dyer formula)**.** *The leading term of the Taylor* *series of $L(E,s)$ centered at $s=1$ is given by*

$$\frac{L^{(r)}(E,1)}{r!} = \frac{\Omega_E \operatorname{Reg}_E \prod_p c_p}{\#E(\mathbb{Q})_{tor}^2} \#\text{Ш}(E).$$

The BSD formula has been proven to hold for many curves of rank 0 and 1 [17, 31], but it is still an open problem to show that it is true for even a single higher-rank curve. In fact, it is not even known that $\text{Ш}(E)$ is finite or that the ratio $L^{(r)}(E,1)\Omega_E^{-1}\operatorname{Reg}_E^{-1}$ of transcendental numbers is rational in even a single case for a curve of rank $r \geq 2$, clearly necessary conditions for the formula to hold. We can, however, give numerical evidence for this formula. Given the (possibly conjectural) rank r of E, the left hand side can be computed to high precision using, e.g., the algorithm outlined in this paper. Likewise, all the invariants on the right hand side of the formula can be computed in full generality with the exception of $\#\text{Ш}(E)$. One can then solve for the expected value of $\#\text{Ш}(E)$, denoted $\#\text{Ш}(E)_{an}$, which should be an integer (in fact, a perfect square [6]). For the rank 2 elliptic curve of conductor 389 we have verified that $|\#\text{Ш}(E)_{an} - 1| < 2^{-10000}$. Equivalently, assuming that $\text{Ш}(E)$ is trivial, we have proved that the BSD formula holds for this curve to at least 10,000 bits of precision, see §6.1.

2.3 BSD over Number Fields

Though originally only stated for elliptic curves over $\mathbb{Q}$, the BSD conjecture is believed to hold, suitably generalized, to a much larger class of objects. Later on, we will need the case where E is defined over a number field, so we state the appropriate generalizations here.

Given an elliptic curve E over a number field K, we still have the Mordell-Weil theorem and get the rank and torsion subgroup. The L-series is defined in essentially the same way as that for $K=\mathbb{Q}$, taking the product over all prime ideals of K rather than rational primes.

$$L(E/K,s) = \prod_{\mathfrak{p}|N} \frac{1}{1-a_{\mathfrak{p}}N(\mathfrak{p})^{-s}} \prod_{\mathfrak{p}\nmid N} \frac{1}{1-a_p\mathfrak{p}N(\mathfrak{p})^{-s}+N(\mathfrak{p})^{1-2s}}.$$

Again, the $a_{\mathfrak{p}}$ arise from counting points on the reduce curve, or equivalently, from the action of Frobenius on the Tate module. Likewise, the definition of the Tamagawa numbers and Shafarevich-Tate group are extended to the number field case by considering prime ideals of K rather than rational primes. The canonical height decomposes into local parts

$$\hat{h}(P) = \sum_v n_v \lambda_v(P)$$

where the n_v are the local degrees $[K_v : Q_v]$ of v. In this case there may be more than one archimedean height to consider, but they are simply the archimedean heights of each possible embedding of $E(K) \to E(\mathbb{C})$. There is another a choice of normalization–one gives a height that is constant for all embeddings $E(K) \to E(F)$ for any finite extension F of K (and can hence be defined on all of $E(\overline{\mathbb{Q}})$), and the other which is exactly $[K : \mathbb{Q}]$ times as large, and the is the correct normalization for the BSD formula over K. Finally, we come to the analogue of the real period Ω_E. Let v be a place of K. If v corresponds to a real embedding $\sigma : K \to \mathbb{R}$ we evaluate the real volume

$$\Omega_E, v = \int_{E_\sigma(\mathbb{R})} \omega_\sigma.$$

For v corresponding to a pair of complex embeddings $\sigma, \overline{\sigma}$ we have

$$\Omega_E, v = \int_{E_\sigma(\mathbb{C})} \omega_\sigma \wedge i\overline{\omega_\sigma}.$$

Let $\Omega_{E/K}$ be the the product $\prod_{v|\infty} \Omega_{E,v}$.

One again has the rank conjecture, and the conjectural formula for the leading term of the Taylor series is:

$$\frac{L^{(r)}(E/K,1)}{r!} = \frac{\Omega_{E/K}\,\mathrm{Reg}_{E/K} \prod_{\mathfrak{p}} c_{\mathfrak{p}}}{\sqrt{|\Delta_K|}\#E(K)^2_{tor}} \#\text{Ш}(E/K).$$

Chapter 3

DOKCHITSER'S METHOD

Let

$$L(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

be a Dirichlet series. We make the following assumptions on $L(s)$:

- $L(s)$ converges on some right half plane. Equivalently, a_n grows at most polynomially in n.

- $L(s)$ admits a meromorphic continuation to the entire complex plane and with a functional equation of the form

  $$\Lambda(s) = \epsilon\Lambda(w - s)$$

  where

  $$\Lambda(s) = A^s\gamma(s)L(s)$$

  for some *weight* $w \geq 0$, *sign* $\epsilon \in \mathbb{C}$, *exponential factor* $A > 0$, and $\gamma(s) = \Gamma\left(\frac{s+\lambda_1}{2}\right) \cdots \Gamma\left(\frac{s+\lambda_d}{2}\right)$ for $d \geq 1$ Hodge numbers $\lambda_1, ..., \lambda_d \in \mathbb{C}$.

- $\Lambda(s)$ has finitely many simple poles p_j with corresponding residues r_j and no other singularities.

All L-functions we are interested in studying satisfy these assumptions. (The Legendre duplication formula $\Gamma(s) = \pi^{-1/2}2^{s-1/2}\Gamma(\frac{s}{2})\Gamma(\frac{s+1}{2})$ often comes in handy for writing the gamma factors of many well known L-functions in the above form.) The requirement that the poles be simple is not essential, but simplifies the presentation. Some L-functions of non-motivic origin also satisfy the above criteria, such as Maass forms [39], and this algorithm may be applied there as well.

Given such a function, we would like to be able to evaluate it and its derivatives to (numerically) compute special values and verify functional equations. In [13], Dokchitser outlines a procedure to do this in general using Mellin Transforms, which we summarize here in the next two sections.

3.1 Formula

Using the functional equation of $\Lambda(s)$ and the theory of Mellin transforms, one can deduce a rapidly converging formula $\Lambda(s)$ for all $s \in \mathbb{C} - \{p_j\}_j$.

Theorem 3.1.1. *Let $\Lambda(s)$ be defined as above. Let $\phi(t)$ be the inverse Mellin transform of $\gamma(s)$, that is*

$$\gamma(s) = \int_0^\infty \phi(t) t^s \frac{dt}{t}.$$

Let $x^s G_s(x)$ be the incomplete Mellin transform of $\phi(t)$

$$G_s(x) = x^{-s} \int_x^\infty \phi(t) t^s \frac{dt}{t}.$$

Then

$$\Lambda(s) = \sum_{n=1}^\infty a_n G_s\left(\frac{n}{A}\right) + \epsilon \sum_{n=1}^\infty a_n G_{w-s}\left(\frac{n}{A}\right) + \sum_j \frac{r_j}{p_j - s}.$$

Moreover this series converges exponentially fast.

Proof. Let $\Theta(t)$ be the inverse Mellin transform of $\Lambda(t)$, which can be easily expressed in terms of this inverse Mellin transform of γ:

$$\Theta(t) = \sum_{n=1}^\infty a_n \phi\left(\frac{nt}{A}\right),$$

as

$$\int_0^\infty \Theta(t) t^s \frac{dt}{t} = \int_0^\infty \sum_{n=1}^\infty a_n \phi\left(\frac{nt}{A}\right) t^s \frac{dt}{t} = \sum_{n=1}^\infty a_n \int_0^\infty \phi\left(\frac{nt}{A}\right) t^s \frac{dt}{t}$$

$$= \sum_{n=1}^\infty a_n \int_0^\infty \phi\left(t'\right) \left(\frac{At'}{n}\right)^s \frac{dt'}{t'} = \sum_{n=1}^\infty a_n n^{-s} A^s \gamma(s) = \Lambda(s).$$

Mellin's inversion formula [14] tells us that, for sufficiently large c (in particular c lying to the right of all the poles of $\Lambda(s)$) we have

$$\Theta(t) = \int_{c-i\infty}^{c+i\infty} \Lambda(s) t^{-s} ds.$$

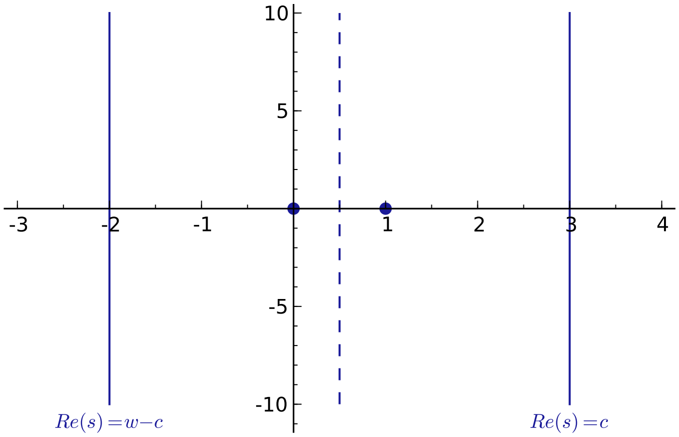
Now compute

$$\Theta(1/t) = \int_{c-i\infty}^{c+i\infty} \Lambda(s) t^{s} ds = \int_{c-i\infty}^{c+i\infty} \epsilon\Lambda(w-s) t^{s} ds = \int_{w-c-i\infty}^{w-c+i\infty} \epsilon\Lambda(s) t^{w-s} ds$$

This is almost an expression for $\epsilon t^w \Theta(t)$, only the path of integration occurs to the left of all the poles of $\Lambda(s)$. Shifting the path of integration to the right we pick up the residues of all the poles (noting that they occur in pairs about the line $\operatorname{Re} s = \frac{w}{2}$, see figure 3.1), yielding

$$\Theta(1/t) = \epsilon t^w \Theta(t) - \sum_j r_j t^{p_j}.$$

Figure 3.1: Path integrals defining $\Theta(t)$ and $\Theta(1/t)$.



(In the case that the poles of $\Lambda(s)$ are not simple, one would pick up extra $\log t$ factors in the residues.)

Using this functional equation we can write

$$\Lambda(s) = \int_0^\infty \Theta(t) t^s \frac{dt}{t} = \int_1^\infty \Theta(t) t^s \frac{dt}{t} + \int_0^1 \Theta(t) t^s \frac{dt}{t} = \int_1^\infty \Theta(t) t^s \frac{dt}{t} + \int_1^\infty \Theta(1/t) t^{-s} \frac{dt}{t}$$

$$= \int_1^\infty \Theta(t) t^s \frac{dt}{t} + \int_1^\infty \epsilon t^w \Theta(t) t^{-s} \frac{dt}{t} - \int_1^\infty \sum_j r_j t^{p_j} t^{-s} \frac{dt}{t}$$

$$= \int_1^\infty \Theta(t) t^s \frac{dt}{t} + \epsilon \int_1^\infty \Theta(t) t^{w-s} \frac{dt}{t} - \sum_j \frac{r_j}{p_j - s}.$$

Recalling the definitions of $\Theta(t)$ and $G_s(t)$ we can rewrite the first integral as

$$\int_1^\infty \Theta(t) t^s \frac{dt}{t} = \sum_{n=1}^\infty a_n \int_1^\infty \phi\left(\frac{nt}{A}\right) t^s \frac{dt}{t} = \sum_{n=1}^\infty a_n \int_{n/A}^\infty \phi(t) t^s n^{-s} A^s \frac{dt}{t} = \sum_{n=1}^\infty a_n G_s(t).$$

The second integral may be similarly rewritten, yielding the formula above. Finally observe that $\phi(t)$, and hence $G_s(t)$ decay exponentially fast. □

This formula for $\Lambda(s)$ can be differentiated term by term giving the $\frac{\partial^r}{\partial s^r}\Lambda(s)$ in terms of $\frac{\partial^r}{\partial s^r}G_s(t)$. The difficult part is accurately computing $G_s(t)$, which is considered in the next section.

