ASIAN J. MATH.
Vol. 4, No. 1, pp. 183–212, March 2000

018

LANDAU-SIEGEL ZEROES AND BLACK HOLE ENTROPY*

STEPHEN D. MILLER† AND GREGORY MOORE‡

There has been some speculation about relations of D-brane models of black holes to arithmetic. In this note we point out that some of these speculations have implications for a circle of questions related to the generalized Riemann hypothesis on the zeroes of Dirichlet L-functions.

1. Introduction. In [1] some connections were made between string theoretic models of black holes and certain issues in arithmetic. One of the more speculative suggestions in [1] was a proposal that the entropy of BPS black holes is related to an arithmetic height of certain arithmetic varieties. In the present note we remark on some connections between these speculations and some issues related to the Riemann hypothesis. In particular, the speculations seem most relevant to the question of "Landau-Siegel zeroes," which are hypothetical zeroes of L-functions very close to $s = 1$. (A precise definition is given in definition 4.4.)

In section 2 we review the theory of Strominger-Vafa [2]. In section 3 we summarize, reformulate, and sharpen the statements from [1] whose implications we wish to explore. In section 4 we provide some background information on analytic number theory and Dirichlet L-functions, and in section 5 we see how everything fits together. In particular, we discuss a close interplay between the Strominger-Vafa prediction about black hole entropy, the Landau-Siegel zero, and the minimal discriminant of an elliptic curve with complex multiplication.

Warnings: This note is written for a mixed audience of both string theorists and analytic number theorists, so some trivial things are explained. We caution the reader at the outset that the evidence for the Height Conjecture of [1] is slim, to say the least. Thus, this paper should be regarded as an exercise in Pascal's Wager.

2. Summary of the attractor mechanism and the Strominger-Vafa proposal. Let X be a compact Calabi-Yau 3-fold, and let $\gamma \in H_3(X; \mathbb{Z})$ be an integral homology class. String theory associates two interesting mathematical objects to the pair (X, γ):

a.) A finite-dimensional Hermitian vector space $\mathcal{H}(\gamma)$.

b.) Another Calabi-Yau X_γ, in the same complex structure moduli space as X.

We can interpret (a) and (b) both mathematically and physically. The physical setting is the theory of BPS black holes in $d = 4, \mathcal{N} = 2$ compactifications of type II strings on a Calabi-Yau manifold X. We now explain (a) and (b) in a little more detail.

First consider (a). In physics $\mathcal{H}(\gamma)$ is the space of BPS states of charge γ. The definition of this space has not been completely rigorously formulated mathematically, although this should be possible using the theory of D-branes. Very roughly speaking $\mathcal{H}(\gamma)$ should be defined mathematically as follows. Consider the moduli space $\mathcal{M}(\gamma)$ of pairs (Σ, A) where Σ is a smooth special Lagrangian submanifold of X in the homology

* Received December 11, 1999; accepted for publication January 13, 2000. hep-th/9903267, YCTP-P8-99, IASSNS-HEP-99/34

† Department Of Mathematics, Yale University, New Haven, Connecticut 06520, USA (steve@math.yale.edu).

‡ Department Of Physics, Yale University, New Haven, Connecticut 06520, USA and School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540 (gmoore@physics.rutgers.edu).

class γ, and A is a flat $U(1)$ connection on Σ. The moduli space $\mathcal{M}(\gamma)$ inherits a metric from the Calabi-Yau metric on X, and $\mathcal{H}(\gamma)$ is the L^2-cohomology of $\mathcal{M}(\gamma)$.[1]

Now consider (b). By Yau's theorem, a Calabi-Yau manifold may be specified by its complex structure and its Kähler class. X will belong to a family of Calabi-Yau manifolds with complex structures in $\mathcal{M}_{cplx}$. (For what follows we will need to work on the universal cover $\widetilde{\mathcal{M}}_{cplx}$.) The map (b) is provided by the "attractor mechanism" of Ferrara, Kallosh, and Strominger [3][4][5]. For each γ there is a dynamical system on the moduli space $\widetilde{\mathcal{M}}_{cplx}$ and, for suitable γ,[2] the dynamical system will have a unique fixed point X_γ. Associated with the Calabi-Yau X_γ is the normalized period

$$|\mathcal{Z}(\gamma)|^2 := \frac{|\int_\gamma \Omega_\gamma|^2}{|\int_{X_\gamma} \Omega_\gamma \wedge \bar{\Omega}_\gamma|} \tag{2.1}$$

where Ω_γ is a nowhere vanishing holomorphic $(3,0)$ form on X_γ.

The connection between objects (a) and (b) is provided by the Strominger-Vafa proposal for the microstates accounting of black hole entropy [2]. The idea may be summarized in the following four steps:

1. Given a Calabi-Yau X, one may write a system of partial differential equations for a Minkowski-signature 4-manifold M equipped with certain geometrical data (e.g. a rank $\frac{1}{2}b_3(X)$ torus bundle with connection). These equations generalize the Einstein equations on M and are called the supergravity equations. A choice of charge vector γ is equivalent to a choice of boundary conditions for these equations. For appropriate vectors γ (those for which the normalized period has an isolated minimum in $\widetilde{\mathcal{M}}_{cplx}$), the supergravity equations admit black hole solutions $\mathcal{B}(\gamma)$. Using the supergravity equations one may compute the horizon area $A(\gamma)$ of the black hole solution $\mathcal{B}(\gamma)$. It turns out that
$$A(\gamma) = 4\pi|\mathcal{Z}(\gamma)|^2. \tag{2.2}$$
We work in Planck units $\ell_{planck} = 1$.
2. According to Bekenstein and Hawking a black hole is a thermodynamical object. It has a temperature and an entropy, and the latter is given by:
$$S_{\text{sugra}}(\gamma) = \frac{A(\gamma)}{4}. \tag{2.3}$$
3. According to Boltzmann and Planck, we have the exact formula:
$$S_{\text{micro}}(\gamma) = k\log[W(\gamma)] \tag{2.4}$$
where $W(\gamma)$ is the dimension of the space of available states in the microcanonical ensemble specified by the charges γ and the energy $M = |\mathcal{Z}(\gamma)|$. Here k is Boltzmann's constant; we henceforth choose units with $k = 1$.
4. According to Strominger and Vafa, for charge vectors γ which are in some sense large, physical states formed by BPS configurations of D3-branes in X with charge γ can be described macroscopically, in a supergravity approximation, by the black hole solutions $\mathcal{B}(\gamma)$ to the supergravity equations. From the microscopic, D-brane viewpoint we identify $W = \dim \mathcal{H}(\gamma)$, for all γ, large or small. The correspondence

[1] In the mirror formulation γ would specify the Chern classes of a coherent sheaf in the mirror Calabi-Yau $\tilde{X}$, and $\mathcal{M}(\gamma)$ would be the L^2 cohomology of the moduli space of such sheaves.

[2] Some discussion of which γ are suitable is given in [1].

between configurations of D-branes and the black hole solutions is formulated mathematically as the statement that for large γ

$$S_{\text{micro}}(\gamma) \sim S_{\text{sugra}}(\gamma). \tag{2.5}$$

Putting all the above statements together one arrives at the mathematically startling proposal that:

$$\log[\dim \mathcal{H}(\gamma)] \sim \pi |\mathcal{Z}(\gamma)|^2 \tag{2.6}$$

for large γ. Both sides of equation (2.6) are susceptible of precise mathematical definition. Moreover, such a statement would be a deep and surprising mathematical fact. Before that becomes a reality, several clarifications of the meaning of (2.6) must be carried out. In particular we need:

1. A precise definition of $\mathcal{H}(\gamma)$ and a definition of $\dim \mathcal{H}(\gamma)$ (since $\mathcal{H}(\gamma)$ is a graded vector space, we might well have to use a graded dimension.)
2. A precise statement of the meaning of "large γ."
3. A precise statement of the meaning of the asymptotic symbol $\sim$.

It is quite necessary to take a limit of "large γ," in order to justify the supergravity approximation to string theory, and thereby the connection to black holes. We will take it to mean operationally that we consider sequences γ_n of charges such that $z_n := |\mathcal{Z}(\gamma_n)| \to \infty$. We refer to such a sequence of charges as a *big sequence.*

Another subtlety which we would like to mention is the concept of *effective* asymptotics. For example, one could make a statement that a certain quantity is bounded, but without being able to determine anything about the nature of the bound or even being able to begin computing it. We will assume that the limiting behaviors of the physical quantities are effective: that is, any time a constant (implicit or named) is asserted to exist, one could furthermore explicitly compute or name such a constant.

2.1. Versions of the Strominger-Vafa conjecture. Point 3 above (below (2.6)) has not really been addressed in the literature. The Strominger-Vafa conjecture (2.6) in fact admits several inequivalent formulations. Two plausible versions of the conjecture are:

SSV Conjecture. (*Strong SV conjecture*): If $\{\gamma_n\}$ is a big sequence then $S_{\text{micro}}(\gamma)$ has an asymptotic expansion in $z_n := |\mathcal{Z}(\gamma_n)|$. More precisely,

$$\frac{S_{\text{micro}}(\gamma_n)}{\pi z_n^2} = 1 + \mathcal{O}\left(\frac{\log z_n}{z_n^2}, \frac{1}{z_n^\delta}\right), \tag{2.7}$$

where we use $\mathcal{O}$-notation in the precise sense of asymptotics of sequences in n, and δ (a quantity discussed below) is positive.

This is an expectation based on the supergravity approach to black hole entropy for CY 3-fold black holes. See, for examples, the discussions in [6][7][8]. The corrections arise from both M-theory corrections to the 11D supergravity and from quantum effects associated with the compactification. The quantity δ in (2.7) depends on details of the sequence γ_n, e.g., on the details of the directions of the γ_n in $H^3(X;\mathbb{Z})$ as z_n goes to infinity. The SSV conjecture might appear to be obvious from the point of view of supergravity. However, a systematic scheme for calculating corrections to the leading order supergravity approximation to M-theory is unkown. Moreover, it is known that supergravity, which treats charges as continuous, can miss subtle arithmetic properties. (For example, the number of U-duality classes of γ's with a fixed value of $|\mathcal{Z}(\gamma)|$ can be 1 in supergravity, but in fact is given by class numbers in the exact formulation [1].) The SSV conjecture is also suggested by the proposed formulae for exact results on $S_{\text{micro}}(\gamma)$ in the case when $X = S \times E$ with S a K3 surface and E an elliptic curve [9]. Using the Cardy formula[3] one can justify the SSV conjecture for this special class of Calabi-Yau 3-folds. However, the existing proposals, based on elliptic genera of symmetric products are only firmly justified for 5D black holes, and the extension to 4D black holes - even for compactification on a six-dimensional torus - is nontrivial. Moreover, the generalization of the existing proposals for $\dim \mathcal{H}(\gamma)$ to generic Calabi-Yau 3-folds will be much more subtle and intricate.