3.2 Computing $G_s(x)$ and its derivatives

The first step to computing $G_s(x)$ is understanding how to compute $\phi(t)$. The function $\phi(t)$ has an expansion about zero:

$$\phi(t) = \sum_j t^{\lambda_j} p_j(t^2), \qquad p_j(t) \in \mathbb{C}[\log t][[t]].$$

The powers of $\log t$ appearing in $p_j(t)$ are bounded for each j, and the coefficients arise from a linear recurrence coming from the recurrence $\Gamma(s+1) = s\Gamma(s)$. The p_j can be computed as follows:

1. Group the λ_j into equivalence classes H_j where λ_i is equivalent to λ_j whenever $\lambda_i - \lambda_j \in 2\mathbb{Z}$. Let $m_j = 2 - \lambda_{k_j}$ where $\operatorname{Re}\lambda_{k_j} = \min_{\lambda \in H_j} \operatorname{Re}\lambda$, that is λ_{k_j} is the element of H_j with least real part.

2. Let $c_j^{(0)}(z)$ be the Taylor series of $\gamma(z + m_j)$ about $z = 0$.

3. For $1 \leq j \leq |H_j|$ and $n \geq 1$ define $c_j^{(n)}(z)$ recursively as

$$c_j^{(n)}(z) = \frac{c_j^{(n-1)}(z)}{\prod_{k=1}^{d} \left(\frac{z+\lambda_k+m_j}{2} - n \right)}$$

considered as a Laurent series in z about 0. Let $c_{j,k}^{(n)}$ denote the coefficient of z^{-k} in $c_j^{(n)}(z)$.

For positive real t the expansion of $\phi(t)$ about 0 is

$$\phi(t) = \sum_{j=1}^{N} t^{-m_j} \sum_{n=1}^{\infty} t^{2n} \sum_{k=1}^{|H_j|} \frac{(-\log t)^{k-1}}{(k-1)!} c_{j,k}^{(n)}.$$

It is easy to see that $|c_{j,k}^{(n)}| = O((n!)^{-d})$ as $n \to \infty$, so this series converges exponentially fast.

Recall that $G_s(t)$ is given by

$$G_s(x) = x^{-s} \int_x^{\infty} \phi(t) t^s \frac{dt}{t}.$$

Now $\lim_{x \to 0} x^s G_s(x) = \gamma(s)$ for s not a pole of γ. This allows us to write

$$x^s G_s(x) = \gamma(s) - \int_0^x \phi(t) t^s \frac{dt}{t}.$$

Given an expansion for $\phi(t)$ we can integrate this term-by-term to get a similar series for $G_s(x)$. Differentiating under the integral sign with respect to s allows us to compute $\frac{\partial^r}{\partial s^r} G_s(x)$ for any r as well.

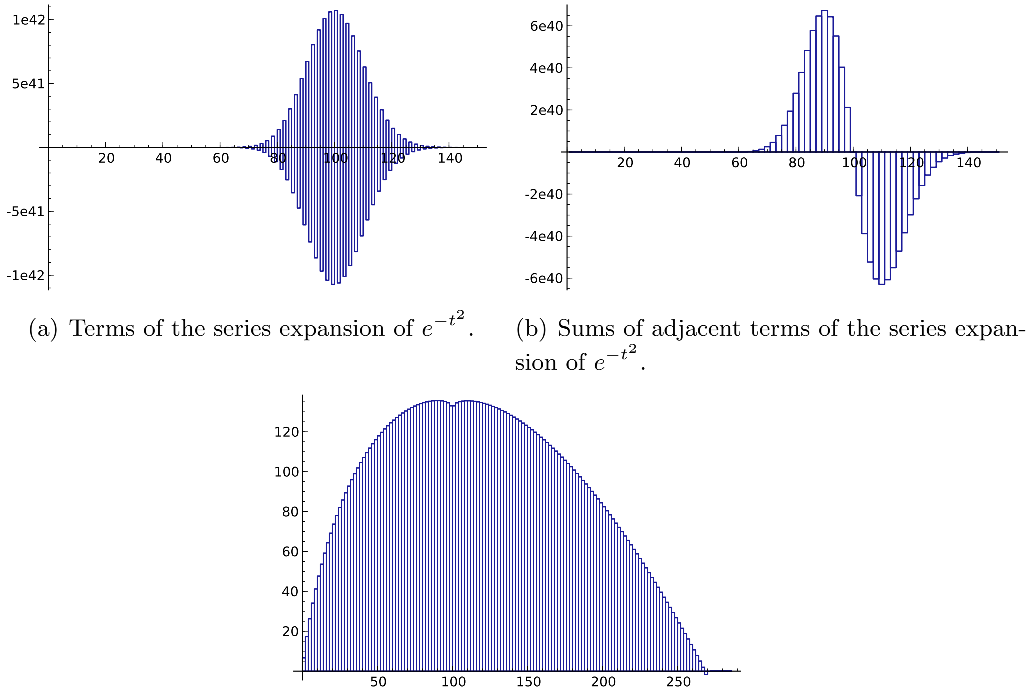
3.3 *Considerations for* $G_s(x)$ *at large* x

Unfortunately, for large values, the terms in these series tend to get very large before canceling out to what is typically a very small result. For example, the inverse Mellin transform of $\Gamma(s/2)$ is $2e^{-t^2}$ and one recovers the expected series at zero,

$$\phi(t) = \sum_{n=0}^{\infty} \frac{(-t^2)^n}{n!}.$$

To compute $\phi(10)$ to even 20 bits of absolute precision requires computing 285 terms of the sequence with intermediate precision of 140 bits, all canceling out to the final result of 0.0000.

Figure 3.2: Typical behavior of $\phi(t)$ for large t.



(a) Terms of the series expansion of e^{-t^2}.

(b) Sums of adjacent terms of the series expansion of e^{-t^2}.

(c) Bits in sums of successive terms of the series expansion of e^{-t^2}.

We can see this visually in figure 3.2. In plot (a) we see $\frac{(-t^2)^n}{n!}$ for various values of n, which grow to 10^{42} before finally decaying exponentially. As this is an alternating series, it makes sense to take advantage of the cancelation between successive terms, as illustrated in (b). The areas above and below the x-axis are still very large and nearly equal—we are interested in their miniscule difference. Finally, in (c) we plot the log of the absolute value of the values in (b), giving an idea of their magnitude throughout the whole range, and also illustrating the positive and negative region of the graph are not as symmetrical as one would like for any any kind of termwise cancelation.

In full generality, being the inverse Mellin transform of the product of several gamma factors, $\phi(t)$ is the following special case of the Meijer G-function

$$\phi(t) = 2G^{0,d}_{d,0}\left(t^2;;\frac{\lambda_j}{2}\right)$$

which is well studied [30] but not so easy to evaluate at large t. (Even commercial math packages are known to return incorrect results for these functions in some cases.) One option is to use continued fraction approximations to the asymptotic expansion of $\phi(t)$ and $G_s(x)$ at infinity, with the hopes that they converge to the actual values. This is the approach taken in Dokchitser's work. Though this seems to be work well in practice, and numerical checks can be made to raise confidence in the result, it is far from rigorous. To quote Dokchitser's original paper [13, p. 12],

> Unfortunately, it seems very difficult to provide explicit bounds for K_n. It appears that $C_n(x)$ converge rapidly to $\Psi(x)$ but to prove either "converge" or "rapidly" or "to $\Psi(x)$" in any generality seems hard. So the last step of the algorithm is based purely on empirical observations concerning the convergence of the continued fractions.

For low degree ($d \leq 2$) the functions $\phi(t)$ and, in turn, $G_s(x)$ can be recognized as specific special functions with known convergent continued fraction expansions at infinity. However, even when explicit expansions with effective bounds can be found, they are are often not tight enough for efficient, provably correct evaluation for all but the most trivial examples.

Another option is to use numerical integration to evaluate $G_s(x)$. For $c >> 0$ and $0 \leq \operatorname{Re} s < c$ we have

$$G_s(x) = \int_x^\infty \phi(t)t^s\frac{dt}{t} = \int_x^\infty \int_{c-i\infty}^{c+i\infty} t^{-u}\gamma(u)t^s du\frac{dt}{t}$$

$$= \int_{c-i\infty}^{c+i\infty} \gamma(u) \int_x^\infty t^{s+u} \frac{dt}{t} du = \int_{c-i\infty}^{c+i\infty} \gamma(u) \frac{x^{s-u}}{s+u} du$$

This integral is highly oscillatory, but this can be mitigated by introducing an auxiliary "smoothing function" $g(s)$ and computing $\Lambda(s)g(s)$ by applying inverse Mellin transform to $\gamma(s)g(s)$. This was suggested by Lagarias-Odlyzko [27] and worked out by Rubinstein [33] in the cases $d = 1$ and $d = 2$ when the Legendre duplication formula can be used to write the gamma factor as $\gamma(s) = \Gamma(s + \lambda)$ under an additional (though easily satisfied) hypothesis on the asymptotic behavior of $L(s)$ as $|\Im s| \to \infty$. Some indication is given how this may be applied to higher degree L-series in [33, §3.3.3]. This smoothing function also mitigates the cancelation effects that arise when computing $L(s)$ for s with large imaginary part, which is especially relevant when studying behavior within the critical strip. This has been implemented in [34] and works well in practice, though the error bounds are base on "estimates..., experimentation, and intuition" [33, p. 75] rather than rigorous analysis.

We are most interested in looking at special values of higher degree L-functions, so our approach is to compute the expansion about zero to whatever precision is necessary to guarantee accuracy of the resulting answer.

Chapter 4
AN ALGORITHM

As mentioned in section 6 of [13], there are several steps which need to be made rigorous to make the procedure explained above into a provably correct algorithm. The issues involved are truncating the various infinite series, controlling the rounding error, and maintaing adequate precision (especially in the case of catastrophic cancelation). The general asymptotic behavior of these series is usually easy to understand, but effective bounds are required to make the algorithms rigorous.

4.1 Real Number Representations

In order to do rigorous computation over the real field, it is necessary to understand a little bit about how real numbers are represented and operated on on a computer. There are two main issues to deal with. The first is that almost all real numbers cannot be represented exactly in a fixed finite amount of space, thus one is most often dealing with inexact approximations to a specified precision of the true values one is trying to compute. The second is that issue is errors introduced due to (repeated) rounding. As Kernighan and Plauger put it in their classic *The Elements of Programming Style*, "Floating-point numbers are a lot like sandpiles: Every time you move one you lose a little sand and pick up a little dirt." [25] From a more mathematical standpoint, the set of real numbers, as represented on a computer, do *not* satisfy the associative, distributive, and additive/multiplicative identity axioms needed to be a field.

Given an approximation $\tilde{x}$ to a real number x, one can speak of the relative and absolute precision of $\tilde{x}$. The absolute precision of $\tilde{x}$ in base b is defined to be $-\log_b|x-\tilde{x}|$, and the relative precision of $\tilde{x}$ base b is $-\log_b|1-\tilde{x}/x|$. These two notions of precision correspond to the two standard real number representations. Absolute precision behaves well with respect to addition and subtraction, and relative precision behaves well with respect to multiplication and division.

For nearly all applications, complex numbers are simply represented as pairs of real numbers indicating the imaginary and real parts.

4.1.1 Precision Models

The most common format for representing real numbers on a computer is the *floating point* representation. Numbers are represented by a mentissa of bounded size and an exponent, much like scientific notation. Nearly all modern processors have native (and very fast) support for 53-bit mentissa (11-bit exponent) arithmetic, and software libraries such as MPFR [15] exists for doing floating point computations for arbitrarily large relative precision, limited only by the physical constraints of the machine. When arithmetic is performed, the mentissa is rounded to fit into the representation. One significant drawback of this representation is that it is susceptible to catastrophic cancelation which is the drastic loss of (relative) precision when, for example, subtracting two very close numbers. For a very simple example, lets subtract $10^{12} + \pi$ and $10^{12} - \pi$ with 15 digits of precision:

$$\begin{array}{rrcr} & 10^{12}+\pi & \approx & 1000000000003.14 \\ - & 10^{12}-\pi & \approx & 999999999996.858 \\ \hline & 2\pi & \approx & 6.282 \end{array} .$$

Despite having a decently large working precision, the final result is only correct to 4 digits. Had the working precision been lower, or the terms relatively closer, we might not have been able to obtain any correct digits of the difference at all.

Another format for storing real numbers is the *fixed point* representation. In this case the exponent is fixed, and the mentissa is allowed to be an arbitrarily large integer. The main advantage of fixed point arithmetic is that it is not as susceptible to catastrophic cancelation, which is especially important when summing or integrating highly oscillatory series such as those encountered in §4.4. Of course this comes at a cost, which is that the memory and cost of doing arithmetic is no longer fixed but is (in most cases) proportional to the magnitude of the values involved. Also, fixed point arithmetic does not behave as well with respect to multiplication and division. Going back to the example above, with a fixed absolute precision of 15 digits, we have

$$\begin{array}{rrcr} & 10^{12}+\pi & \approx & 1000000000003.141592653589793 \\ - & 10^{12}-\pi & \approx & 999999999996.858407346410207 \\ \hline & 2\pi & \approx & 6.283185307179586 \end{array}$$

which is a much higher precision result.

4.1.2 Interval Arithmetic

A very useful tool in doing rigorous computations over the real field is interval arithmetic [26, §4.2.2D], [32]. Let $\overline{\mathbb{R}}$ denote the two-point compactification $\mathbb{R} \cup \{\pm\infty\}$ of $\mathbb{R}$. A real number x is represented by an interval $I_x = [x_-, x_+]$ containing x where $x_-, x_+ \in \overline{\mathbb{R}}$. (Note that $\pm\infty$ can be represented as well.) Of course such a representation is not unique, but if the diameter of I_x is small enough it often encodes enough information about x to be useful. For example, we may be able to deduce that x is non-zero, or if x is known to be an integer (perfect square, element of a number field of bounded degree and denominator, etc.) we may be able to identify x exactly. In all cases, it gives us an upper and lower bound on the true value of x. In some sense, I_x encapsulates both an approximation to x and a specific precision to which it is known.

All the ordinary arithmetic operations on $\mathbb{R}$ can be extended to operations on intervals, using the definition

$$I_x \star I_y = [\inf_{\tilde{x}\in I_x, \tilde{y}\in I_y} \tilde{x} \star \tilde{y}, \sup_{\tilde{x}\in I_x, \tilde{y}\in I_y} \tilde{x} \star \tilde{y}].$$

Clearly for any $x, y \in \mathbb{R}$ we have $x \star y \in I_x \star I_y$. This does *not* however turn the set of intervals into a field. For example, intervals with positive diameter do not have additive or multiplicative inverses, and the "distributive law" is only one of containment rather than equality.

For any function $f : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ we can define a function on intervals

$$\bar{f}(I_x) = [\inf_{\tilde{x}\in I_x} f(\tilde{x}), \sup_{\tilde{x}\in I_x} f(\tilde{x})]$$

and analogous definitions for multivariate functions. It is important to note that while we always have $f(x) \in f(I_x) \subseteq \bar{f}(I_x)$, it is often the case that $\bar{f}(I_x)$ is a proper superset of $f(I_x)$, especially as the diameter of I_x grows.

Interval arithmetic is particularly useful for doing computations on a computer. As operations on a computer can only be caried out to a finite amount of precision, the results need to be repeatedly truncated or rounded. Using intervals allows one to keep track of the possible errors rounding may introduce. For example, to compute $\bar{f}(I_x)$, the infimum and supremum are computed to finite precision, the former rounded towards $-\infty$ and the latter rounded to $+\infty$. This may result in a slightly larger interval $I_{f(x)} \supset \bar{f}(I_x)$, but the endpoints are finitely represented and we still have the guarantee that $f(x) \in I_{f(x)}$. Likewise, to add two intervals $[a, b]$ and $[c, d]$ one notes that the infimum is $a+c$ and supremum $b+d$, giving $[a, b] + [c, d] = [a + c, b + d]$. The left endpoint $a + c$ is computed rounding towards $-\infty$ and the right endpoint $b + d$ is computed rounding towards $+\infty$.

In a sense, the intervals do the bookkeeping of how much inaccuracy is introduced by only using finite approximations, and the final result after any number of operations is an interval known to contain the correct value despite all of the intermediate approximations made. This is especially helpful where explicit error analysis is infeasible or impractical, such as deducing *a priori* bounds on the errors in the coefficients $c_{j,k}^{(n)}$ obtained by (repeated) power series division. However, because intervals always account for the worst case rounding error, if the intermediate precisions are not high enough one may end up with the unhelpful interval $[-\infty, \infty]$. In practice it helps to group operations as much as possible, as one can easily have situations like $\inf_{x\in I} f(x) + \inf_{x\in I} g(x)$ being a poor bound for $\inf_{x\in I}(f(x) + g(x))$. In other words, interval arithmetic doesn't prevent loss of accuracy, it simply measures it. Returning again to our example, using an interval with 15-digit floating point entries,

$$\begin{array}{rrcrr} & 10^{12}+\pi & \in & [1000000000003.14\ \ , & 1000000000003.15\ \ ] \\ - & 10^{12}-\pi & \in & [\ 999999999996.858\ , & 999999999996.859]. \\ \hline & 2\pi & \in & [\qquad\qquad 6.281\ , & 6.292] \end{array}$$

4.2 Bounding $G_s(x)$

We now return to bounding the series in question.

An effective bound on $G_s(x)$ follows directly from such a bound on $\phi(t)$. The following proposition about the incomplete gamma function will be useful.

Proposition 4.2.1. *The upper incomplete gamma function satisfies* $\Gamma(s,x) < x^s e^{-x}$ *for all positive real numbers* $x > s+1$.

Proof. This is easy to see using the definition and the fundamental theorem of calculus.

$$\Gamma(s,x) = \int_x^\infty t^s e^{-t} \frac{dt}{t} < \int_x^\infty (t-s) t^s e^{-t} \frac{dt}{t} = \int_x^\infty \frac{d}{dt}\left(-t^s e^{-t}\right).$$

Here the condition on x is used to force $t - s > 1$ on the domain of integration. □

We can now get the following explicit bound on ϕ which, as expected, looks much like the asymptotic expansion at infinity.

Lemma 4.2.2. *Suppose there are d Hodge numbers $\lambda_1, ..., \lambda_d$. Let $\bar{\lambda}_i = \max(0, \operatorname{Re}\lambda_i)$ and $\kappa = (\bar{\lambda}_1 + \cdots + \bar{\lambda}_d)/d$. Then for $t > 1$ we have the bound $|\phi(t)| \leq 2d! t^\kappa e^{-t^{2/d}}$.*

Proof. For $d = 1$ this is clear, as the inverse Mellin transform of $\Gamma\left(\frac{s}{2}\right)$ is $2e^{-t^2}$, so by properties of the Mellin Transform [14], the inverse Mellin transform of $\Gamma\left(\frac{s+\lambda_1}{2}\right)$ is $2t^{\lambda_1}e^{-t^2}$.

Consider $d > 1$. Let $\gamma_{(d-1)}(s) = \Gamma\left(\frac{s+\lambda_1}{2}\right) \cdots \Gamma\left(\frac{s+\lambda_{d-1}}{2}\right)$ and $\phi_{(d-1)}(t)$ be the inverse Mellin transform of $\gamma_{(d-1)}(s)$. Let $\kappa' = (\bar{\lambda}_1 + \cdots \bar{\lambda}_{d-1})/(d-1)$. Using the Mellin convolution theorem [14] we have

$$\phi(t) = \int_0^\infty 2u^{\lambda_d} e^{-u^2} \phi_{(d-1)}\left(\frac{t}{u}\right) \frac{du}{u}.$$

This is as follows

$$\begin{aligned}
|\phi(x)| &\leq \left| \int_0^\infty 2u^{\lambda_d} e^{-u^2} \cdot 2(d-1)! \left(\frac{x}{u}\right)^{\kappa'} e^{-\left(\frac{x}{u}\right)^{2/(d-1)}} \frac{du}{u} \right| \\
&= \int_0^\infty 2u^{\operatorname{Re}\lambda_d} e^{-u^2} \cdot 2(d-1)! \left(\frac{x}{u}\right)^{\kappa'} e^{-\left(\frac{x}{u}\right)^{2/(d-1)}} \frac{du}{u} \\
&= 2(d-1)! \int_0^\infty x^{\kappa'} v^{(\operatorname{Re}\lambda_d - \kappa')/2} e^{-v} e^{-\left(\frac{x^2}{v}\right)^{1/(d-1)}} \frac{dv}{v} \\
&= 2(d-1)! \left(\int_0^{x^{2/d}} x^{\kappa'} v^{(\operatorname{Re}\lambda_d - \kappa')/2} e^{-v} e^{-\left(\frac{x^2}{v}\right)^{1/(d-1)}} \frac{dv}{v} \right. \\
&\qquad\qquad \left. + \int_{x^{2/d}}^\infty x^{\kappa'} v^{(\operatorname{Re}\lambda_d - \kappa')/2} e^{-v} e^{-\left(\frac{x^2}{v}\right)^{1/(d-1)}} \frac{dv}{v} \right) \\
&= 2(d-1)! \left(\int_{x^{2/d}}^\infty (d-1) x^{\operatorname{Re}\lambda_d} w^{(d-1)(\kappa' - \operatorname{Re}\lambda_d)/2} e^{-x^2 w^{-(d-1)}} e^{-w} \frac{dw}{w} \right. \\
&\qquad\qquad \left. + \int_{x^{2/d}}^\infty x^{\kappa'} v^{(\operatorname{Re}\lambda_d - \kappa')/2} e^{-v} e^{-\left(\frac{x^2}{v}\right)^{1/(d-1)}} \frac{dv}{v} \right) \\
&< 2(d-1)! \left(\int_{x^{2/d}}^\infty (d-1) w^{d \operatorname{Re}\lambda_d/2 + (d-1)(\kappa' - \lambda_d)/2} e^{-x^2 w^{-(d-1)}} e^{-w} \frac{dw}{w} \right. \\
&\qquad\qquad \left. + \int_{x^{2/d}}^\infty v^{(d\kappa' + \lambda_d - \kappa')/2} e^{-v} e^{-\left(\frac{x^2}{v}\right)^{1/(d-1)}} \frac{dv}{v} \right) \\
&\leq 2(d-1)! \left(\int_{x^{2/d}}^\infty (d-1) w^{d\kappa/2} e^{-x^2 w^{-(d-1)}} e^{-w} \frac{dw}{w} \right. \\
&\qquad\qquad \left. + \int_{x^{2/d}}^\infty v^{d\kappa/2} e^{-v} e^{-\left(\frac{x^2}{v}\right)^{1/(d-1)}} \frac{dv}{v} \right) \\
&= 2(d-1)! \int_{x^{2/d}}^\infty C_{d,x}(u) u^{d\kappa/2} e^{-u} \frac{du}{u} \\
&\leq 2(d-1)! \left(\max_{u \geq x^{2/d}} C_{d,x}(u) \right) \Gamma\left(\frac{d\kappa}{2}, x^{2/d}\right) \\
&< 2d! x^\kappa e^{-x^{2/d}}
\end{aligned}$$

where $C_{d,x}(u) = (d-1)e^{-x^2 u^{-(d-1)}} + e^{-\left(x^2/u\right)^{1/(d-1)}}$, which is clearly bounded above by d for positive x and u. □

A bound for for $|\frac{\partial^r}{\partial s^r} G_s(x)|$ for $x > 1$, can be derived directly in terms of the bound for

$\phi(t)$:

$$\begin{aligned}
\left|\frac{\partial^r}{\partial s^r}G_s(x)\right| &= \left|\frac{\partial^r}{\partial s^r}\int_x^\infty \phi(t)t^s\frac{dt}{t}\right| \\
&< \int_x^\infty |\phi(t)(\log t)^r t^s|\frac{dt}{t} \\
&< \int_x^\infty 2d!t^\kappa e^{-t^{2/d}}(\log t)^r t^{\operatorname{Re} s}\frac{dt}{t} \\
&= 2d!\int_x^\infty e^{-t^{2/d}}(\log t)^r t^{\operatorname{Re} s+\kappa}\frac{dt}{t} \\
&= 2d!\frac{d}{2}\int_{x^{2/d}}^\infty e^{-t}(\log t^{d/2})^r t^{\frac{d}{2}(s+\kappa)}\frac{dt}{t} \\
&= 2d!\frac{d^{r+1}}{2^{r+1}}\int_{x^{2/d}}^\infty e^{-t}(\log t)^r t^{\frac{d}{2}(\operatorname{Re} s+\kappa)}\frac{dt}{t}.
\end{aligned}$$

If we further assume that $0 < (\log x^{2/d})^r < x^{2/d}$, giving $(\log t)^r < t$, when computing the r-th derivative (which is reasonable, as we are interested in the behavior as $x \to \infty$) we get that $\left|\frac{\partial^r}{\partial s^r}G_s(x)\right| < 2^{-r}d!d^{r+1}\Gamma(d(\operatorname{Re} s+\kappa)/2+1, x^{2/d})$.