For all these reasons, we are also willing to entertain the

WSV Conjecture. (*Weak SV conjecture*): If $\{\gamma_n\}$ is a big sequence then

$$\lim_{n\to\infty} \frac{\log S_{\text{micro}}(\gamma_n)}{\log z_n^2} = 1. \tag{2.8}$$

Of course, the SSV conjecture implies the WSV conjecture, but the converse is false. The motivation behind the formulation of the WSV conjecture is the following. It might be that for large γ, the microscopic entropy actually behaves like

$$S_{\text{micro}}(\gamma_n) \sim \pi z_n^2 (\log z_n)^{\alpha}, \tag{2.9}$$

where α is a random variable, chosen using a measure on $H_3(X;\mathbb{Z})$. Presumably the distribution would become more sharply concentrated around $\alpha = 0$ as γ becomes larger. As far as we know there is nothing wrong with this idea, and from some viewpoints it even seems likely. For example, the existence of BPS states can depend on arithmetic properties of γ. So there might well be a large amount of "scatter" and "noise" in the behavior of $\dim \mathcal{H}(\gamma)$ as $\gamma \to \infty$. (See Fig. 1 in section 5 below for an illustration of the kind of scatter we mean.) In this case the Bekenstein-Hawking formula would not be the leading term in a systematic expansion, but would only hold in some average sense.

If the dimensions $\dim \mathcal{H}(\gamma)$ indeed behave in the way just suggested, then it is quite possible that the limit (2.8) does not exist in the standard sense. In such a situation it is more appropriate to use the lim-sup, (also denoted as $\overline{\lim}$) which *always exists* for any sequence of real numbers.[4] Thus we could replace the limits in the WSV conjecture by $\underline{\lim}$ or $\overline{\lim}$. We refer to these as the $\underline{WSV}$ and $\overline{WSV}$ conjectures, respectively.

There are yet other inequivalent formulations of the SV conjecture.

3. The height conjecture.

3.1. Sharpening the height conjecture. The main conjectures of [1] (conjectures 8.2.1, 8.2.2, and 8.2.3) posit that the attractor variety X_γ can be defined as an arithmetic variety over some number field K_γ. This was verified in [1] in some special cases, for example when X is a free quotient of $K3 \times E$ (where E is an elliptic curve), or when X is a complex torus. In a much more speculative section, [1] also suggested a

[3] The Cardy formula is a generalization of the Hardy-Ramanujan formula for the asymptotics of the partition function and gives the asymptotics of conformal field theory partition functions.

[4] We recall the definition. If a_n is any sequence of real numbers let $b_n = \{a_n, a_{n+1}, \dots\}$. Then let $c_n := \sup b_n$ be the smallest constant bounding b_n from above. Plainly, the c_n form a strictly decreasing sequence of real numbers. We define $\overline{\lim}\, a_n = \lim\, c_n$.

possible connection between the Faltings height for a metrized line bundle $\mathrm{ht}(X_\gamma; K_\gamma)$ and the entropy. We want to investigate the consequences of these latter conjectures.

We begin by making the "height conjecture" more precise. The first point to note is that if an attractor variety X_γ indeed satisfies the attractor conjectures, then it might nevertheless admit several different arithmetic "models" with quite different arithmetic properties. First of all, changing the field of definition of X can alter the arithmetic properties. We will not investigate the issues of "base change" systematically, although we note that the Faltings height does stabilize under field extension [10]. A second ambiguity, of more direct relevance to what follows, is that two varieties X, X' over a field K might be isomorphic as varieties over an algebraic closure $\bar{K}$ (such as $\bar{K} = \mathbb{C}$ for a number field), but fail to be isomorphic as varieties over K. We will quote some relevant examples in section 3.3 below.

As with the SV conjecture, there are various inequivalent precise formulations of the height conjecture:

Strong Height (SH) Conjecture: Assume a family of arithmetic attractor varieties $(X_{\gamma_n}, K_{\gamma_n})$ corresponds to a big sequence, that is, $z_n := |\mathcal{Z}(\gamma_n)| \to \infty$. Then there exists a finite positive constant κ (possibly depending on the family) and a choice of model for X_{γ_n} such that if $\mathrm{ht}(X_\gamma; K_\gamma)$ is the Faltings height for the metrized line bundle provided by the Calabi-Yau data, then

$$\exp\left[\frac{1}{\kappa}\mathrm{ht}(X_{\gamma_n}; K_{\gamma_n})\right] = S_{\mathrm{micro}}(\gamma_n) + o(|z_n|^2). \tag{3.1}$$

Alternatively, with the same hypotheses, we can formulate the

Weak Height (WH) Conjecture:

$$\lim_{n\to\infty} \frac{\mathrm{ht}(X_{\gamma_n}; K_{\gamma_n})}{\log[S_{\mathrm{micro}}(\gamma_n)]} = \kappa. \tag{3.2}$$

Of course, since not much is known about either numerator or denominator in (3.2), it is prudent to allow alternative formulations $\overline{WH}$, and $\underline{WH}$ of the conjectures, with constants $\overline{\kappa}$ and $\underline{\kappa}$, respectively. If the WH or SH conjectures turn out to be true, then the extent to which the constant κ depends on the family of charges will become an interesting question, as will become apparent in section 5.

Let us return to the ambiguities of base-change and choice of model for X_γ. Regarding base change, our hope is that the choice of base field, while necessary to define the heights, should do no more than change the value of the constant κ in the height conjectures.

Regarding the choice of model, the attractor equation of supergravity only specifies the attractor variety X_γ as a variety over $\mathbb{C}$. If X_γ can be defined over a number field K_γ, then since there can be inequivalent models over K_γ we must ask if the choice of model has any relevance for the physics of string compactification on X. A physicist's natural reaction to this question would be that the choice of model should be irrelevant, by general covariance.[5] However, as seen in [1], the arithmetic of the number fields associated to attractor varieties *is* related to such physical quantities

[5] Roughly speaking, general covariance refers to invariance of a theory of gravity under C^∞ diffeomorphisms. A blind application of this principle would also suggest that complex structure is physically irrelevant (which is hardly the case). The main thesis of this exploratory paper is that not only complex structure, but even arithmetic structure is physically significant.

as the BPS mass spectrum. Whether or not the choice of arithmetic model is also of physical relevance remains to be seen. In fact, the height conjectures above are the first instance, of which we are aware, in which such a choice really matters. This raises the interesting question of whether physics indeed selects a distinguished arithmetic model.[6]

3.2. A special class of attractor varieties. Some weak evidence for the height conjecture was given in [1] in the case when X is a complex torus. In this case it was shown that X_γ is isogenous to a product of 3 elliptic curves with complex multiplication:

$$X_\gamma \sim E_{\tau(\gamma)} \times E_{\tau(\gamma)} \times E_{\tau(\gamma)} \tag{3.3}$$

where $E_\tau := \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ and $\tau(\gamma) = i\sqrt{I_4(\gamma)}$. Here $I_4(\gamma)$ is an integral quartic polynomial on $H_3(X;\mathbb{Z})$ related to the quartic $E_{7,7}$ invariant (see [1] for precise definitions). Moreover, one easily verifies in this example that:

$$|\mathcal{Z}(\gamma)|^2 = \sqrt{I_4(\gamma)}, \tag{3.4}$$

while the supergravity analysis shows that indeed $S_{\text{sugra}}(\gamma) = \pi\sqrt{I_4(\gamma)}$ [12][13][14].

Equation (3.3) is only a statement about isogeny classes. In order to estimate the height it is more convenient (but not absolutely necessary) to pin down the attractor variety exactly. We will eliminate the unkown isogeny by choosing "diagonal charges." That is, using equations (6.2)–(6.5) of [1], with $P^{ij} = p^i\delta^{ij}, Q_{ij} = q_i\delta_{ij}$ in the notation of that paper, we may choose lattice vectors $\gamma \in H^3(X;\mathbb{Z})$ depending on eight integers $r, p^1, p^2, p^3, s, q_1, q_2$, and q_3 such that we have the equality

$$X_\gamma = E_{\tau_1} \times E_{\tau_2} \times E_{\tau_3}, \tag{3.5}$$

where now we have

$$I_4(\gamma) = 2\big[(\sum_{i=1}^{3} p^i q_i)^2 - \sum_{i=1}^{3}(p^i q_i)^2\big] - (rs + \sum_{i=1}^{3} p^i q_i)^2 + 4\big(rq_1q_2q_3 - sp^1p^2p^3\big) \tag{3.6}$$

and, defining $D = -I_4 < 0$, τ_i are given by:

$$\tau_1 = \frac{2(p^1q_1 - p^2q_2 - p^3q_3 - rs) + \sqrt{D}}{4(p^2p^3 + rq_1)} \tag{3.7}$$

with cyclic permutations on 123 giving the formulae for τ_2, τ_3.

Now let us discuss some arithmetic associated with the elliptic curves E_{τ_i}. In general, if τ is of the form $(-b+\sqrt{D})/2a$ for integral a, b, D, with $D < 0$, then it follows that $\mathbb{Z} + \tau\mathbb{Z}$ is a proper fractional ideal for some order $\mathcal{O}$ of the field $K_D := \mathbb{Q}(\sqrt{D})$. To be more precise, let the minimal polynomial of τ be $A\tau^2 + B\tau + C = 0$ where A, B, C are integral. (Note: it need not be true that $a = A$, or $b = B$.) Then by [15], Lemma 7.5, $\mathbb{Z} + \tau\mathbb{Z}$ is a proper fractional ideal for the order $\mathcal{O} = \mathbb{Z} + A\tau\mathbb{Z}$ of K_D. In general, the conductor of the order $\mathcal{O}$ will be larger than one. By [15], Theorem 11.1, the value of the modular function $j(\tau)$ generates the ring class field of this order, i.e., the ring class field is $K_D(j(\tau))$. It follows from (3.5) that X is at least defined over

[6] A good place to start thinking about this might be Witten's linear sigma model formulation of CY sigma models, where a definite choice of projective model (albeit over $\mathbb{C}$) is made by the quantum field theory [11].

the compositum of three such ring class fields (it might well be defined over a smaller field).

We would like to eliminate some of the complications of general ring class fields and work instead with the Hilbert class field $\widehat{K}_D$ of K_D. This will simplify the height computation below. As explained at length in [15], the distinct ideal classes of the ring of integers $\mathcal{O}(K_D)$ of K_D have representatives $\mathbb{Z} + \tau_k\mathbb{Z}$. Here $k = 1, \ldots, h(D)$, and $h(D)$ is the class number of K_D. The complex numbers τ_k have the form $\tau_k = (-b_k + \sqrt{D})/2a_k$ and are the solutions in the upper half-plane to $a_k\tau^2 + b_k\tau + c_k = 0$ where $a_kx^2 + b_kxy + c_ky^2$ runs over representatives of inequivalent primitive binary quadratic forms of discriminant D. We can choose representatives in the standard "keyhole" fundamental domain for $PSL(2,\mathbb{Z})$. We will refer to the τ_k as "Heegner points." One of the beautiful statements of the theory of complex multiplication is that for any k, $j(\tau_k)$ is an algebraic integer and $\widehat{K}_D \cong K_D(j(\tau_k))$.

Now let us consider some explicit charge vectors γ. First we take $p = p^1 = p^2 = p^3$, and $q = q_1 = q_2 = q_3$ so that we only have to work with a single elliptic curve with modular parameter $\tau = \tau_1 = \tau_2 = \tau_3$ given by

$$\tau = \frac{-2(pq + rs) + \sqrt{D}}{4(p^2 + rq)}. \tag{3.8}$$

Here D reduces to $-I_4$, with I_4 given by

$$I_4(p, q, r, s) = 12p^2q^2 + 4(rq^3 - sp^3) - (3pq + rs)^2. \tag{3.9}$$

Next we choose p, q, r, s so that τ is one of the Heegner points. We will content ourselves with finding charge vectors γ_D which yield the principal class (the trivial class). This is given by

$$\tau = \frac{-1 + \sqrt{D}}{2} \tag{3.10}$$

if $D = 1 \bmod 4$ and $\tau = \sqrt{D}$ if $D = 0 \bmod 4$. Specifically, if we choose $r = 0$ and $s = 3 + 4D$, then with $p = q = 4$, we get (3.10). If $p = 1, q = 2$, we get

$$\tau = -1 + \sqrt{D}, \tag{3.11}$$

which is clearly $PSL(2,\mathbb{Z})$-equivalent to $\tau = \sqrt{D}$.

Presumably, other charges γ can lead to other ideal classes in K_D, but we will focus on the above sequence of charges in this paper, and will call them γ_D. As explained in the next section, we will take $\widehat{K}_D$ as the field of definition of the special attractor varieties associated with the charges γ_D.

3.3. Silverman's formula for the height of an arithmetic elliptic curve.

Now that we have focused on the attractors X_{γ_D} let us compute their height. Since X_{γ_D} is a product of 3 elliptic curves we have

$$\mathrm{ht}(X_{\gamma_D}; \widehat{K}_D) = 3\mathrm{ht}(E_{\tau(\gamma_D)}; \widehat{K}_D), \tag{3.12}$$

and therefore we need only compute the height of an elliptic curve. This has been done in a very explicit way by Silverman in chapter 10 of [10]. We now review his formula. In the next section we apply the formula to our considerations.

To begin, we must review a few standard definitions. (See any textbook on elliptic curves, for examples [16][17][18].) Let K be a field. An elliptic curve E/K can be given by a Weierstrass model

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.13}$$

where $a_i \in K$. Two equivalent Weierstrass models for the same curve over K are related by a change of coordinates

$$\begin{aligned} x' &= u^2x + r \\ y' &= u^3y + su^2x + t \end{aligned} \tag{3.14}$$

with $u \in K^*, r, s, t \in K$.

We introduce the discriminant Δ through the standard definitions:

$$\begin{aligned} b_2 &:= a_1^2 + 4a_2 \\ b_4 &:= a_1a_3 + 2a_4 \\ b_6 &:= a_3^2 + 4a_6 \\ b_8 &:= a_1^2a_6 - a_1a_3a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2 \\ c_4 &:= b_2^2 - 24b_4 \\ c_6 &:= -b_2^3 + 36b_2b_4 - 216b_6 \\ \Delta &:= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6 \end{aligned} \tag{3.15}$$

If the elliptic curve is nonsingular then $\Delta \neq 0$ and we can define the j-invariant:

$$j := c_4^3/\Delta. \tag{3.16}$$

When the characteristic is not 2 or 3 it is useful to note that $2^63^3\Delta = c_4^3 - c_6^2$.

Under the change of variables (3.14) we have:

$$\begin{aligned} c_4' &= u^4c_4 \\ c_6' &= u^6c_6 \\ \Delta' &= u^{12}\Delta \\ j' &= j \end{aligned} \tag{3.17}$$

Curves with the same j-invariant are isomorphic over the algebraic closure of K, but need not be isomorphic over K.

Now let K be a number field and E an elliptic curve over K. The field K has a set of valuations. These consist of archimedean valuations ("the places at infinity") and nonarchimedean valuations ("the finite places"). The places at infinity correspond to inequivalent embeddings of K into $\mathbb{R}$ or $\mathbb{C}$. In our example below all the embeddings will be complex embeddings $\psi_i : K \hookrightarrow \mathbb{C}$, so we henceforth assume this is the case for K. Under these embeddings $E(\mathbb{C})$ will be isomorphic to $\mathbb{C}/(\mathbb{Z} + \tau_i\mathbb{Z})$. The finite places correspond to valuations labelled by the different prime ideals in K. Let $\mathbf{p}$ be a prime ideal in K. We can then consider the curve $E_{\mathbf{p}}$ over the $\mathbf{p}$-adic completion $K_{\mathbf{p}}$ which is just given by the same equation for E as before, but now considered over this larger field. Let $(\Delta)_{\mathbf{p}}$ be the discriminant of $E_{\mathbf{p}}$ (this is an ideal in $\mathcal{O}(K_{\mathbf{p}})$). A *minimal model for* $E_{\mathbf{p}}$ (at $\mathbf{p}$) is obtained by making changes of coordinates of the form (3.14) (with the field now taken to be $K_{\mathbf{p}}$) such that the power of $\mathbf{p}$ dividing $(\Delta)_{\mathbf{p}}$ is the smallest possible. For the minimal model let us say $\mathbf{p}^{n_{\mathbf{p}}}|(\Delta)_{\mathbf{p}}$, but $\mathbf{p}^{n_{\mathbf{p}}+1} \nmid (\Delta)_{\mathbf{p}}$. In particular, if E is smooth at $\mathbf{p}$ then $n_{\mathbf{p}} = 0$, and $n_{\mathbf{p}} > 0$ only at those primes which divide (Δ) in the "global model" of E. Taking the product over all the finite places of K defines the *minimal discriminant* of E as an ideal in K:

$$\mathcal{D}_{E/K} := \prod_{\mathbf{p}} \mathbf{p}^{n_{\mathbf{p}}}. \tag{3.18}$$

We can now state Silverman's formula for the height of E/K (Proposition 1.1, p. 254 in [10].)[7] The formula for the height involves the sum of the contributions from the finite and infinite places:

$$12[K:\mathbb{Q}]\mathrm{ht}(E;K) = \log |N_{K/\mathbb{Q}}(\mathcal{D}_{E/K})| - 2\sum_i \log\left[(\Im\tau_i)^6|\eta^{24}(\tau_i)|(2\pi)^{-12}\right] \tag{3.19}$$

where $N_{K/\mathbb{Q}}$ is the norm of the ideal, the second term is the sum over inequivalent complex embeddings of K, and $\eta(\tau)$ is the Dedekind function.

Now, to apply (3.19) we must *choose* a field of definition and a Weierstrass model for our attractor varieties with τ given by (3.10) or (3.11). This will introduce some arbitrariness into our discussion. We motivate our choice as follows. If $j \neq 0, (12)^3$ then one can always write a model for an elliptic curve over $\mathbb{Q}(j)$ with invariant j by taking:

$$y^2 = 4x^3 - \frac{27j}{j-1728}(x+1). \tag{3.20}$$

Thus, one obvious model for the attractor varieties is obtained by making a transformation of the form (3.14) to get

$$\begin{aligned}
y^2 &= 4x^3 - g_2 x - g_3\\
g_2 &= 27j(j-(12)^3)\\
g_3 &= 27j(j-(12)^3)^2\\
\Delta &= g_2^3 - 27g_3^2 = 2^6\cdot 3^{12}\cdot j^2(j-(12)^3)^3.
\end{aligned} \tag{3.21}$$

Note that we are taking the coefficients in the ring of integers $\mathcal{O}_{\hat{K}_D}$. Note too that the Hilbert class field is not the minimal field of definition of this curve. We could take $\mathbb{Q}(j(\tau))$. However, this field has many different conjugates inside the Hilbert class field, depending on which Heegner point τ is taken. Thus we find the Hilbert class field more natural.

It is important to note that (3.21) is not the only Weierstrass model we could choose. We can illustrate the kinds of ambiguities we face in the following two simple sets of examples. Consider two families of elliptic curves over $\mathbb{Q}$ labelled by an integer n

$$\begin{aligned}
y^2 &= x^3 + n\\
y^2 &= x^3 + nx.
\end{aligned} \tag{3.22}$$

The first family has $j=0$, $c_4 = 0, c_6 = -2^5 3^3 n, \Delta = -2^4 3^3 n^2$ and complex multiplication by $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$; the second family has $j = 2^6 3^3, c_4 = -2^3 3n, c_6 = 0, \Delta = -2^6 n^3$ and complex multiplication by $\mathbb{Z}[\sqrt{-1}]$. Nevertheless, if we consider two curves for

[7] We will not define the height. See, for examples, [19][10][20] for definitions. For expository discussions see [21][22]. Curiously, the Faltings height has appeared before in string theory. See, for example, [23], and related papers. We do not understand any connection to this work.

n_1, n_2 then, if n_1/n_2 is not a sixth power in the first example or a fourth power in the second example, the curves are inequivalent over $\mathbb{Q}$ and only become equivalent over extensions of degree 6 and 4, respectively. Moreover, *the minimal discriminants depend on the choice of n.* Less trivial examples of inequivalent Weierstrass models (with $j \neq 0, 1728$) can be gleaned from [24].

3.4. Bounding the height. We now specialize Silverman's formula further and put a bound on the height so that we can test the height conjecture.