4.3 Truncating the main series

We may now use the bound computed in the previous section to truncate the main series.

Recall that one of the conditions on our L-series was that the a_n grow at most polynomially in n. Let C and D be such that $|a_n| \leq Cn^D$ for all n. For example in the elliptic curve case we have $C = D = 1$. As $G_s(x)$ and its derivatives decay exponentially fast, it is easy to see that the series in Theorem 3.1.1 converges, but an explicit bound is needed for computational purposes. For simplicity of notation, let $\alpha = d(\operatorname{Re} s+\kappa)/2+1$, $\alpha' = \alpha + d(D+1)/2$, and $C' = 2^{-r}d!d^{r+1}C$. Choose N such that $(\log(N/A)^{2/d})^r < (N/A)^{2/d}$, and note that $G_s(x)$ and its derivatives are strictly decreasing as x increases. Then an explicit bound on the tail is given by

$$\begin{aligned}\left|\sum_{n=N+1}^{\infty} a_n \frac{\partial^r}{\partial s^r} G_s\left(\frac{n}{A}\right)\right| &< \int_N^{\infty} C u^D G_s\left(\frac{u}{A}\right) du \\ &< C' \int_N^{\infty} u^D \Gamma\left(\alpha, \left(\frac{u}{A}\right)^{2/d}\right) du \\ &= C' \int_N^{\infty} u^D \int_{(u/A)^{2/d}}^{\infty} e^{-t} t^{\alpha} \frac{dt}{t}\, du \\ &= C' \int_{(N/A)^{2/d}}^{\infty} \int_N^{At^{d/2}} u^D e^{-t} t^{\alpha} du\, \frac{dt}{t} \\ &= \frac{C'}{D+1} \int_{(N/A)^{2/d}}^{\infty} (At^{d/2})^{D+1} - N^{D+1}) e^{-t} t^{\alpha} du \frac{dt}{t} \\ &= \frac{C'}{D+1} \left(A^{D+1} \Gamma\left(\alpha', (N/A)^{2/d}\right) - N^{D+1} \Gamma\left(\alpha, (N/A)^{2/d}\right)\right)\end{aligned}$$

which allows us to know what precision is achieved when truncating the series after a given number of terms.

4.4 Truncating the series for $\phi(t)$ *and* $G_s(x)$

Rather than using an a priori bound for the number of terms needed to compute $G_s(x)$, it is more profitable to proactively sum the series until the terms get small enough to bound the tail. This is possible because, although the intermediate terms get quite large, after a certain point they converge quickly to zero. Recall from section 3.2 that the series expansion for $\phi(t)$ is

$$\phi(t) = \sum_{j=1}^{N} t^{-m_j} \sum_{k=1}^{|H_j|} \frac{(-\log t)^{k-1}}{(k-1)!} \sum_{n=1}^{\infty} c_{j,k}^{(n)} t^{2n}.$$

where the recursive formula for the $c_j^{(n)}$ is

$$c_j^{(n)}(z) = \frac{c_j^{(n-1)}(z)}{\prod_{a=1}^{d} \left(\frac{z+\lambda_a+m_j}{2} - n\right)}.$$

Let $N_j = \frac{1}{2}\max_a |\lambda_a + m_j|$. For $n > N_j$, the denominator is invertible as a power series yielding

$$(-2)^d \prod_{a=1}^{d} \sum_{b=0}^{\infty} (2n - \lambda_a - m_j)^{-1-b} z^b.$$

and it follows from the definition if m_j that the leading terms of the laurent series $c_j^{(n)}$ has exponent exactly $-|H_j|$. Expanding the quotient one finds

$$\begin{aligned}
\left|c_{j,k}^{(n)}\right| &= \left|(-2)^d \sum_{b_0=k}^{|H_j|} c_{j,b_0}^{(n-1)} \sum_{b_1+\cdots+b_d=b_0-k} \prod_{a=1}^{d} (2n-\lambda_a-m_j)^{-1-b_a}\right| \\
&\leq 2^d \sum_{b_0=k}^{|H_j|} \left|c_{j,b_0}^{(n-1)}\right| \sum_{b_1+\cdots+b_d=b_0-k} \left(\min_{1\leq a\leq d} |2n-\lambda_a-m_j|\right)^{-d(1+b_0-k)} \\
&\leq (n-N_j)^{-d} \left(\max_{k\leq b_0\leq |H_j|} \left|c_{j,b_0}^{(n-1)}\right|\right) \sum_{\substack{b_1+\cdots+b_d-b_0=-k \\ b_0\leq |H_j|}} 1.
\end{aligned}$$

That last sum is simply the number of ways to write $|H_j|-k$ as an ordered sum of $d+1$ non-negative integers, which is given by the binomial coefficient $\binom{d}{|H_j|-k}$. Because d is typically quite small, and to simplify analysis, we use the simpler upper bound 2^d giving

$$\frac{\max_{k\leq |H_j|} \left|c_{j,k}^{(n)}\right|}{\max_{k\leq |H_j|} \left|c_{j,k}^{(n-1)}\right|} \leq \left(\frac{2}{n-N_j}\right)^d$$

which is clearly decreasing with n. Choose an $N > N_j$ such that $\alpha = \left(\frac{2}{N-N_j}\right)^d t^2 < 1$. For each $k < |H_j|$,

$$\begin{aligned}
\left|\sum_{n=N}^{\infty} c_{j,k}^{(n)} t^{2n}\right| &\leq \sum_{n=N}^{\infty} \max_{k\leq |H_j|} \left|c_{j,k}^{(n)}\right| t^{2n} \\
&\leq \max_{k\leq |H_j|} \left|c_{j,k}^{(N)}\right| t^{2N} \sum_{n=0}^{\infty} \alpha^n \\
&= \max_{k\leq |H_j|} \left|c_{j,k}^{(N)}\right| t^{2N} \frac{1}{1-\alpha}.
\end{aligned}$$

Thus for sufficiently large N the error in the truncated series is bounded by a simple multiple of the first omitted term, allowing computation of the series to any desired precision. As $G_s(t)$ is computed via termwise integration of the series for $\phi(t)$, it can be truncated in a similar manner. Let $\tilde{G}_s(t) = t^{-s}\gamma(s) - G_s(t)$. Then

$$\begin{aligned}
\frac{\partial^r}{\partial s^r}\tilde{G}_s(x) &= \frac{\partial^r}{\partial s^r}\left(x^{-s}\int_0^x \phi(t)t^s\frac{dt}{t}\right) \\
&= \sum_{q=0}^{r}\binom{r}{q}(-\log x)^{r-q}x^{-s}\int_0^x \phi(t)(\log t)^q t^s\frac{dt}{t} \\
&= \sum_{q=0}^{r}\binom{r}{q}(-\log x)^{r-q}x^{-s}\int_0^x \left(\sum_{j=1}^{N} t^{-m_j}\sum_{k=1}^{|H_j|}\frac{(-\log t)^{k-1}}{(k-1)!}\sum_{n=1}^{\infty}c_{j,k}^{(n)}t^{2n}\right)(\log t)^q t^s\frac{dt}{t} \\
&= \sum_{q=0}^{r}\binom{r}{q}(-1)^q\sum_{j=1}^{N}x^{-m_j}\sum_{k=1}^{|H_j|}\frac{(k-1+q)!}{(k-1)!}\sum_{a=0}^{k-1+q}\frac{(-\log x)^{a+r-q}}{a!}\sum_{n=1}^{\infty}\frac{c_{j,k}^{(n)}}{(2n+s-m_j)^{k+q-a}}x^{2n}.
\end{aligned}$$

This algorithm has been implemented using Sage [38], and has been submitted for inclusion in a future release.

```
sage: zeta = LFunction(coefficients=[1]*8,
                       exponential_factor=1/sqrt(pi),
                       hodge_numbers=[0], poles=[(1, -1)])
sage: zeta(2, proof=True)
1.6449340668482264364724151666460251892189499012067984377355582?

sage: E = EllipticCurve('37a')
sage: L = LFunction(E)
sage: L.taylor_series(1.0, 2, proof=True)
0.?e-22 + 0.305999773834052372?*T + 0.186547797268162016?*T^2 + O(T^3)
sage: 2*E.period_lattice().real_period()*E.gen(0).height() / L(1, r=1, proof=True)
1.000000000000000?
```

Here, the question mark indicates an interval which is approximately ± the last digit.

4.5 Complexity

Combining the bound in section 4.3 with proposition 4.2.1, we get

$$\left| \sum_{n=N+1}^{\infty} a_n \frac{\partial^r}{\partial s^r} G_s \left(\frac{n}{A} \right) \right| = O\left(e^{-(N/A)^{2/d}+\varepsilon} \right)$$

for any $\varepsilon > 0$. This lets us estimate the number of terms in the main series needed for a given precision: $O(A)$ Dirichlet coefficients and $O(A)$ evaluations of $G_s(x)$ are needed to compute $L(s)$ to a fixed precision. Note further that applying proposition 4.2.1 requires $N/A > \alpha' + 1$, which agrees with the heuristic observation that $O(A)$ terms are required to get any precision at all. For example, when L is the Hasse-Weil L-function of an elliptic curve E then $d = 2$ and $A = \sqrt{N_E}/2\pi$, where N_E is the conductor of E. This agrees with the rule of thumb that the number of Dirichlet coefficients needed to evaluate L and its derivatives to a fixed precision is roughly proportional to the square root of the conductor.

Now consider varying the desired precision. To get P digits of precision requires $O(AP^{d/2})$ evaluations of $G_s(x)$. For $d = 1$ or $d = 2, \lambda_1 + 1 = \lambda_2$ we have seen earlier that we can evaluate $G_s(x)$ with a single incomplete gamma function (plus a constant number of P-digit arithmetic operations), and the incomplete gamma function can be evaluated uniformly to P digits of precision using $O(P)$ P-digit arithmetic operations [41]. Assuming asymptotically fast arithmetic [35] is used, this is $\tilde{O}(P^2)$ bitwise operations per G-function evaluation, giving a total complexity of $\tilde{O}(AP^{2+d/2})$. Though we have not worked it out in full generality, the analysis of the G-functions that arise all seem to follow the same pattern as that of the incomplete gamma function, giving the same complexity result (though with possibly larger constants).

Chapter 5

HEEGNER POINTS AND GROSS-ZAGIER FORMULAE

Let $E/\mathbb{Q}$ be an optimal elliptic curve of conductor N. Choose a square free $D < -4$ such that all primes dividing N split in $K = \mathbb{Q}(\sqrt{D})$. The ideal $N\mathcal{O}_K$ can be factored as $\mathcal{N}\bar{\mathcal{N}}$ such that $\mathcal{O}_K/\mathcal{N} \cong \mathbb{Z}/N\mathbb{Z}$. Viewing $\mathcal{O}_K$ as a lattice in $\mathbb{C}$, $\mathbb{C}/\mathcal{O}_K$ is an elliptic curve, and $\mathcal{N}^{-1}/\mathcal{O}_K$ is a cyclic subgroup of order N on this curve. Let x_1 be the point $(\mathbb{C}/\mathcal{O}_K, \mathcal{N}^{-1}/\mathcal{O}_K)$ on $X_0(N)$. By complex multiplication, x_1 is defined over K_1, the Hilbert Class Field of K. Using the modular parameterization $\varphi_E : X_0(N) \to E$, one obtains the point $y_1 = \varphi_E(x_1) \in E(K_1)$. Let y_K be the trace of y_1 down to K. The point y_K is called the Heegner point with discriminant D and is well defined up to sign. These Heegner points have many uses, and play an essential role in the following theorem:

Theorem 5.0.1. *(Gross-Zagier, Kolyvagin) Let $E/\mathbb{Q}$ be an elliptic curve with analytic rank $r_{an}(E/\mathbb{Q}) \leq 1$. Then the Shafarevich-Tate group $\text{Ш}(E/\mathbb{Q})$ is finite and $r_{an}(E/\mathbb{Q}) = r_{alg}(E/\mathbb{Q})$.*

In other words, the rank part of the BSD conjecture is true for curves of analytic rank ≤ 1.