In order to specialize Silverman's formula, we note that the values of the j-function at the different Heegner points τ_k define the distinct embeddings of $\widehat{K}_D \hookrightarrow \mathbb{C}$, so we may rewrite Silverman's formula for the height of (3.21) as

$$\begin{aligned} 24h(D)h(E_\gamma/\widehat{K}_D) &= \log|N_{\widehat{K}_D/\mathbb{Q}}(\mathcal{D}_{E_\gamma/\widehat{K}_D})| + \sum_{k=1}^{h(D)} \mathcal{R}(\tau_k) \\ \mathcal{R}(\tau) &:= -2\ \log\left[(\Im\tau)^6|\eta^{24}(\tau)|(2\pi)^{-12}\right]. \end{aligned} \tag{3.23}$$

The sum in (3.23) is over representatives for the ideal class group. In general we will write:

$$\langle f(\tau)\rangle_D := \frac{1}{h(D)} \sum_{k=1}^{h(D)} f(\tau_k) \tag{3.24}$$

for any function f. For $f(\tau) = \mathcal{R}(\tau)$ we denote $R(D) := \langle \mathcal{R}(\tau)\rangle_D$. Thus, for our special charges γ_D

$$\mathrm{ht}(X_{\gamma_D}; \widehat{K}_D) = \frac{1}{8}\left[R(D) + \frac{\log|N_{\widehat{K}_D/\mathbb{Q}}(\mathcal{D}_{E/\widehat{K}_D})|}{h(D)}\right]. \tag{3.25}$$

The term $R(D)$ in (3.25) only depends on γ through the $E_{7,7}$ invariant $D = -I_4(\gamma)$. As we have discussed above, the minimal discriminant can depend on the choice of Heegner point defining the curve (3.21), and even on the choice of $\widehat{K}_D$-isomorphism class for the curve. We have, somewhat arbitrarily, chosen (3.10)(3.11) corresponding to the principal class and moreover have chosen the Weierstrass model (3.21).

Now we can put some useful bounds on the height. The minimal discriminant is a subtle object and is hard to estimate, even for a CM curve. For some CM curves its norm can be as small as 1 (over $F = \mathbb{Q}(j)$ under certain conditions on D – see [25]). In [26] it is calculated for an interesting family of elliptic curves. It is quite possible that further results for the model (3.21) can be obtained from the deep work of Gross and Zagier [27]. Nevertheless, while it is subtle, we *do* know that in the model (3.21) it is a certain integral ideal in $\widehat{K}_D$ which divides the principal ideal (Δ) explicitly calculated in (3.21). From this we may derive some easy inequalities, as in [1]:

$$\frac{1}{8}R(D) \leq \mathrm{ht}(X_{\gamma_D}, \widehat{K}_D) \leq \frac{6}{8}R(D) + \langle f(\tau)\rangle_D \tag{3.26}$$

where

$$f(\tau) = 6\log\left[(\Im\tau)^4|E_4(\tau)|^2\right] + 6\log\left[(\Im\tau)^6|E_6(\tau)|^2\right] + \mathit{const}. \tag{3.27}$$

The first inequality in (3.26) does not depend on the choice of Weierstrass model, but the second does..

The function $f(\tau)$ grows like $60\log(\Im\tau)$ for $\tau \to i\infty$ and is therefore square integrable in the Poincaré measure. It follows from a theorem of W. Duke [28] that the average $\langle f(\tau)\rangle_D$ converges to

$$\langle f(\tau)\rangle_D \to \frac{3}{\pi}\int_{\mathcal{F}} \frac{dxdy}{y^2} f(x+iy) \tag{3.28}$$

as $D \to -\infty$. In particular, the integral and hence the limit is finite.

4. Summary of some analytic number theory. In sections 4.1 and 4.2 we summarize some well-known facts and definitions from analytic number theory. In section 4.3 we summarize some more technical facts needed in section 5.

4.1. *L*-functions.

DEFINITION 4.1: LEGENDRE-JACOBI-KRONECKER SYMBOL. This is the unique, real, nontrivial Dirichlet character of modulus D. Its value for n is denoted $(\frac{D}{n})$.

The LJK symbol $(\frac{D}{n})$ can be computed for $D, n \neq 0$ as follows. First of all, it is completely multiplicative in both arguments:

$$(\frac{D_1 D_2}{n}) = (\frac{D_1}{n})(\frac{D_2}{n}) \qquad (\frac{D}{n_1 n_2}) = (\frac{D}{n_1})(\frac{D}{n_2}) \tag{4.1}$$

Thus it suffices to give its value for $n = -1$ and for n prime:

1. If $n = -1$ then

$$\begin{aligned} (\frac{D}{-1}) &= +1 \qquad D \geq 0 \\ &= -1 \qquad D < 0. \end{aligned} \tag{4.2}$$

2. If $n = 2$ then

$$\begin{aligned} (\frac{D}{2}) &= 0 \qquad && D \text{ even} \\ &= (-1)^{(D^2-1)/8} \qquad && D \text{ odd.} \end{aligned} \tag{4.3}$$

3. Finally, if $n = p$ is an odd prime then $(\frac{D}{p})$ is the Legendre symbol, i.e.

$$\begin{aligned} (\frac{D}{p}) &= 0 \qquad && if \qquad D = 0 \bmod p \\ &= +1 \qquad && if \qquad D = x^2 \bmod p, \text{ for some } x \neq 0 \\ &= -1 \qquad && if \qquad D \neq x^2 \bmod p, \text{ for any } x \end{aligned} \tag{4.4}$$

One can check that if $D = 0, 1 \bmod 4$, (as in this paper), $(\frac{D}{n})$ only depends on the residue class $n \bmod |D|$.

DEFINITION 4.2: FUNDAMENTAL DISCRIMINANTS. $D < 0$ is called a fundamental discriminant if it is the product of relatively prime factors of the form $-4, 8, -8$, or $(-1)^{(p-1)/2}p, p \geq 3$.

Equivalently, $D < 0$ satisfies either (a.) $D = 1 \bmod 4$ and D is squarefree, or (b.) $D = 0 \bmod 4$, $D/4 \neq 1 \bmod 4$, and $D/4$ is squarefree. These are the discriminants of quadratic imaginary fields.

DEFINITION 4.3: L-FUNCTIONS. L-functions of conductor D are by definition the infinite series

$$L(s,D) := \sum_{n=1}^{\infty} \left(\frac{D}{n}\right) n^{-s}. \tag{4.5}$$

For example, choosing $D = 1$ we get the Riemann ζ function. Other examples of $L(s,D)$ are:

$$\begin{aligned} L(s,-3) &= \sum_{n=0}^{\infty} \left(\frac{1}{(3n+1)^s} - \frac{1}{(3n+2)^s} \right) \\ L(s,-4) &= \sum_{n=0}^{\infty} \left(\frac{1}{(4n+1)^s} - \frac{1}{(4n+3)^s} \right) \\ L(s,-7) &= \sum_{n=0}^{\infty} \left(\frac{1}{(7n+1)^s} + \frac{1}{(7n+2)^s} - \frac{1}{(7n+3)^s} + \right. \\ &\qquad \left. + \frac{1}{(7n+4)^s} - \frac{1}{(7n+5)^s} - \frac{1}{(7n+6)^s} \right) \end{aligned} \tag{4.6}$$

The analytic properties of $L(s,D)$ are well known [29]. The series is absolutely convergent for $Re(s) > 1$ and admits an analytic continuation as an entire function of s (for $D \neq 1$). Moreover, for $D < 0$ we may define

$$\xi(s,D) := (q/\pi)^{(s+1)/2} \Gamma((s+1)/2) L(s,D). \tag{4.7}$$

Here q is a positive integer defined to be the minimal period in n of the function $\left(\frac{D}{n}\right)$. In general it is a positive integer dividing D, but in our case where D is a fundamental discriminant, we have $q = |D|$. That is, the character $\left(\frac{D}{n}\right)$ is primitive with period $|D|$ exactly when D is a fundamental discriminant.

It can be shown that $\xi(s,D)$ is an entire function of complex order one. Moreover the zeroes of $\xi(s,D)$ come in complex conjugate pairs. Thus $\xi(s,D)$ has a product formula[8]

$$\xi(s,D) = e^A \prod_{\rho} \left(1 - \frac{s}{\rho}\right). \tag{4.8}$$

The completed function $\xi(s,D)$ has the functional equation

$$\xi(s,D) = \xi(1-s,D), \tag{4.9}$$

which is valid only for primitive characters ($q = |D|$). Otherwise one must make a modification to account for prime factors of $|D|/q$. The zeroes ρ, called the *critical zeroes* of $L(s,D)$, lie in the critical strip $0 \leq Re(s) \leq 1$.

It is known that $s = 1$ is *not* a zero for any D, and in fact

$$L(1,D) = \frac{2\pi h(D)}{w(D)|D|^{1/2}}. \tag{4.10}$$

(We will recall a proof of this below.) Here $w(D)$ is the order of the group of units in K_D; for $D < -4$ we have $w(D) = 2$.

[8] This notation is a little sloppy because one should include a convergence factor for the infinite product. Alternatively, one can eliminate this by grouping the factors with $\Im\rho \neq 0$ into complex-conjugate pairs (since the coefficients of the Dirichlet series are real.)

4.2. The zeroes of $L(s, D)$. The LJK symbol is completely multiplicative as a function of n. Combining this with the prime factorization of integers one can write *another* product formula

$$L(s, D) = \prod_{\text{p prime}} \left(1 - (\frac{D}{p})p^{-s}\right)^{-1} \tag{4.11}$$

for the function $L(s, D)$. This is analogous to Euler's product formula for $\zeta(s)$. While it is not immediately obvious, the product formula (4.11) for $L(s, D)$ encodes the structure of primes in the field K_D. Thus, comparing the two product formulae (4.8) and (4.11) gives information about primes. Accordingly, the nature of the zeroes of $L(s, D)$ is of interest in understanding the arithmetic of K_D.

This astute remark leads to the central problem of analytic number theory, the generalized Riemann hypothesis. It is "generalized" because we are considering the family of functions $L(s, D)$; henceforth we use the abbreviation GRH.[9] A formulation of the GRH, suitable for our present purposes, is the conjecture that the critical zeros of $L(s, D)$ all lie on the line $Re(s) = \frac{1}{2}$. The GRH is of course very difficult, so various sub-problems have been studied. One of these concerns the possible existence of critical zeroes near $s = 1$. These are known as "Landau-Siegel zeroes" (LSZ's). Their existence would falsify the GRH. More precisely:

DEFINITION 4.4: LANDAU-SIEGEL ZEROES. Choose a constant $c > 0$. A "*Landau-Siegel zero for c, D*" is a real zero $s = \beta$ of $L(s, D)$ with

$$1 - \frac{c}{\log |D|} < \beta < 1. \tag{4.12}$$

See [29] (chapter 21) for equivalent formulations of the Landau-Siegel zero.

It is possible to explicitly compute an effective constant $c_2 > 0$ such that the neighborhood of radius $c_2 / \log |D|$ about the point $s = 1$ contains at *most* one zero of $L(s, D)$. This rules out complex zeroes in this neighborhood since they come in conjugate pairs. However, the important challenge remains to eliminate the real zeroes. That is, to find an *effective* constant c_1 such that there are no Landau-Siegel zeroes for c_1 and any D, whether positive or negative. In this paper we only study $D < 0$.

4.3. The large $|D|$ behavior of $L(1, D)$ and the zeroes of $L(s, D)$. In this section we explain how the zeroes of $L(s, D)$ are related to the large $|D|$ behavior of

$$\lambda(D) := \frac{L'(1, D)}{L(1, D)}, \tag{4.13}$$

where the derivative is with respect to s. We will begin with some heuristic remarks, and conclude with some precise estimates (Theorems 4.1 and 4.2). These estimates will be useful in the next section.