Explicit computation of Heegner points is a well studied topic [12], [40], [9]. The most computationally feasible algorithms boil down to numerically approximating a representitive of $x_1 \in X_0(N)(\mathbb{C})$ as an element in the upper half plane and numerically applying the modular parameterization and Weierstrass $\wp$-function to get an approximation to the x-coordinate of y_1 or y_K on the Weierstrass model for E. These calculations are done with enough accuracy to recognize this x-coordinate as an element of K_1. To determine the precision needed to recognize $y_K \in E(K)$, a height bound on y_K is needed. This is given by

Theorem 5.0.2. *(Gross-Zagier [19, §5.2]) Let $E/\mathbb{Q}$ be an elliptic curve and K an imaginary quadratic field satisfying the Heegner hypothesis. Then the Néron-Tate canonical height of* *the Heegner point $\hat{h}(y_K)$ is given by*

$$\hat{h}(y_K) = \frac{\sqrt{|D|}}{4}\frac{L'(E/K,1)}{\Omega_E}.$$

As the points in $E(K)$ are not a discrete subset of $E(\mathbb{C})$ for positive-rank curves, a provable height bound on the Heegner point is needed to make the numerical methods of computing Heegner points rigorous.

5.1 Higher Heegner Points

Kolyvagin's cohomology classes are constructed from generalized Heegner points with a conductor c. Fix an integer c whose prime divisors are inert in K and coprime to N. Let $\mathcal{O}_c = \mathbb{Z} + c\mathcal{O}_K$ and $\mathcal{N}_c = \mathcal{N} \cap \mathcal{O}_c$. Again, $\mathbb{C}/\mathcal{O}_c$ defines an elliptic curve with a cyclic subgroup $\mathcal{N}_c^{-1}/\mathcal{O}_c$ of order N. This gives a point $x_c \in X_0(N)$ which is defined over K_c, the ring class field of conductor c of K. As before, we use the modular parameterization to map x_c to a point y_c on $E(K_c)$. Jetchev, Lauter, and Stein give an method to explicitly compute y_c in a manner analogous to the computation of y_K [23]. These higher Heegner points are used in Kolyvagin's Euler systems to construct elements in the Selmer group. It is hoped that understanding Kolyvagin classes may shed light on both the BSD conjecture for higher rank curves and the behavior of $\text{Ш}(E/K)$ for curves of all rank [37].

As before, a key component in making this rigorous is a provable height bound for y_c, using a generalization of the Gross-Zagier formula [19]. Let χ be a character $\chi : \mathrm{Gal}(K_c/K) \to \mathbb{C}^\times$ and e_χ the idempotent

$$e_\chi = \frac{1}{\#\,\mathrm{Gal}(K_c/K)} \sum_{\sigma\in\mathrm{Gal}(K_c/K)} \chi^{-1}(\sigma)\sigma \in \mathbb{C}[\mathrm{Gal}(K_c/K)].$$

The heights $\hat{h}(e_\chi y_c)$ are related to the special values of certain L-functions, as described below. The e_χ subspaces of $E(K_c)\otimes\mathbb{C}$ are orthogonal with respect to the height pairing, as, for $\chi \neq \chi'$,

$$\begin{aligned}\langle e_\chi P, e_{\chi'} Q\rangle &= \sum_\sigma \sum_{\sigma'} \chi^{-1}(\sigma)\overline{\chi'^{-1}(\sigma')}\left\langle P^\sigma, Q^{\sigma'}\right\rangle = \sum_\sigma \sum_{\sigma'} \chi^{-1}(\sigma)\chi'(\sigma')\left\langle P^\sigma, Q^{\sigma'}\right\rangle \\ &= \sum_\sigma \sum_{\sigma'} \chi^{-1}(\sigma)\chi'(\sigma\sigma')\left\langle P^\sigma, Q^{\sigma\sigma'}\right\rangle = \left\langle \sum_\sigma \chi^{-1}(\sigma)\chi'(\sigma)P^\sigma \ , \ e_{\bar{\chi}'} Q^\sigma \right\rangle \\ &= \langle 0 \ , \ e_{\bar{\chi}'} Q^\sigma \rangle = 0.\end{aligned}$$

Thus $\hat{h}(y_c)$ is simply given by

$$\hat{h}(y_c) = \hat{h}\left(\sum_\chi e_\chi y_c\right) = \sum_\chi \hat{h}(e_\chi y_c).$$

Let f be the newform corresponding to E, and let $L(f, \chi, s)$ be the Rankin-Selberg convolution $L(f \otimes g_\chi, s)$ as described in [18, section III]. In [42], Zhang proves some generalization of the Gross-Zagier formula that relates the heights of the higher Heegner points to the special value $L'(f, \chi, 1)$. For non-trival χ, [23] claims that this specializes to

$$L'(f, \chi, 1) = \frac{4}{\sqrt{|D|}}(f, f)\hat{h}(e_\chi y_c)$$

The earlier paper [21] conjectures that the formula should be

$$L'(f, \chi, 1) = \frac{h_c}{\sqrt{|D|}}||\omega_f||^2 \hat{h}(e_\chi y_c)$$

where $h_c = [K_c : K]$. Based on numerical evidence and consistency checks with the BSD conjecture, neither of these appear to be correct. Instead, we have the following conjecture:

Conjecture 5.1.1. *For non-trivial χ, he formula relating the special value of $L(f, \chi, s)$ to the heights of the Heegner points is*

$$L'(f, \chi, 1) = \frac{h_c}{\mathrm{cond}(\chi)\sqrt{|D|}}||\omega_f||^2 \hat{h}(e_\chi y_c)$$

where $\mathrm{cond}(\chi)$ *is the conductor of* χ.

5.2 Recognizing Heegner points

Though the points in $E(K)$ are arbitrarily close in $E(\mathbb{C})$, the fact that there are a finite number of points of bounded height allows us to recognize a point in $E(K)$ of known height exactly from a sufficiently high precision numerical approximation. Though an exhaustive point search gives an algorithm, it is clearly impractical for anything but the smallest heights and extensions of $\mathbb{Q}$. To come up with a feasible algorithm, we first recall the methods used to recognize algebraic numbers. For classical Heegner points y_K, the coordinates are known to be rational, so may be recognized via continued fractions. The generalized Heegner points y_c are not in general defined over $\mathbb{Q}$ (indeed, for curves of higher rank they are never defined over $\mathbb{Q}$ unless they are torsion) so we need more sophisticated algorithms to recognize the coordinates of y_c as elements of K_c.

5.2.1 Recognizing algebraic numbers

Recognizing algebraic numbers from numerical approximations is typically done using integer relation algorithms. Given a numerical approximation x to an algebraic number α of degree n, one attempts to find integers $a_0, ..., a_n$, not all 0, such that $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$. There are several algorithms such as PSLQ that given a set of real (or complex) numbers $x_1, ..., x_n$ and upper bound on $|a_i|$ will either produce $(a_1, ..., a_n)$ such that $a_1x_1 + \cdots + a_nx_n = 0$ (to the precision of the inputs) or certify that no such integers exist. To provably recognize an algebraic number of a given height, we need a slightly stronger statement, namely that the $(a_1, ..., a_n)$ is the unique (up to scalar multiples) vector satisfying this relation and height bound. The LLL algorithm can be adapted to give us this stronger statement, and works as follows.

Let L be the lattice spanned by $\{\langle 1, 0, ..., 0, x_1\rangle, \langle 0, 1, ..., 0, Cx_2\rangle, ..., \langle 0, ..., 0, 1, Cx_n\rangle\}$ for some large constant C. The short vectors in this lattice correspond to vectors of the form $\langle a_1, ..., a_n, C\cdot(a_1x_1 + \cdots a_nx_n)\rangle$ where the a_i are minimized and $a_1x_1 + \cdots a_nx_n$ is very small. Although finding shortest vectors is NP-hard (at least in the generic case), an LLL-reduced basis can be computed in polynomial time and has strong enough properties to give us what we need [29]. In particular, we have the following theorem:

Theorem 5.2.1 (Lenstra-Lenstra-Lovász)**.** *Let $b_1, ..., b_n$ be an LLL-reduced basis for a lattice L, with parameter $\delta = \frac{3}{4}$. Then*

1. *$|b_1| \leq 2^{\frac{n-1}{2}}|v|$ for every non-zero $v \in L$ and, more generally,*

2. *$|b_t| \leq 2^{\frac{n-1}{2}} \max\{|v_1|, ..., |v_t|\}$ for every set of t linearly independent vectors $v_1, ..., v_t \in L$.*

We now have the immediate corollary which applies when any of the basis vectors are sufficiently different in size.

Corollary 5.2.2. *Let $b_1, ..., b_n$ be an LLL-reduced basis for a lattice L, with parameter $\delta = \frac{3}{4}$, and suppose that $|b_i| < 2^{-\frac{n-1}{2}}|b_t|$ for each $i < t$. Then any $w \in L$ linearly independent from $b_1, ..., b_{t-1}$ is bound from below by $2^{-\frac{n-1}{2}}|b_t| \leq |w|$.*

Proof. Suppose w is linearly independent from $b_1, ..., b_t$ and $|w| < 2^{-\frac{n-1}{2}}|b_{t+1}|$. Then by the above theorem we have $|b_t| \leq 2^{\frac{n-1}{2}} \max\{|b_1|, ..., |b_{t-1}|, |w|\} < 2^{\frac{n-1}{2}} \cdot 2^{-\frac{n-1}{2}}|b_t|$ which is a contradiction. □

In particular, this tells us that when $|b_1| < 2^{\frac{n-1}{2}}|b_2|$ then all points of L of norm less than $2^{-\frac{n-1}{2}}$ are in fact scalar multiples of b_1 and b_1 is the shortest vector in L.

Note that due to the finite nature of computers, the algorithm cannot be run with the actual x_i, but only approximations thereof. In particular, running LLL on the above lattice will gives an approximate solution which could be falsified but, for general x_i, cannot be *numerically* verified to hold exactly. However, if we know a unique solution with height at most B exists, we can provably recognize it.

Theorem 5.2.3. *If there is a unique (up to scalars) integer solution $(s_1, ..., s_n)$ of $a_1x_1 + \cdots + a_nx_n = 0$ of bounded height B, then for sufficiently large C the algorithm above gives proof of correctness of the returned solution.*

Proof. Let $X = \{|a_1x_1 + \cdots + a_nx_n| : a_i \in \mathbb{Z}, |a_i| \leq 2^{\frac{n-1}{2}}\sqrt{n}B\}$. Because X is finite, it has a minimum positive element ϵ. Let $C > 2^{\frac{n-1}{2}}\sqrt{n}B\epsilon^{-1}$. Let $v = \langle v_1, ..., v_{n+1}\rangle \in \mathbb{Z}^{n+1}$ be linearly independent from s. If $v \in X$ then $|v| \geq |v_{n+1}| \geq C\epsilon > 2^{\frac{n-1}{2}}\sqrt{n}B$. Alternatively, if

$v \notin X$ then we also have $|v| \geq \max_{1 \leq i \leq n} |v_i| > 2^{\frac{n-1}{2}} \sqrt{n} B$. Thus any non-solution vector v has norm at least $2^{\frac{n-1}{2}} \sqrt{n} B$.

Let $b_1, ..., b_n$ be an LLL-reduced bases of L. As $s = \langle s_1, ..., s_n, 0 \rangle \in L$, and $|s| < \sqrt{n} B$, we see that $|b_1| < 2^{\frac{n-1}{2}} |s| = 2^{\frac{n-1}{2}} \sqrt{n} B$ and hence b_1 must represent a solution. This solution is unique up to scalars, so b_1 spans the set of solutions, and the linearly independent b_2 is not a solution. Thus $|b_2| > 2^{\frac{n-1}{2}} \sqrt{n} B > 2^{\frac{n-1}{2}} |b_1|$, verifying that the span of b_1 contains all elements of L of bounded height B. □

In practice, much smaller C often gives large enough $|b_2|$ to verify the solution.