The relation between the "size" or rate of growth of $\lambda(D)$ as $|D| \to \infty$ and the existence of zeroes near $s = 1$ follows from the key identity

$$\frac{\xi'}{\xi}(s, D) = \frac{1}{2}\log(|D|/\pi) + \frac{1}{2}\psi((s+1)/2) + \frac{L'}{L}(s, D) = \sum_{\rho|\xi(\rho, D)=0} \frac{1}{s - \rho}. \tag{4.14}$$

[9] In this paper GRH always stands for the hypothesis for Dirichlet L-functions, not for elliptic curve L-functions.

Here

$$\psi(x) := \frac{\Gamma'(x)}{\Gamma(x)} \tag{4.15}$$

is the digamma function. Equation (4.14) follows immediately from logarithmic differentiation in s of the product formula (4.8).

It follows from (4.14) that $\frac{L'}{L}(1,D)$ can grow rapidly with $|D|$ if $L(s,D)$ has some zero ρ which is "close" (as a function of $|D|$) to the point $s=1$. On the other hand, zeroes further away from $s=1$ have less of an impact on the sum in (4.14); their effect is governed by their "density of states." It is only the lower-lying zeroes which are important for the size of $\lambda(D)$. Therefore, if we can say $\lambda(D)$ is "small" then we will have checked a prediction of the GRH, at least in some neighborhood of $s=1$. The smaller the bound on $\lambda(D)$ the larger the zero-free region around $s=1$.

We will now explain a trivial bound on the growth of $\lambda(D)$ obtained from bounding the numerator and denominator separately. Our argument is sloppy and far from optimal, yet it does explain in simple terms the overall reason why the Landau-Siegel zero is related to a small value of $L(1,D)$.

We first bound the numerator by showing there is a constant C such that

$$|L'(1,D)| \le C(\log|D|)^2. \tag{4.16}$$

We prove this as follows. By (4.5)

$$L'(1,D) = -\sum_{n=1}^{|D|}(\frac{D}{n})(\log n)n^{-1} - \sum_{n=|D|+1}^{\infty}(\frac{D}{n})(\log n)n^{-1}. \tag{4.17}$$

The first sum is trivially

$$\mathcal{O}(\sum_{n=1}^{|D|}(\log n)n^{-1}) = \mathcal{O}((\log|D|)^2). \tag{4.18}$$

The second sum can easily be bounded by $\mathcal{O}((\log|D|)^2)$ using partial summation, since $\frac{\log n}{n}$ changes slowly and

$$\sum_{n=k}^{k+|D|}(\frac{D}{n}) = 0, k = 1,2,3,\ldots \tag{4.19}$$

(Roughly speaking, this means there is a lot of cancellation due to the oscillation of the Dirichlet character which modulates $(\log n)n^{-1}$.)

The bound $L'(1,D) = \mathcal{O}((\log|D|)^2)$ can be improved. Nevertheless, the intuition is that the fraction $\lambda(D)$ can be extremely large only if its denominator is extremely small. Therefore we now bound the denominator. A trivial bound comes from (4.10). Since $h(D) \ge 1$ it follows that

$$\frac{1}{L(1,D)} = \frac{\sqrt{|D|}}{\pi h(D)} \le \frac{\sqrt{|D|}}{\pi}. \tag{4.20}$$

Of course, $h(D)$ *does* grow with D so this estimate is far from optimal![10]

[10] Note that if $h(D)/|D|^{1/2}$ is large there cannot be a zero near $s=1$, while if $h(D)/|D|^{1/2}$ is small there might be a nearby zero. This is the essence of the connection to the class number problem. Indeed, the work of Goldfeld and Gross-Zagier provides better bounds on $1/L(1,D)$, reverses the logic

Taken together, (4.16) and (4.20) already show that

$$|\lambda(D)| \leq C'(\log|D|)^2\sqrt{|D|}. \tag{4.21}$$

From (4.14), the inequality (4.21) rules out zeroes at (*roughly*) a distance of order

$$(\log|D|)^{-2}|D|^{-1/2}$$

from $s = 1$. Of course, the sizes of $L'(1, D)$ and $L(1, D)$ are related to each other (by integration of $L'(s, D)$ in s), but it is surprisingly difficult to improve on $|D|^{1/2}$-factor in the trivial bound (4.21).

Let us now see what the GRH has to say about the growth of $\lambda(D)$. First of all, Littlewood proved that the GRH implies $1/L(1, D) = \mathcal{O}(\log\log|D|)$ as $D \to -\infty$ [30]. Moreover, he showed that the GRH implies

$$|\lambda(D)| = \mathcal{O}(\log\log|D|), \tag{4.22}$$

and hence, in particular, that the GRH forces

$$\lim_{D\to-\infty} \frac{\lambda(D)}{\log|D|} = 0. \tag{4.23}$$

In fact, one may produce sequences of D's such that $\lambda(D)$ grows roughly as $\log\log|D|$. All of this is evident from the following

THEOREM 4.1. [Miller – in the appendix to this paper, [31]]. Then

$$\begin{aligned} \limsup_{D\to-\infty} \frac{\lambda(D)}{\log\log|D|} &\geq \frac{1}{2} \\ \liminf_{D\to-\infty} \frac{\lambda(D)}{\log\log|D|} &\leq -\frac{1}{2} \end{aligned} \tag{4.24}$$

where $\lambda(D) = L'(1, D)/L(1, D)$. These inequalities hold, in fact, if $D \to -\infty$ along fundamental discriminants.

Of course, the GRH might actually be false! Evidently, there is a good deal of room between the trivial bound (4.21) and the consequence (4.23) of the GRH. The elimination of Landau-Siegel zeroes is thought to constitute an important beachhead on the way towards understanding the GRH. The following theorem gives a criterion to rule out the existence of Landau-Siegel zeroes:

THEOREM 4.2. Suppose D_n is a sequence of fundamental discriminants with $D_n \to -\infty$. Suppose furthermore that the limit

$$\delta := \overline{\lim}_{D_n\to-\infty} \frac{\lambda(D_n)}{\log|D_n|} \tag{4.25}$$

is finite. Then there exists a positive constant c,

$$c = \left(\frac{1}{2} + \sup\frac{|\lambda(D_n)|}{\log|D_n|}\right)^{-1} \tag{4.26}$$

such that there are no Landau-Siegel zeroes for c among the D_n.

---

of the present discussion, and proves a growth rate of $h(D)$ for $|D|$ increasing. This provides a solution to Gauss' problem of finding which imaginary quadratic fields have a given fixed class number.

Proof. Note that the zeroes ρ are either real or come in complex conjugate pairs. Writing $\rho = \beta + i\gamma$ for the real and imaginary parts we have:

$$\sum_\rho \frac{1}{1-\rho} = \sum_\rho \frac{1-\beta}{(1-\beta)^2+\gamma^2} \tag{4.27}$$

and since the zeroes are in the critical strip, this is a sum of nonnegative terms. Thus, if $\rho_* = \beta_*$ is a real zero of $L(s,D)$ then

$$\begin{aligned} \frac{1}{1-\beta_*} &< \sum \frac{1}{1-\rho} \\ &= \frac{1}{2}\log(|D|/\pi) + \frac{1}{2}\psi(1) + \lambda(D) \end{aligned} \tag{4.28}$$

where in the second line we have used (4.14). Recall that $\psi(x)$ is the digamma function (4.15).

Since (4.27) is positive and the constant $-\frac{1}{2}\log\pi + \frac{1}{2}\psi(1) \cong -0.86$ is negative it follows from (4.28) that

$$-\frac{1}{2} < \frac{\lambda(D)}{\log|D|} \tag{4.29}$$

for any D. Therefore, finiteness of (4.25) implies that

$$\sup \frac{|\lambda(D_n)|}{\log|D_n|} < \infty, \tag{4.30}$$

Now we use

$$\lambda(D_n) \le \left(\sup \frac{|\lambda(D_n)|}{\log|D_n|}\right) \log|D_n|. \tag{4.31}$$

Then (4.28) shows that if $\beta_{*,n}$ is a real zero of $L(s,D_n)$, then

$$\frac{1}{1-\beta_{*,n}} < \left(\frac{1}{2} + \sup \frac{|\lambda(D_n)|}{\log|D_n|}\right) \log|D_n|. \tag{4.32}$$

Now comparing with the definition (4.12) we see there is never a Landau-Siegel zero for (c, D_n) for $c = \left(\frac{1}{2} + \sup\frac{|\lambda(D_n)|}{\log|D_n|}\right)^{-1}$.

An alternative formulation of this proof (which has the advantage of naming the constant c at the sacrifice of ignoring a finite number of cases) is to use the literal definition of the lim sup: for each $\epsilon > 0$ one can find an $N(\epsilon)$ such that

$$\left|\delta - \sup_{m\ge n} \frac{\lambda(D_m)}{\log|D_m|}\right| < \epsilon \tag{4.33}$$

for all $n \ge N(\epsilon)$. Then

$$\frac{1}{1-\beta_{*,n}} < (\frac{1}{2} + \delta + \epsilon)\log|D_n| \tag{4.34}$$

for $n \ge N(\epsilon)$, which leads to the conclusion that there is no Landau-Siegel zero for (c, D_n) for $c = (\frac{1}{2}+\delta+\epsilon)^{-1}$ and $n \ge N(\epsilon)$. □

REMARKS.

1. Put differently, if there are c's $\to 0$ such that each has some Landau-Siegel zero, then those discrimiants $D_n \to -\infty$ satisfy

$$\lim \frac{\lambda(D_n)}{\log|D_n|} = \infty.$$

2. The issue of effectiveness enters in the following ways. In the first argument, one might know δ exactly but not know what $\sup \frac{|\lambda(D_n)|}{\log|D_n|}$ is – we may know the limiting behavior of this sequence but not be able to bound how long it takes for this sequence to exhibit it. In the second argument, we may not be able to compute $N(\epsilon)$ effectively.

5. Black hole entropy and critical zeroes. In this section we will explore come consequences for analytic number theory of the various conjectures of sections 2 and 3.

The main technical observation is that the height conjectures have implications for the behavior of $L(s,D)$ thanks to an equation sometimes called the Chowla-Selberg formula [32]. This is a consequence of class field theory and the Kronecker limit formula. We next recall the derivation of this formula.

The derivation begins with the the nonholomorphic Eisenstein series

$$E(\tau,s) := \sum_{n,m}{}' \frac{y^s}{|m\tau+n|^{2s}} \tag{5.1}$$

where $y = \Im\tau$ and the sum is over all pairs of integers $(n,m) \neq (0,0)$. The sum is absolutely convergent for $Re(s) > 1$, and admits a meromorphic continuation to the entire s-plane. The (first) Kronecker limit formula is the statement (see, e.g., [33], p. 273)

$$E(\tau,s) = \frac{\pi}{s-1} - 2\pi\log\Big[2e^{-\gamma}y^{1/2}|\eta(\tau)|^2\Big] + \mathcal{O}(s-1) \tag{5.2}$$

as $s \to 1$. Now, let τ_k be a set of representatives of the $h(D)$ Heegner points as in section 3. A simple application of class field theory shows that

$$w(D)\big(\frac{|D|}{4}\big)^{s/2}\zeta(s)L(s,D) = \sum_{i=1}^{h(D)} E(\tau_i,s). \tag{5.3}$$

We now take $s \to 1$ and compare the two sides of the equation. Equating the residue of the pole leads to the standard result (4.10). Comparing the constant terms leads to the Chowla-Selberg formula

$$R(D) = 12\Big[\log\sqrt{-D} + \lambda(D) + \log\big[8\pi^2 e^{-\gamma}\big]\Big] \tag{5.4}$$

where $R(D)$ is defined by (3.23)(3.24).

We are now ready to combine the conjectures of sections 2 and 3 with the results from section 4. The first point to make is that the SSV and SH conjectures are "probably" *incompatible*. Let $D_n \to -\infty$ be a sequence of fundamental discriminants. Applying the SSV and SH conjectures to the family of attractors $X_{\gamma_{D_n}}$ constructed in sections 3.2 and 3.3 we find

$$\mathrm{ht}(X_{\gamma_{D_n}}, \widehat{K}_{D_n}) = \kappa\log[\pi\sqrt{-D_n}] + \mathcal{O}\Big(\frac{\log[\sqrt{-D_n}]}{|D_n|^{1/2}}, \frac{1}{|D_n|^{\delta/4}}\Big). \tag{5.5}$$

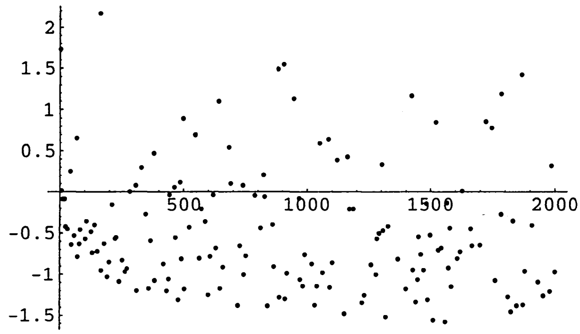

Fig. 1: A plot of $\lambda(D)$ against $|D|$, as D runs through the negative prime fundamental discriminants with $|D| < 2000$. Unfortunately $\log\log|D|$ grows too slowly to numerically see the envelope predicted by Theorem 4.1.

This implies some very interesting cancellation in (3.25) for the following reason. Roughly speaking, Theorem 4.1 says that there exist real numbers $\alpha_\pm$ with $\alpha_- < 0 < \alpha_+$, such that $\lambda(D)$ can actually grow like $\alpha_\pm \log(\log|D|)$ for some sequence of D's [31]. By (5.4) this means that $R(D)$ has scatter as illustrated in fig. 1. For large $|D|$ the envelope is of width $(\alpha_+ - \alpha_-)\log(\log|D|)$. Suppose, for the moment, that the minimal discriminant term in (3.25) were absent. Then using the SH conjecture we would conclude that $\kappa = 3/2$ and moreover, by choosing suitable sequences of discriminants D_n, the charges γ_{D_n} would produce black holes with entropies $S_{\text{micro}}(\gamma_{D_n})$ that actually grow with $|D_n|$ like

$$S_{\text{micro}}(\gamma_{D_n}) \sim \pi\sqrt{|D_n|}(\log|D_n|)^C \tag{5.6}$$

for various constants C in the range $\alpha_- \le C \le \alpha_+$. Moreover, Theorem 4.1 asserts that there exist sequences $\{D_n\}$ which actually realize the extreme cases $\alpha_\pm$.

Now let us restore the minimal discriminant term

$$\frac{\log|N_{\widehat{K}_D/\mathbb{Q}}(\mathcal{D}_{E_\gamma/\widehat{K}_D})|}{h(D)} \tag{5.7}$$

in (3.25). As we have discussed, this depends on the choice of Heegner point and the Weierstrass model. It would be fascinating (though we think unlikely) if for general sequences of charges γ_{D_n} with $D_n \to -\infty$ one could systematically choose Weierstrass models such that the minimal discriminant in (3.25) fluctuates to match the changes in $\lambda(D_n)$ in the way demanded by the SSV conjecture. Since we think it is highly unlikely, we conclude that SSV and SH are probably incompatible, as stated above. We could be more precise if we had better information on the size of the minimal discriminant $\mathcal{D}_{E_\gamma/\widehat{K}_D}$.

Put differently, if one could choose families of attractor points with a bounded value of (5.7) then the strong height conjecture implies that the distribution of

$$\frac{\log\dim\mathcal{H}(\gamma)}{\pi\sqrt{I_4(\gamma)}} \tag{5.8}$$

plotted against $\sqrt{I_4(\gamma)}$ would have a lot of scatter, similar to the scatter in fig. 1, around the average value 1.

Let us now assume WSV and $\overline{WH}$ (or WH). Then, from (2.8)(3.2) and the first inequality in (3.26) we get

$$\overline{\lim_{n}} \frac{R(D_n)}{\log |D_n|^{1/2}} \leq 8\bar{\kappa}. \tag{5.9}$$

Then, from the Chowla-Selberg formula we get

$$\overline{\lim_{n}} \frac{\lambda(D_n)}{\log |D_n|^{1/2}} \leq \frac{8\bar{\kappa} - 12}{12}. \tag{5.10}$$

Now, using Theorem 4.2 we arrive at the main statement in this paper: The WSV conjecture and WH conjecture together rule out Landau-Siegel zeroes for $D < 0$. More precisely, we have

THEOREM 5.1. *The WSV conjecture and the $\overline{WH}$ conjecture together imply the following restrictions on Landau-Siegel zeroes: For any sequence $\{D_n\}$ of fundamental discriminants with $D_n \to -\infty$, there exists a positive constant c such that there are no Landau-Seigel zeroes for (c, D_n).*

We would like to remark that a similar result (with $\log |D|$ replaced by $|D|^{\epsilon}$, for a positive constant ϵ) was proven unconditionally by Tatuzawa in [34]. Tatuzawa's theorem, while weaker, is sufficient for many applications. Two novelties of the present discussion are that we have $\log |D|$ rather than a power, and, moreover, the constant c is in principle computable from physics (granted the WSV and $\overline{WH}$ conjectures).

Let us now turn the logic around and suppose that the correct configuration of conjectures is WSV, WH, but that there are Landau-Siegel zeroes for $D < 0$ (thus falsifying the GRH). Then there would be sequences of charges γ_n with D-brane configurations of anomalously large entropy compared to the Bekenstein-Hawking entropy.

Even more ambitiously, we can turn things around and assume the "safest" set of conjectures: GRH, WSV, and WH, and see what they predict. From the GRH we have:

$$\lim_{n \to \infty} \frac{R(D_n)}{\log |D_n|^{1/2}} = 12 \tag{5.11}$$

Now, the WSV and the $\overline{WH}$ conjecture imply information on the mysterious minimal discriminants of the elliptic curves (3.21). From (3.26) we get

$$\overline{\lim_{n}} \frac{\log |N_{\widehat{K}_{D_n}/\mathbb{Q}}(\mathcal{D}_{E_{\gamma_n}/\widehat{K}_{D_n}})|}{h(D_n) \log(|D_n|^{1/2})} = 8\bar{\kappa} - 12, \tag{5.12}$$

and in particular, $\bar{\kappa} \geq 3/2$. Note that the second inequality in (3.26) shows that $\bar{\kappa} \leq 9$. The second inequality depends on the choice of Weierstrass model (3.21), so we are assuming our choice of models is suitable for the conjecture. Of course, replacing $\overline{WH}$ by WH we replace $\bar{\kappa}$ by κ and $\overline{\lim}$ with lim and get a stronger "prediction."

Curiously, some examples support (5.12). As we mentioned before, the minimal discriminant was computed in [26] for a certain family of curves over $\widehat{K}_{D=-p}$ for $p = 3 \bmod 4$, and prime. Denoting these curves by $A(p)$, Theorem 12.2.1 of [26] shows that $\mathcal{D}_{A(p)/\widehat{K}_{D=-p}} = (-p^3)$, so the norm over $\mathbb{Q}$ is $(-p^3)^{2h}$ and we thus get $\bar{\kappa} = 3$ for this family. On the other hand, there are also families of CM curves with $N_{\widehat{K}_{D_n}/\mathbb{Q}}(\mathcal{D}) = 1$ [25]. Such a family would have $\bar{\kappa} = 3/2$.

6. Conclusion. The remarks of section 5 are built on a house of cards, namely, on a chain of conjectures about some relations between D-branes, black holes, and number theory. The weakest link by far in the chain of conjectures is the relation between D-branes and arithmetic suggested in [1]. Admittedly, it is a long shot. Nevertheless, as we have shown, it would have dramatic consequences if true. At worst, there are a couple of interesting coincidences, so perhaps it deserves some closer scrutiny. It is curious that the choice of arithmetic model has some relevance for "predictions" such as (5.12). Whether this turns out to be an interesting feature or a fatal flaw of our discussion remains to be seen.

ACKNOWLEDGEMENTS. We would like to thank P. Candelas, E. Diaconescu, D. Goldfeld, R. Ramakrishna, D. Rohrlich, P. Sarnak, J. Silverman, and S. Zhang for helpful discussions and correspondence. We are particularly grateful to B. Gross for important critical remarks on the draft. GM is supported by DOE grant DE-FG02-92ER40704. SM is supported by an NSF postdoctoral fellowship and by a Yale Hellman fellowship. GM also thanks the Institute for Advanced Study for hospitality and the Monell foundation for support during the completion of this work; SM is similarly indebted to Harvard University for their hospitality.

Appendix A. Large Values of $\frac{L'}{L}(1,-p)$, by Stephen D. Miller. In the above paper, we made use of the fact that the generalized Riemann hypothesis (GRH) implies that

$$\lambda(D) = O(\log\log|D|),$$

where

$$\lambda(D) = \frac{L'(1,D)}{L(1,D)}.$$

Also, we discussed the implications of how the $\log\log|D|$ rate is in fact optimal. The upper bound is well-known; a proof can be found in [35], for example. The purpose of this appendix is to prove the lower bound:

THEOREM A(=4.1 ABOVE). *Let $\lambda(D) = \frac{L'(1,D)}{L(1,D)}$ and $D = -p$. As $p \to \infty$ among the primes which are 3 modulo 4,*

$$\limsup \frac{\lambda(D)}{\log\log|D|} \geq \frac{1}{2}$$

and

$$\liminf \frac{\lambda(D)}{\log\log|D|} \leq -\frac{1}{2}.$$

The constant $\frac{1}{2}$ is far from optimal; any non-zero constant will do for the application in section 5 above. We remark that the discriminants $D = -p$ are a very special kind; for one thing, they are fundamental discriminants. Therefore, the conclusion of the theorem still holds if we weaken our conditions on D, e.g. if it passes to $-\infty$ over all discriminants.