The bounds between and canonical and naïve heights of points on elliptic curves [11] now allow us to put a bound on the height of the minimal polynomial of the x-coordinate of y_c, allowing us to recognize Heegner points exactly.

5.3 Heegner index

Let E be an elliptic curve over $\mathbb{Q}$ of analytic rank ≤ 1. In this case, the subgroup generated by the Heegner point plays an essential role in the proof of the BSD conjecture. In particular, the nontorsion point $y_K = \mathrm{Tr}_{K_1/K}(y_1)$ is used to bound the rank of $E(K)$ from below. As a consequence of the Gross-Zagier formula and conjectural BSD formula, one can show that

$$[E(K) : \mathbb{Z} y_K] = c_E \sqrt{\# \text{Ш}(E/K)} \prod c_p$$

where c_E here is the Manin constant of E. (For simplicity of presentation, from here on we will assume that $c_E = 1$, which is conjectured to be the case for optimal curves). Several results have been proved bounding the p-part of Sha in terms of p-divisibility of the Heegner index and the Tamagawa numbers [24, 7, 17, 22], which play an essential role in verifying the full BSD conjecture for specific elliptic curves.

We now investigate what the implications are if we assume the BSD formula in conjunction with the generalizations of the Gross-Zagier formula (Conjecture 5.1.1) for higher rank curves.

Let E be an elliptic curve over $\mathbb{Q}$ of rank greater than 1, and choose a Heegner discriminant D with $K = \mathbb{Q}(\sqrt{D})$. Let $y_c \in E(K_c)$ be the Heegner point of square-free conductor

c, $G = \mathrm{Gal}(K_c/K)$, and $h_c = [K_c : K]$. Bertolini and Darmon [3] showed that the Galois orbit of y_c spans $E(K_c)/E(K)$ up to finite index whenever $y_c \neq 0$, under the assumption that E does not have complex multiplication. Denote the span of this orbit by W.

Proposition 5.3.1. *If the rank of $E(K)$ is greater than one then W is orthogonal to $E(K)$ with respect to the canonical height pairing.*

Proof. Choose $P \in E(K)$ and $Q \in W$.The Gross-Zagier formula tells us that $\hat{h}(\mathrm{Tr}_{K_c/K}\, y_c) = \hat{h}(\mathrm{Tr}_{K_1/K}\, a_c y_1) = a_c^2 \hat{h}(y_K) = 0$, and hence $\mathrm{Tr}_{K_c/K}\, Q$ is torsion for all $Q \in W$. Due to Galois equivariance of the trace paring we have $\langle P, Q\rangle = \langle \sigma P, \sigma Q\rangle = \langle P, \sigma Q\rangle$ for all $\sigma \in G$. Thus

$$\langle P, Q\rangle = \frac{1}{h_c}\sum_{\sigma\in G}\langle P, \sigma Q\rangle = \frac{1}{h_c}\langle P, \mathrm{Tr}_{K_c/K}\, Q\rangle = 0.$$

□

Note that we do not necessarily have orthogonality when the rank of $E(K)$ is 1. For example, for the elliptic curve defined by $y^2 + y = x^3 - x$, $K = \mathbb{Q}(\sqrt{-7})$, and $P = (0,0) \in E(K)$ we have $\langle P, y_5\rangle = 0.017037...$

We now derive a conjectural formula for the index $[E(K_c) : E(K) \oplus W]$ when E has rank at least 2 and c is a product of distinct primes.

First, recall the the BSD conjectural formula for a curve E over a number field F of rank r [28]:

$$\frac{L^{(r)}(E/F)}{r!} = \frac{\Omega_{E/F}\,\mathrm{Reg}_{E(F)}\,\#\text{Ш}(E/F)\prod_v c_v}{\sqrt{|\Delta_F|}\#E(F)_{tor}^2}.$$

If E is defined over $\mathbb{Q}$ and K is totally imaginary, as it is in our case, we have $\Omega_{E/F} = ||\omega_f||^{[F:Q]}$.

The L-function of $E(K_c)$ breaks down as

$$L(E/K_c, s) = \prod_\chi L(f,\chi,s) = L(E/K,s)\prod_{\chi\neq\chi_0} L(f,\chi,s)$$

where the product is over characters $\chi : G \to \mathbb{C}^*$ with χ_0 being the trivial character. The sign in the functional equation for $L(f,\chi,s)$ is -1, showing it vanishes to odd order at

$s = 1$. By [3] we have that $e_\chi y_c$ is non-torsion for $\chi \neq \chi_0$ whenever y_c is non-torsion, so the corresponding $L(f, \chi, s)$ vanish to order exactly 1 at $s = 1$. Thus we get equality for the leading coefficients of Taylor series expansions about $s = 1$:

$$\frac{L^{(r+h_c-1)}(E/K_c, 1)}{(r + h_c - 1)!} = \frac{L^{(r)}(E/K, 1)}{r!} \prod_{\chi \neq \chi_0} L'(E/K, \chi, 1)$$

Applying the BSD formula and the generalized Gross-Zagier formula on both sides yields

$$\frac{||\omega_f||^{2h_c} \operatorname{Reg}_{E(K_c)} \#\text{Ш}(E/K_c) \prod c_{v,K_c}}{\sqrt{|\Delta_{K_c}|} \#E(K_c)^2_{tor}} = \frac{||\omega_f||^2 \operatorname{Reg}_{E(K)} \#\text{Ш}(E/K) \prod c_{v,K}}{\sqrt{|D|} \#E(K)^2_{tor}} \prod_{\chi \neq \chi_0} \frac{h_c ||\omega_f||^2}{\operatorname{cond}(\chi) \sqrt{|D|}} \hat{h}(e_\chi y_c).$$

Rearranging the terms on the left and the right gives

$$\frac{\sqrt{|D|^{h_c}} \prod_{\chi \neq \chi_0} \operatorname{cond}(\chi)}{\sqrt{|\Delta_{K_c}|}} \frac{\prod c_{v,K_c}}{\prod c_{v,K}} \frac{\#\text{Ш}(E/K_c)}{\#\text{Ш}(E/K)} = \frac{\operatorname{Reg}_{E(K)} h_c^{h_c-1} \prod_{\chi \neq \chi_0} \hat{h}(e_\chi y_c)}{\operatorname{Reg}_{E(K_c)}} \cdot \frac{\#E(K_c)^2_{tor}}{\#E(K)^2_{tor}}. \tag{5.1}$$

Because we understand the ramification behavior of $K_c/\mathbb{Q}$ completely [20], we can compute the discriminant Δ_{K_c} explicitly. For each prime $p|c$, write $c = pc'$. The unique prime $(p) \subset K$ above p splits completely in $K_{c'}/K$ as it is inert in K. Going from $K_{c'}$ to K_c the primes above p are totally ramified (with ramification index $[K_c : K_{c'}] = p + 1$). Thus the different $\delta_{K_c/K}$ is $\prod_{p|c} \prod_{\mathfrak{p}|p} \mathfrak{p}^p$. The different ideal is multiplicative over towers, and the discriminant is the norm of the different, thus we have

$$\Delta_K = \operatorname{Norm}_{K_c/\mathbb{Q}}(\delta_{K/\mathbb{Q}} \delta_{K_c/K}) = \operatorname{Norm}_{K_c/\mathbb{Q}}(\delta_{K/\mathbb{Q}}) \prod_{p|c} \prod_{\mathfrak{p}|p} \operatorname{Norm}_{K_c/\mathbb{Q}}(\mathfrak{p})^p = D^{h_c} \prod_{p|c} p^{\frac{2h_c p}{p+1}}.$$

Now consider the set of ring class characters $\chi : G \to \mathbb{C}^*$. For every $p|c$ we see that $p|| \operatorname{cond}(\chi)$ for every χ except for χ in the coset that acts trivially on the cyclic factor of size $p+1$ associated to p. In other words, $p|| \operatorname{cond}(\chi)$ for $h_c - h_c/(p+1)$ of the χ. As $\operatorname{cond}(\chi)|c$

we have $\prod_{\chi\neq\chi_0}\operatorname{cond}(\chi) = \prod_{p|c} p^{h_c - h_c/(p+1)} = \prod_{p|c} p^{h_c p/(p+1)}$, showing that

$$\frac{\sqrt{|D|^{h_c}}\prod_{\chi\neq\chi_0}\operatorname{cond}(\chi)}{\sqrt{|\Delta_{K_c}|}} = 1. \tag{5.2}$$

To relate the product of the heights of the $e_\chi y_c$ to the regulator of W, let $(e_\chi y_c)_{\chi\neq\chi_0}$ be the basis for a lattice L in $W_{/tor}\otimes\mathbb{C}$ (extending the height paring linearly via $\langle\alpha P, \beta Q\rangle = \alpha\bar{\beta}\langle P, Q\rangle$). Because the $e_\chi y_c$ are orthogonal, we can easily compute the regulator $\operatorname{Reg}_L = \prod_{\chi\neq\chi_0}\hat{h}(e_\chi y_c)$. Let $(y_c^\sigma)_{1\neq\sigma\in G}$ be a basis for $W_{/tor}$, noting that because $\operatorname{Tr} y_c$ is torsion, $y_c = -\sum_{\sigma\neq 1} y_c^\sigma$ in $W_{/tor}$, and let M be the change of basis matrix from $W_{/tor}$ to L with respect to these bases. This gives us $\operatorname{Reg}_L = |\det M|^2 \operatorname{Reg}_W$, so we simply need to compute the determinant of M. By definition of e_χ we have $e_\chi y_c = \frac{1}{|G|}\sum_{\sigma\in G}\chi^{-1}(\sigma) y_c^\sigma = \frac{1}{h_c}\sum_{1\neq\sigma\in G}(\chi^{-1}(\sigma)-1)y_c^\sigma$. For any two rows M_{χ_i}, M_{χ_j} of M,

$$\begin{aligned} M_{\chi_i}\cdot M_{\chi_j} &= \frac{1}{h_c^2}\sum_{1\neq\sigma\in G}(\chi_i^{-1}(\sigma)-1)(\chi_j^{-1}(\sigma)-1) = \frac{1}{h_c^2}\sum_{\sigma\in G}(\chi_i^{-1}(\sigma)-1)(\chi_j^{-1}(\sigma)-1) \\ &= \frac{1}{h_c^2}\sum_{\sigma\in G}(\chi_i\chi_j)^{-1}(\sigma) - \chi_i^{-1}(\sigma) - \chi_j^{-1}(\sigma) + 1 = \begin{cases} \frac{2}{h_c} & \text{if } \chi_i = \chi_j^{-1} \\ \frac{1}{h_c} & \text{otherwise.} \end{cases} \end{aligned}$$

Thus

$$(\det M)^2 = \det MM^T = \det(M_{\chi_i}\cdot M_{\chi_j})_{i,j} = \pm\begin{vmatrix} \frac{2}{h_c} & \frac{1}{h_c} & \cdots & \frac{1}{h_c} \\ \frac{1}{h_c} & \frac{2}{h_c} & & \vdots \\ \vdots & & \ddots & \frac{1}{h_c} \\ \frac{1}{h_c} & \cdots & \frac{1}{h_c} & \frac{2}{h_c} \end{vmatrix}$$

where the columns in the final matrix have been permuted to make the matrix diagonal which only affects the determinant up to sign. To evaluate this we can use the following lemma which we will prove at the end of this section:

Lemma 5.3.2. *Let $M_m(a,b)$ be the $m\times m$ matrix with $a+b$ along the diagonal and all other entries equal to b. Then* $\det M_m(a,b) = (a+mb)a^{m-1}$.