One can prove this theorem under the assumption of GRH by modifying the technique of Littlewood ([30]), who was the first to prove the analogous result for $L(1,D)$. We will present a different proof here based on the methods of [36] and [37], which has the advantage that it is unconditional and that it allows us to infer the theorem over a set as sparse as the primes. Shorter proofs are possible, but we present the one here because all of the background material is contained in the standard reference [29]. Extremal theorems of this type are proved in two steps: first one truncates the dirichlet series definition for each individual $L(s, D)$, and then averages the remaining finite sum over many D to show the theorem in the mean. Since one term must always meet the size of the average, we conclude that there are large individual values. We will use sieve methods for each of these steps.

A.1. Character sums. The material in this section is modified from [37], chapter 22.

LEMMA 1. [Lemma 22.4 of [37]] Let $\chi_1, \dots, \chi_J$ be distinct primitive characters to moduli $\leq Q$, and let a_n be arbitrary complex numbers. Then

$$\sum_{j=1}^{J} \left| \sum_{n=1}^{N} a_n \chi_j(n) \right|^2 \leq (N + J c_0 Q \log Q) \sum_{n=1}^{N} |a_n|^2$$

for some absolute constant $c_0 > 0$.

We will use Lemma 1 to make a series of estimates. The exponents we use are somewhat arbitrary and are hardly optimal.

PROPOSITION 2. Let Q be large and $(\log Q)^{20} \leq U \leq Q^2$, $b_p \in \mathbb{C}$, $|b_p| \leq 1$ for prime p. Then at most $O(Q^{3/4})$ of the distinct primitive characters χ to moduli $\leq Q$ violate

$$\left| \sum_{U<p<2U} \frac{b_p \log(p) \chi(p)}{p} \right| \leq U^{-1/10}. \tag{A.1}$$

Proof. Let $\chi_1, \dots, \chi_J$ be these violating characters. Let a_n be defined by the expansion

$$\sum_{n=1}^{N} a_n \chi(n) = \left(\sum_{U<p<2U} \frac{b_p \log(p) \chi(p)}{p} \right)^m, \quad N = (2U)^m.$$

Then

$$a_{p_1 \cdots p_m} = \frac{b_{p_1} \cdots b_{p_m}}{p_1 \cdots p_m} (\log p_1) \cdots (\log p_m) \mu(p_1, \dots, p_m) \leq \frac{m!}{U^m} (\log 2U)^m,$$

where $\mu(p_1, \dots, p_m) \leq m!$ is the multiplicity from the expansion. Also,

$$\sum_{U<p<2U} \left| \frac{b_p \log p}{p} \right| \leq \sum_{U<p<2U} \frac{\log p}{p} = \log(2U/U) + o(1) \leq 1$$

for large Q. Consequently

$$\sum_{n=1}^{N} |a_n| \leq \sum_{U<p_1, \dots, p_m<2U} \frac{|b_{p_1}| \cdots |b_{p_m}| (\log p_1) \cdots (\log p_m)}{p_1 \cdots p_m} \mu(p_1, \dots, p_m)$$

$$= \left(\sum_{U<p<2U} \left| \frac{b_p \log p}{p} \right| \right)^m \leq 1,$$

and so

$$\sum_{n=1}^{N} |a_n|^2 \leq (\max |a_n|) \sum_{n=1}^{N} |a_n| \leq \frac{m!}{U^m} (\log 2U)^m.$$

By Lemma 1 applied to the exceptional characters $\chi_1, \ldots, \chi_J$,

$$JU^{-m/5} \leq ((2U)^m + Jc_oQ \log Q)\, m! \left(\frac{\log 2U}{U}\right)^m.$$

This estimate will be more than sufficient for our purposes, and in fact we will use $m! \leq m^m$ to deduce the weaker

$$JU^{-m/5} \leq (2m \log 2U)^m + Jc_0Q \log Q \left(\frac{m \log 2U}{U}\right)^m. \tag{A.2}$$

Now we shall set $m = \left[\frac{11}{8} \frac{\log Q}{\log U}\right] + 1$, where $[\cdot]$ denotes the greatest-integer function. In Lemma 3 we show that

$$U^{-m/5} > 2c_0Q \log Q \left(\frac{m \log 2U}{U}\right)^m$$

holds for large Q. Granted this, (A.2) now implies

$$JU^{-m/5} \leq 2 \cdot (2m \log 2U)^m$$

for large Q. We claim that

$$J \leq 2 \left(2mU^{1/5} \log 2U\right)^m = O(Q^{3/4}).$$

If this were not so, then for any $C > 0$ we would have that

$$m \log(2mU^{1/5} \log 2U) \geq \frac{3}{4} \log Q + C.$$

Adjusting C,

$$m \log 2 + \frac{1}{5} m \log U + \log(m \log U) \geq \frac{3}{4} \log Q + C$$

$$m \log 2 + \left(\frac{11}{40} \log Q + \frac{1}{5} \log U\right) + \log\left(\frac{11}{8} \log Q + \log U\right) \geq \frac{3}{4} \log Q + C$$

$$m \log 2 + \frac{27}{40} \log Q + \log(\frac{27}{8} \log Q) \geq \frac{3}{4} \log Q + C.$$

Readjusting C again slightly, we find

$$\frac{11}{8} \frac{\log Q}{\log U} \log 2 + \log \log Q \geq \frac{3}{40} \log Q + C$$

and finally

$$\frac{11}{8} \frac{\log Q \log 2}{20 \log \log Q} + \log \log Q \geq \frac{3}{40} \log Q + C,$$

which is absurd. □

LEMMA 3. Keeping the notation of the above proof, let C be any fixed constant. Then for Q sufficiently large,

$$U^{-m/5} > CQ \log Q \left(\frac{m \log 2U}{U}\right)^m.$$

Proof. Otherwise

$$\frac{4m}{5}\log U \leq \log C + \log Q + \log\log Q + m\log(m\log 2U),$$

which implies

$$\frac{11}{8}\frac{4}{5}\log Q \leq (1+\delta)\log Q + m\log(m\log 2U)$$

for some $\delta > 0$ and Q large. Continuing

$$(\frac{1}{10} - \delta)\log Q \leq m\log(m\log 2U)$$

$$\leq \left(\frac{11}{8}\frac{\log Q}{\log U} + 1\right)\log\left(\left(\frac{11}{8}\frac{\log Q}{\log U} + 1\right)(\log U + 2)\right)$$

$$\leq \left(\frac{11}{8}\frac{\log Q}{\log U} + 1\right)\log\left(\frac{11}{8}\log Q + \log U + \frac{11}{4}\frac{\log Q}{\log U} + 2\right).$$

Using $(\log Q)^{20} \leq U \leq Q^2$, introducing a tiny $\delta' > 0$, and taking Q larger yet, we get

$$(\frac{1}{10} - \delta)\log Q \leq \left(\frac{11}{8}\frac{(1+\delta')\log Q}{20\log\log Q}\right)\log\left(\frac{11}{4\cdot 20}\frac{\log Q}{\log\log Q} + \frac{27}{8}\log Q + 2\right)$$

$$\leq (1+\delta')\frac{11}{160}\frac{\log Q}{\log\log Q}\log(7/2\log Q).$$

This is a contradiction for large Q, because $\frac{1}{10} > \frac{11}{160}$. □

PROPOSITION 4. With the same notation as in Proposition 2 and Lemma 3, but instead requiring that $U > Q^2$, at most $O(Q^{2/3}\log U)$ distinct primitive characters to moduli $\leq Q$ have

$$\left|\sum_{U<p<2U}\frac{b_p\log(p)\chi(p)}{p}\right| \geq Q^{-1/3}.$$

Proof. Again, let $\chi_1, \dots, \chi_J$ be these characters. By Lemma 1,

$$JQ^{-2/3} \leq (U + Jc_0 Q\log Q)\sum_{U<p<2U}\left|\frac{b_p\log p}{p}\right|^2.$$

Since $|b_p| \leq 1$, the prime number theorem guarantees that

$$\sum_{U<p<2U}\left|\frac{b_p}{p}\log p\right|^2 \leq \sum_{U<p<2U}\frac{(\log p)^2}{p^2} = O\left(\int_U^{2U}\frac{(\log\xi)^2}{\xi^2}\frac{d\xi}{\log\xi}\right) = O\left(\frac{\log U}{U}\right).$$

Therefore

$$JQ^{-2/3} \leq C(U + Jc_0 Q\log Q)\frac{\log U}{U} = C_1\log U + C_2 J\frac{Q\log Q}{U}\log U,$$

or

$$J\left(1 - C_2 U^{-1}Q^{5/3}\log Q\log U\right) \leq C_1 Q^{2/3}\log U.$$

Since $U > Q^2$,

$$\frac{Q^{5/3}}{U}\log Q\log U < 2Q^{-1/3}(\log Q)^2$$

becomes arbitrarily small for large Q, and we conclude

$$J = O(Q^{2/3} \log U).$$

□

We will require a form of Dirichlet's theorem on the distribution of primes in residue classes. To fix notation we will set $li(x)$ to be the logarithmic integral

$$li(x) = \int_2^\infty \frac{dt}{\log t}$$

and

$$\pi(x,k,l) = \#\{p \leq x | p \; prime, p \equiv l \; (mod \; k)\}.$$

See [29] for more details on this topic. Siegel and Walfisz proved

THEOREM SW. Let $\epsilon > 0$ be fixed. Then there exists a constant c depending on ϵ such that

$$\pi(x,k,l) = \frac{li(x)}{\phi(q)} + O(x \exp(-c\sqrt{x})) \;,\; (k,l) = 1$$

for $x \geq \exp k^\epsilon$.

Page proved a result which is also valid for smaller x, but with a possible exception (related to the Landau-Siegel zero):

THEOREM P. Let Q be a positive integer. Then there exists a constant $b > 0$ such that for any modulus $k \leq Q$ – except for perhaps one exceptional modulus k_1 and its multiples – we have

$$\pi(x,k,l) = \frac{li(x)}{\phi(q)} + O(x \exp(-b\sqrt{\log x})) \;,\; (k,l) = 1$$

for $x \geq \exp{(\log Q)^2} \Leftrightarrow Q \leq \exp(\sqrt{\log x})$. The possible exception k_1 grows to infinity with Q.