Now we have $\operatorname{Reg}_W = (\det M)^{-2}\operatorname{Reg}_L = h_c^{h_c}\prod_{\chi\neq\chi_0}\hat{h}(e_\chi y_c)$. As W and $E(K)$ are orthogonal, and $\operatorname{Reg}_{E(K)}$ was computed with respect to the heigh pairing over K rather than K_c we have $\operatorname{Reg}_{E(K)\oplus W} = h_c^r \operatorname{Reg}_{E(K)} \operatorname{Reg}_W$ giving

$$\begin{aligned} \frac{\operatorname{Reg}_{E(K)} h_c^{h_c-1} \prod_{\chi\neq\chi_0} \hat{h}(e_\chi y_c)}{\operatorname{Reg}_{E(K_c)}} &= \frac{\operatorname{Reg}_{E(K)} h_c^{-1} \operatorname{Reg}_W}{\operatorname{Reg}_{E(K_c)}} \\ &= \frac{h_c^{r-1} \operatorname{Reg}_{E(K)\oplus W}}{\operatorname{Reg}_{E(K_c)}} \end{aligned} \tag{5.3}$$

Combining (5.1), (5.2), and (5.3) gives

Conjecture 5.3.3. *For an elliptic curve $E/\mathbb{Q}$ of rank at least 2, Heegner field K, and square-free c such that all the primes dividing c are inert in K, we have*

$$[E(K_c) : E(K) \oplus W]^2 = h_c^{r-1} \cdot \frac{\prod c_{v,K_c}}{\prod c_{v,K}} \cdot \frac{\#Ш(E/K_c)}{\#Ш(E/K)}.$$

Proof of lemma 5.3.2. The case for $m = 1$ is clear. For $m > 1$, first consider the determinant of the matrix $M'_m(a,b)$ of size $m \times m$ whose entries are all b except for the first upper off diagonal whose entries are $a+b$. We claim that $\det M'_m(a,b) = (-a)^{m-1}b$. Again, for $m = 1$ this is clear. For larger m we perform a row operation and do expansion by minors.

$$\det M'_m(a,b) = \begin{vmatrix} b & a+b & \cdots & b \\ b & b & \ddots & \vdots \\ \vdots & & \ddots & a+b \\ b & \cdots & b & b \end{vmatrix} = \begin{vmatrix} 0 & a & \cdots & 0 \\ b & b & \ddots & \vdots \\ \vdots & & \ddots & a+b \\ b & \cdots & b & b \end{vmatrix} = -a \det M'_{m-1}(a,b) = (-a)^{m-1}b.$$

Using this we can compute

$$\det M_m(a,b) = \begin{vmatrix} a+b & b & \cdots & b \\ b & a+b & & \vdots \\ \vdots & & \ddots & b \\ b & \cdots & b & a+b \end{vmatrix} = \begin{vmatrix} a & 0 & \cdots & -a \\ b & a+b & & \vdots \\ \vdots & & \ddots & b \\ b & \cdots & b & a+b \end{vmatrix}$$

$$= a \det M_{m-1}(a,b) + (-1)^m(-a) \det M'_{m-1}(a,b) = (a+mb)a^{m-1}.$$

□

Chapter 6

OTHER APPLICATIONS

6.1 High precision BSD verification

Given that we don't know how to verify the BSD conjecture for curves of rank $r \geq 2$, can we at least verify it holds to a million digits? The orders of groups occurring in the formula are of course integers, and hence known to infinite precision, so we only need to concern ourselves with the real values $L^{(r)}(E,1)$, Ω_E, and Reg_E.

Computing values such as $L^{(r)}(E,1)$ is the main topic of this paper, so we will not discuss this here other than to say that in this specific case of an elliptic curve over $\mathbb{Q}$, the G-function $G_s(x)$ can be expressed in terms of incomplete gamma function simplifying much of the computation, especially at $s = 1$ (see, e.g. [10, Chapter 2]). Computing $L^{(r)}(E,1)$ to B bits of precision requires $\widetilde{O}(B\sqrt{N})$ evaluations of (derivatives of) the incomplete Gamma function, each of which can be done in $\widetilde{O}(B^2)$ bit operations [41], yeilding a total runtime of $\widetilde{O}(B^3\sqrt{N})$.

The constant Ω_E can be computed using the arithmetic-geometric mean. As mentioned in Section 2, it is either the real period ω_1 of the period lattice of E or its double. Given a Weierstrass model for E, we can write down the 2-division polynomial $\psi_2(x) \in \mathbb{Q}[x]$ whose roots are precisely the x-coordinates of the points of order 2 on E. Let e_i be these three roots. The period lattice of E in $\mathbb{C}$ is then spanned by

$$\begin{aligned}\omega_1 &= \frac{\pi}{AGM\left(\sqrt{e_3 - e_1}, \sqrt{e_3 - e_2}\right)}\\ \omega_2 &= \frac{\pi i}{AGM\left(\sqrt{e_3 - e_1}, \sqrt{e_2 - e_1}\right)}.\end{aligned}$$

Though the complex AGM is only defined up to a choice of square root, any consistent choice will give a valid basis. If our curve is defined over a totally real number field, we can avoid the complex AGM by ordering the roots $e_3 > e_2 > e_1$ (in the case that there are three real roots) or letting e_1 and e_2 be the complex conjugate pair (in the case there is exactly one real root). In this latter case, let $z = \sqrt{e_3 - e_1}$ and compute $\omega_1 = \pi/AGM\left(z, \bar{z}\right), = \pi/AGM\left(\frac{1}{2}(z+\bar{z}), z\bar{z}\right)$, which can be done over the real numbers. Both Newtons' method for finding the roots e_i and the AGM converge quadratically, so computing ω_1 (or $2\omega_1$) to any desired precision can be done very quickly.

This leaves the regulator Reg_E, which is the determinant of the height pairing matrix $|\langle P_i, P_j\rangle|_{i,j}$, where $P_1, ..., P_r$ is a set of generators for the non-torsion part of E. The height pairing is a bilinear form defined by

$$\langle P, Q\rangle = \frac{1}{2}\left(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)\right)$$

where $\hat{h}$ is the canonical height for points on E. It should be noted that there are two common normalizations for this height, one of which is twice the other. The larger is the one suitable for the BSD conjecture (the smaller giving a regulator that is too small by 2^r). Given a point $P = (x, y) \in E(\mathbb{Q})$, the naïve height $h(P)$ is defined to be $2\log\max\{|a|, |c|\}$ where $x = a/c$ with a, c relatively prime. The canonical height is then defined as the limit $\hat{h}(P) = \lim_{n\to\infty} 4^{-n} h(2^n P)$. This definition, however, is completely unsuitable for computational purposes. The canonical height decomposes as a sum of local heights

$$\hat{h}(P) = \sum_{v\leq\infty} \lambda_v(P) = h_\infty(P) + \sum_{v|c\Delta} \lambda_v(P)$$

.

The non-archimedian local heights are simple rational multiples of $\log p$ for $p|c\Delta$, that are easily computed to whatever precision is desired (assuming a factorization of $c\Delta$). See, for example, [10, Chapter 4] or [36] for details. The archimedian height λ_∞ is usually computed using a series originating with Tate, whose convergence properties were improved upon by Silverman [10]. This works fine for low precisions, but the number of terms required is linear in the desired precision.

A better series based on theta series is given in [8, algorithm 7.5.7], where the number of terms required for b bits of precision is only $O(\sqrt{b})$, but this is still nowhere near as fast as computing Ω_E for large b. Here we present an algorithm based on unpublished work of Mestre and Bost [4] that uses the AGM to give b bits of precision in only $O(\log n)$ steps which is suitable for very high precision computations. This approach has been known for some time, our contribution is to do the analysis required to obtain effective bounds on the series.

For simplicity of notation, let $\mu_E(P) = \lambda_{\infty,E}(P) + \frac{1}{12}\Delta_E$. For the moment, assume our curve contains full 2-torsion over the $\mathbb{R}$. By a suitable transformation of coordinates, we can write E in the form

$$E : y^2 = x(x + a^2)(x + b^2)$$

where $0 < b < a$. Let (a_n, b_n) be the series used to obtain the arithmetic-geometric mean of a and b defined by $a_0 = a, b_0 = b, a_{n+1} = \frac{a_n+b_n}{2}, b_{n+1} = \sqrt{a_n b_n}$. This yields a series of curves

$$E_n : y^2 = x(x + a_n^2)(x + b_n^2)$$

defined over $\mathbb{R}$ with degree 2 isogonies $\phi_n : E_{n+1} \to E_n$ given by

$$(x_{n+1}, y_{n+1}) \mapsto \left(\frac{x_{n+1}(x_{n+1} + b_{n+1}^2)}{x_{n+1} + a_{n+1}^2}, \frac{(x_n + a_n a_{n+1})(x_n + b_n a_{n+1})}{(x + a_{n+1})^2} y_{n+1} \right).$$

A point P_n on the identity component of E_n lifts to a unique point P_{n+1} in the identity component of E_{n+1}. This gives a sequence of points $P_n = (x_n, y_n)$ whose limit is a point P_∞ on the degenerate curve $E_\infty : y^2 = x\left(x + c^2\right)^2$, where $c = \mathrm{AGM}(a, b)$. We have $\mu_{E_\infty}(P_\infty) = \frac{1}{2}\log|x(P_\infty)|$. The recurrence for the x-coordinate of P_n is

$$x_{n+1} = \frac{1}{2}\left(x_n - a_n b_n + \sqrt{(x_n + a_n^2)(x_n + b_n^2)}\right).$$

By the proposition on page 4 of [1], we can relate heights of points on E_{n+1} to their images on E_n as follows:

$$\mu_{E_n}(P_n) = 2\mu_{E_{n+1}}(P_{n+1}) - \frac{1}{2}\log|x_{n+1} + a_{n+1}^2|.$$

Considering the infinite series of isogonies $E_0 \leftarrow E_1 \leftarrow \cdots \leftarrow E_n \leftarrow \cdots$ and a point P_0 in the identity component of E_0 we obtain

$$\mu_{E_0}(P_0) = \frac{1}{2}\lim_{n\to\infty} 2^{n-1}\log|x_n + a_n^2| - \sum_{k=1}^{n-1} 2^{k-1}\log|x_k + a_k^2| = \frac{1}{2}\log\lim_{n\to\infty}\frac{|x_n + a_n^2|^{2^{n-1}}}{\prod_{k=1}^{n-1}|x_k + a_k^2|^{2^{k-1}}}.$$

In order to turn this formula into an algorithm, we need to determine where to truncate the series and to bound the size of the tail.

$$\begin{aligned}
\mu_{E_0}(P_0) &= \frac{1}{2}\lim_{n\to\infty} 2^{n-1}\log|x_n + a_n^2| - \sum_{k=1}^{n-1} 2^{k-1}\log|x_k + a_k^2| \\
&= \frac{1}{2}\log|x_1 + a_1^2| + \frac{1}{2}\lim_{n\to\infty}\sum_{k=1}^{n-1} 2^k\left(\log|x_{k+1} + a_{k+1}^2| - \log|x_k + a_k^2|\right) \\
&= \frac{1}{2}\log|x_1 + a_1^2| + \frac{1}{2}\sum_{k=1}^{\infty} 2^k\log\frac{|x_{k+1} + a_{k+1}^2|}{|x_k + a_k^2|}
\end{aligned}$$

Let $0 < \varepsilon_n = a_n - b_n$ and note that $\varepsilon_n < \frac{1}{2^{2^n-1}}(a_0 - b_0)$ as the a_n, b_n are an AGM sequence. Now

$$\begin{aligned}
\left|(x_{n+1} + a_{n+1}^2) - (x_n + a_n^2)\right| &= \left|\frac{1}{2}\left(x_n - a_n b_n + \sqrt{(x_n + a_n^2)(x_n + b_n^2)}\right) + \left(\frac{a_n + b_n}{2}\right)^2 - x_n - a_n^2\right| \\
&= \frac{1}{4}\left|2\sqrt{(x_n + a_n^2)(x_n + b_n^2)} + b_n^2 - 2x_n - 3a_n^2\right| \\
&= \frac{1}{4}\left|2\sqrt{(x_n + a_n^2)(x_n + b_n^2)} - 2\sqrt{(x_n + a_n^2)^2} + b_n^2 - a_n^2\right| \\
&\le \frac{1}{4}\left|2\sqrt{(x_n + a_n^2)^2} - 2\sqrt{(x_n + a_n^2)(x_n + b_n^2)}\right| + \frac{1}{4}\left|b_n^2 - a_n^2\right| \\
&< \frac{1}{4}\left|\frac{(x_n + a_n^2)^2 - (x_n + a_n^2)(x_n + b_n^2)}{\sqrt{(x_n + a_n^2)(x_n + b_n^2)}}\right| + \frac{1}{4}\left|b_n^2 - a_n^2\right| \qquad (^1) \\
&= \frac{1}{4}\left(\left|\frac{(x_n + a_n^2)}{\sqrt{(x_n + a_n^2)(x_n + b_n^2)}}\right| + 1\right)|b_n^2 - a_n^2| \\
&= \frac{1}{4}\left(\sqrt{\frac{(x_n + a_n^2)}{(x_n + b_n^2)}} + 1\right)|b_n + a_n|\varepsilon_n \\
&\le \frac{1}{4}\left(\sqrt{\frac{(x_N + a_N^2)}{(x_N + b_N^2)}} + 1\right)2a_N\varepsilon_n
\end{aligned}$$

for any $N \le n$. Let $C(N) = \frac{1}{2}\left(\sqrt{\frac{(x_N+a_N^2)}{(x_N+b_N^2)}}+1\right)a_N$. Choose N sufficiently large so that $\frac{C(N)\varepsilon_N}{x_N+a_N^2} < \frac{1}{2}$ and $N-1 < 2^{N-1}$. Then

$$\begin{aligned}
\left|\sum_{k=N}^{\infty} 2^{k-1}\log\frac{x_{k+1}+a_{k+1}^2}{x_k+a_k^2}\right| &\le \sum_{k=N}^{\infty}\left|2^{k-1}\log\frac{x_{k+1}+a_{k+1}^2}{x_k+a_k^2}\right| \\
&\le \sum_{k=N}^{\infty}\left|2^{k-1}\log\left(1-\frac{C(N)\varepsilon_k}{x_k+a_k^2}\right)\right| \\
&\le \sum_{k=N}^{\infty}\left|2^{k}\frac{C(N)\varepsilon_k}{x_k+a_k^2}\right| \\
&\le \sum_{k=N}^{\infty}\left|2^{k-2^k-1}\frac{C(N)}{b_N^2}(a_0-b_0)\right| \\
&\le \left|\frac{C(N)}{b_N^2}(a_0-b_0)\right|\sum_{k=N}^{\infty}2^{-2^{k-1}} \\
&\le \left|\frac{C(N)}{b_N^2}(a_0-b_0)\right|2^{1-2^{N-1}}.
\end{aligned}$$

This lets us compute $\mu_E(P_0)$ to $O(2^{2^N})$ bits of precision in N steps, as desired.

To handle the case P is not on the identity component, one can compute $\mu_E(P)$ from the relation $\mu_E(2P_0) = 4\mu_E(P) - \log|2y(P))|$. If the curve does not have full 2-torsion over $\mathbb{R}$ one can, after a change of coordinates, write the equation as $y^2 = x(x^2+ux+v)$ and use the isogony

$$(x,y) \mapsto \left(\frac{x^2+ux+v}{x}, y\frac{x^2-v}{x^2}\right)$$

to the curve E_0 given by $y_0^2 = x_0(x_0^2 - 2ux_0 + u^2 - 4v)$ which does have full 2-torsion over $\mathbb{R}$. The relationship between the heights is then $\mu_{E_0}(P_0) = 2\mu_E(P) - \frac{1}{2}\log|x|$.

Using these algorithms, we were able to verify, assuming trivial $Ш(E/\mathbb{Q})$, that the BSD formula holds for the elliptic curve given by $y^2+y = x^3+x^2-2x$ of rank 2 and conductor 389 to 10,000 bits of precision. This computation took about a week and was dominated by computing $L''(E,1)$; given the cubic dependance on precision, it does not seem feasible to push this out to a million digits.

[1] The mean value theorem lets us see that for any $0 < u < v$ we have the estimate $\sqrt{v}-\sqrt{u} < \frac{1}{2\sqrt{u}}(v-u)$.

6.2 Deducing functional equation parameters

The most common use case for evaluating L-functions and their derivatives is to study their behavior at special values which hold special arithmetic significance. However, as pointed out in section 7 of [13], the ability to evaluate a (hypothetical) L-function at arbitrary points in the complex plane allows one to numerically verify whether a functional equation holds for a given set of data. This can be useful when some of the invariants are difficult to compute or even unknown in general. For example, if one has all the data for an L-function except for its sign, one can try $\epsilon = +1$ and $\epsilon = -1$ and see which of the two cases hold for various values of s. Only one value of ϵ will make the functional equation hold for all s. It is often the case that one knows the general form of the missing data, for example everything may be easy to determine except for the powers to which the bad primes divide the conductor and their Euler factors, and things can be narrowed down to a finite list of possibilities. In this case, eliminating all but one possibility via rigorous computation provides a *proof* that the single remaining possibility is correct. As a computational note, rather than verifying $\Lambda(s) = \varepsilon\Lambda(w-s)$ it is cheaper to verify (or disprove)

$$\Theta(1/t) = \varepsilon t^w \Theta(t) - \sum_j r_j t^{p_j}$$

holds for $1 < t < \infty$.

BIBLIOGRAPHY

[1] Dominique Bernardi. Hauteur p-adique sur les courbes elliptiques. In *Seminar on Number Theory, Paris 1979–80*, volume 12 of *Progr. Math.*, pages 1–14. Birkhäuser Boston, Mass., 1981.

[2] Joseph (ed.) Bernstein and Stephen (ed.) Gelbart. *An introduction to the Langlands program. Lectures presented at the Hebrew University of Jerusalem, Jerusalem, Israel, March 12–16, 2001.* Boston, MA: Birkhäuser. viii, 281 p. EUR 48.00/net; sFr. 72.00 , 2003.

[3] Massimo Bertolini and Henri Darmon. Kolyvagin's descent and Mordell-Weil groups over ring class fields. *J. Reine Angew. Math.*, 412:63–74, 1990.

[4] Jean-Benoît Bost and Jean-François Mestre. Calcul de la hauteur archimedienne des points dune courbe elliptique par un algorithme quadratiquement convergent et application au calcul de la capacite de lunion de deux intervalles.

[5] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor. On the modularity of elliptic curves over $\mathbf{Q}$: wild 3-adic exercises. *J. Amer. Math. Soc.*, 14(4):843–939 (electronic), 2001.

[6] J. W. S. Cassels. Arithmetic on curves of genus 1. IV. Proof of the Hauptvermutung. *J. Reine Angew. Math.*, 211:95–112, 1962.

[7] Byungchul Cha. Vanishing of some cohomology groups and bounds for the Shafarevich-Tate groups of elliptic curves. *J. Number Theory*, 111(1):154–178, 2005.

[8] Henri Cohen. *A course in computational algebraic number theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993.

[9] Henri. Cohen. *Number theory II: Analytic and modern methods*, volume 240 of *Graduate Texts in Mathematics*. Springer, 2007.

[10] J. E. Cremona. *Algorithms for modular elliptic curves*. Cambridge University Press, Cambridge, second edition, 1997.

[11] J. E. Cremona, M. Prickett, and Samir Siksek. Height difference bounds for elliptic curves over number fields. *J. Number Theory*, 116(1):42–68, 2006.

[12] C. Delauney. Formes modulaires et invariants de courbes elliptiques définies sur **Q**. *Université Bordeaux I, PhD thesis*, 2002.

[13] T. Dokchitser. Computing special values of motivic L-functions. *Experiment. Math.*, 13(2):137–149, 2004.

[14] Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: harmonic sums. *Theoret. Comput. Sci.*, 144(1-2):3–58, 1995. Special volume on mathematical analysis of algorithms.

[15] Laurent Fousse, Guillaume Hanrot, Vincent Lefèvre, Patrick Pélissier, and Paul Zimmermann. *MPFR: A multiple-precision binary floating-point library with correct rounding*, 2010. `http://www.mpfr.org`.

[16] Stephen Gelbart. An elementary introduction to the Langlands program. *Bull. Am. Math. Soc., New Ser.*, 10:177–219, 1984.

[17] Grigor Grigorov, Andrei Jorza, Stefan Patrikis, William A. Stein, and Corina Tarniţă. Computational verification of the Birch and Swinnerton-Dyer conjecture for individual elliptic curves. *Math. Comp.*, 78(268):2397–2425, 2009.

[18] B. Gross. Heegner points on $X_0(N)$. In *Modular forms (Durham, 1983)*, Ellis Horwood Ser. Math. Appl.: Statist. Oper. Res., pages 87–105. Horwood, Chichester, 1984.

[19] B. Gross, W. Kohnen, and D. Zagier. Heegner points and derivatives of L-series. II. *Math. Ann.*, 278(1-4):497–562, 1987.

[20] Benedict H. Gross. Kolyvagin's work on modular elliptic curves. In *L-functions and arithmetic (Durham, 1989)*, volume 153 of *London Math. Soc. Lecture Note Ser.*, pages 235–256. Cambridge Univ. Press, Cambridge, 1991.

[21] Yoshiki Hayashi. The Rankin's L-function and Heegner points for general discriminants. *Proc. Japan Acad. Ser. A Math. Sci.*, 71(2):30–32, 1995.

[22] Dimitar Jetchev. Global divisibility of Heegner points and Tamagawa numbers. *Compos. Math.*, 144(4):811–826, 2008.

[23] Dimitar Jetchev, Kristin Lauter, and William Stein. Explicit Heegner points: Kolyvagin's conjecture and non-trivial elements in the Shafarevich-Tate group. *J. Number Theory*, 129(2):284–302, 2009.

[24] Kazuya Kato. p-adic Hodge theory and values of zeta functions of modular forms. *Astérisque*, (295):ix, 117–290, 2004. Cohomologies p-adiques et applications arithmétiques. III.

[25] Brian W. Kernighan and P. J. Plauger. *The Elements of Programming Style*. McGraw-Hill, Inc., New York, NY, USA, 1982.

[26] Donald E. Knuth. *The art of computer programming. Vol. 2*. Addison-Wesley Publishing Co., Reading, Mass., third edition, 1998. Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing.

[27] J. C. Lagarias and A. M. Odlyzko. On computing Artin L-functions in the critical strip. *Math. Comp.*, 33(147):1081–1095, 1979.

[28] Serge Lang. *Number theory. III*, volume 60 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1991. Diophantine geometry.

[29] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261(4):515–534, 1982.

[30] Yudell L. Luke. *The special functions and their approximations. Vol. II*. Mathematics in Science and Engineering, Vol. 53. Academic Press, New York, 1969.

[31] Robert Miller. Empirical evidence for the Birch and Swinnerton-Dyer conjecture. *University of Washington, PhD thesis*, 2010.

[32] Ramon E. Moore. *Methods and applications of interval analysis*, volume 2 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1979.

[33] Michael Rubinstein. Evidence for a spectral interpretation of the zeros of l-functions. *Princeton, PhD thesis*, 1998.

[34] Michael Rubinstein. *L: C++ class library and command line program for computing zeros and values of L-functions*, 2010. `http://pmmac03.math.uwaterloo.ca/~mrubinst/L_function_public/L.html`.

[35] A. Schönhage and V. Strassen. Schnelle Multiplikation grosser Zahlen. *Computing (Arch. Elektron. Rechnen)*, 7:281–292, 1971.

[36] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[37] William Stein. Toward a generalization of the Gross-Zagier Conjecture. *IMRN*, 2009. To appear.

[38] William Stein et al. *Sage Mathematics Software (Version 4.4)*. The Sage Development Team, 2010. `http://www.sagemath.org`.

[39] Fredrik Strömberg. Computation of Maass waveforms with nontrivial multiplier systems. *Math. Comp.*, 77(264):2375–2416, 2008.

[40] Mark Watkins. Some remarks on heegner point computations. 2005.

[41] Serge Winitzki. Computing the incomplete gamma function to arbitrary precision. In *Computational science and its applications—ICCSA 2003. Part I*, volume 2667 of *Lecture Notes in Comput. Sci.*, pages 790–798. Springer, Berlin, 2003.

[42] S. W. Zhang. Gross-Zagier formula for GL_2. *Asian J. Math.*, 5(2):183–290, 2001.

VITA

Robert Bradshaw is married to Camille Bradshaw and is the father of two adorable twin girls Lydia and Ruby. Robert recieved a Bachelor degree, majoring in Mathematics and Linguistics, from Brigham Young University in 2004. He received the degree of Doctor of Philosophy in Mathematics from the University of Washington in 2010.