Recall the formula that

$$\frac{L'(1,\chi)}{L(1,\chi)} = - \sum_{p \; prime} \frac{\chi(p) \log p}{p} + O(1).$$

The next proposition shows that one can truncate the sum.

PROPOSITION 5. For Q large and $(\log Q)^{20} \leq y \leq Q^2$, all but $O(Q^{11/12})$ distinct primitive characters χ to moduli $\leq Q$ have

$$\frac{L'(1,\chi)}{L(1,\chi)} = - \sum_{p \leq y} \frac{\chi(p) \log p}{p} + O(1).$$

Proof. Another form of the Siegel-Walfisz theorem gives us that

$$\frac{L'(1,\chi)}{L(1,\chi)} = - \sum_{p \leq \exp Q^{1/8}} \frac{\chi(p) \log p}{p} + O(1).$$

Given y between $(\log Q)^{20}$ and Q^2, we shall dyadically decompose the range from $(y, \exp Q^{1/8}]$ into intervals $(U_k, 2U_k]$, $U_k = 2^k y \leq \exp Q^{1/8}$. Among these $O(Q^{1/8})$ ranges we will distinguish those that come before and after the range containing Q^2, designated as $(U_{k_0}, 2U_{k_0}]$.

First, for $k \le k_0$

$$\left|\sum_{y<p<2^{k_0+1}y} \frac{\chi(p)\log(p)}{p}\right| \le \sum_{k=0}^{k_0}\left|\sum_{U_k<p<2U_k} \frac{\chi(p)\log(p)}{p}\right|$$

$$\le \sum_{k=0}^{k_0}(2^k y)^{-1/10} = O(y^{-1/10}),$$

for characters which satisfy the bound of Proposition 2 in each range. There are at most $O(Q^{3/4}(k_0+1)) = O(Q^{7/8})$ exceptions.

Secondly, for the $k > k_0$

$$\left|\sum_{2^{k_0+1}<p\le\exp Q^{1/8}} \frac{\chi(p)\log p}{p}\right| \le Q^{-1/3} \sum_{2^k y\le\exp Q^{1/8}} 1$$

$$= O(Q^{-1/3+1/8}) = O(Q^{-5/24}) = O(1)$$

with the exception of

$$O(Q^{2/3} \sum_{2^k y\le\exp Q^{1/8}} \log(2^{k+1}y)) = O(Q^{2/3}(Q^{1/8})^2) = O(Q^{11/12})$$

characters. □

A.2. Constructing a set of discriminants. Having proven Proposition 5, we have now established that if D is a fundamental discriminant, then

$$\frac{L'(1,D)}{L(1,D)} = \sum_{p\le(\log|D|)^{20}} \left(\frac{D}{p}\right)\frac{\log p}{p} + O(1)$$

except for D in a very sparse set. In fact, we may usually invoke this conclusion for fundamental discrimants of the form $D = -p$, p a prime with $p \equiv 3\ (4)$. We shall now present a set of candidates for extreme values of $L(1,D)$. While we will not be able to directly show that they are large, we will prove an average result which allows us to conclude that at least one of them is.

The candidates: We will pattern ourselves after the argument in [36]. Let x be a large parameter and set

$$y = \frac{\sqrt{\log x}}{(\log\log x)^2}.$$

Enumerate the odd primes $p_2, p_3, \dots, p_m \le y$ and form the product

$$M = 8p_2p_3\cdots p_m = 4\exp(\sum_{p\le y}\log p) = \exp(y + o(y)).$$

Not all of the moduli

$$\frac{M}{p_m}, \dots, \frac{M}{p_1}$$

are divisible by the exceptional modulus k_1 mentioned in Theorem P (applied to x). Let $k = \frac{M}{p_r}$ be the first of these listed above (i.e. with r the largest) which is not divisible by k_1; if none exists, then set $r = m$. The purpose of this construction is to guarantee that

$$\pi(x,k,l) = \frac{li(x)}{\phi(k)} + O(x\exp(-b\sqrt{\log x}))$$

for l relatively prime to k. Using the Chinese remainder theorem, we can find two special residues $l_\pm$ *mod* k such that

$$\left(\frac{l_\pm}{p}\right) = \pm 1\ ,\ p_r \neq p \leq y.$$

Define the sets

$$\mathcal{S}_\pm = \{-q \mid \sqrt{x} \leq q \leq x, q \ prime, -q \equiv l_\pm \pmod{k}, (q,k_1) = 1\} - E_\pm,$$

where $E_\pm$ is the exceptional set of Proposition 5. It follows that the cardinality of $\mathcal{S}_\pm$ is $S_\pm = \frac{li(x)}{\phi(k)} + O(x\exp(-b\sqrt{\log x}))$. For any $D \in \mathcal{S}_\pm$,

$$\frac{L'(1,D)}{L(1,D)} = \mp \sum_{p_r \neq p \leq y} \frac{\log p}{p} - \sum_{y<p\leq(\log x)^{20}} \left(\frac{D}{p}\right)\frac{\log p}{p} + O(1)$$

$$\frac{L'(1,D)}{L(1,D)} = \mp\frac{1}{2}\log\log x + o(\log\log x) - \sum_{y<p\leq(\log x)^{20}} \left(\frac{D}{p}\right)\frac{\log p}{p} + O(1). \tag{A.3}$$

In the next section, we will show that this last sum is $o(1)$ on average.

A.3. Sifting and averaging. Following [36], we will use a version of the Rényi sieve to average (A.3).

LEMMA R. Let $\mathcal{Z}$ be a set of $Z = |\mathcal{Z}|$ integers in a range of N consecutive integers, $f(p) \leq p$ and $Q(p) > 1$ otherwise-arbitrary functions, and set

$$\tau = \min_{p\leq\sqrt{N}} \frac{f(p)}{p}\ ,\ Q = \max_{p\leq\sqrt{N}} Q(p).$$

Let

$$Z(p,h) = \#\{z \in \mathcal{Z} \mid z \equiv\ h(p)\}.$$

Then for all but at most $\frac{2NQ^2}{Z\tau}$ "abnormal" primes $\leq \sqrt{N}$ we have that

$$\left|Z(p,h) - \frac{Z}{p}\right| < \frac{Z}{pQ(p)}$$

holds except for at most $f(p)$ "irregular" residues.

Proof of Lemma R: This follows from Chebyshev's inequality applied to the sieve inequality

$$\sum_{p\leq Q}\sum_{h=1}^{p}\left(Z(p,h) - \frac{Z}{p}\right)^2 \leq (N+Q^2)Z,$$

([29], p. 158). □

LEMMA 6. ([36]) Keeping the same notation as Lemma R, assume p is a normal prime. Then the number of $z \in \mathcal{Z}$ in the irregular residue classes *mod* p is $\leq Z\left(\frac{f(p)}{p} + \frac{1}{Q(p)}\right)$.

Proof of Lemma 6: The number in regular classes exceeds

$$(p-f(p))\left(\frac{Z}{p}-\frac{Z}{pQ(p)}\right) \geq Z - Z\frac{f(p)}{p} - \frac{Z}{Q(p)}.$$

□

We will always take $f(p) = p/(\log p)^5$ and $Q(p) = (\log p)^5$.
The average of the sum in (A.3) is

$$\text{(A.4)} \qquad \sum_{D\in\mathcal{S}_\pm}\ \sum_{y<p\leq(\log x)^{20}} \left(\frac{D}{p}\right)\frac{\log p}{p} = \sum_{y<p\leq(\log x)^{20}}\ \sum_{h=1}^{p}\left(\frac{h}{p}\right)\frac{\log p}{p}\mathcal{S}_\pm(p,h).$$

LEMMA 7. This sum, restricted to the normal primes p, is

$$\sum_{y<p\leq(\log x)^{20} normal}\ \sum_{h=1}^{p}\left(\frac{h}{p}\right)\frac{\log p}{p}\mathcal{S}_\pm(p,h) = o(S_\pm).$$

Proof of Lemma 7: For normal p,

$$\left|\sum_{h=1}^{p}\left(\frac{h}{p}\right)\mathcal{S}_\pm(p,h)\right| = \left|\sum_{h=1}^{p}\left(\frac{h}{p}\right)\left(\mathcal{S}_\pm(p,h)-\frac{S_\pm}{p}\right)\right|$$

$$\leq \sum_{h\ regular}\frac{S_\pm}{pQ(p)} + \sum_{h\ irreg.}\mathcal{S}_\pm(p,h) + \sum_{h\ irreg.}\frac{S_\pm}{p}$$

$$\leq \frac{S_\pm}{Q(p)} + S_\pm\frac{f(p)}{p} + \left(\frac{S_\pm}{Q(p)} + S_\pm\frac{f(p)}{p}\right) \leq \frac{4S_\pm}{(\log p)^5}.$$

The sum over the normal primes p is thus

$$\leq \sum_{p>y}\frac{4S_\pm}{p(\log p)^4} = o(S_\pm).$$

□

LEMMA 8. The sum in (A.4) over the abnormal primes is also $o(S_\pm)$.

Proof of Lemma 8: First, we demonstrate that there is at most one abnormal prime $p \leq \exp(y\log\log x)$. This is because $\mathcal{S}_\pm(p,h) = \#\{D \in \mathcal{S}_\pm \mid D \equiv h\ (p)\}$. By Theorem P applied to x and the moduli $kp \leq \exp(\sqrt{\log x})$, there is at most one prime $p \leq \exp(y\log\log x)$ which violates

$$\mathcal{S}_\pm(p,h) = \frac{S_\pm}{p-1} + O(x\exp(-b\sqrt{\log x})) = \frac{S_\pm}{p} + O(\frac{S_\pm}{p^2}) + O(x\exp(-b\sqrt{\log x})).$$

This means p qualifies as a normal prime, because $x \sim S_\pm\phi(k)\log x$.

By Lemma R, the number of abnormal primes $p \leq \sqrt{x}$ is bounded by $\frac{2x(\log x)^{15}}{S_\pm} = O(\phi(k)(\log x)^{16})$. We bound the sum over the abnormal primes p individually as

$$\left|\sum_{D\in\mathcal{S}_\pm}\ \sum_{y<p\leq(\log x)^{20}\ abnormal}\left(\frac{D}{p}\right)\frac{\log p}{p}\right| \leq S_\pm \sum_{y<p\leq(\log x)^{20}\ abnormal}\frac{\log p}{p}$$

$$\leq S_{\pm}\left(\frac{\log y}{y}+O\left(\phi(k)(\log x)^{16}\frac{y\log\log x}{\exp(y\log\log x)}\right)\right)=o(S_{\pm}).$$

□

Proof of Theorem A: Lemmas 7 and 8 show that the average of (A.3) over $\mathcal{S}_{\pm}$ is

$$\frac{1}{S_{\pm}}\sum_{D\in\mathcal{S}_{\pm}}\frac{L'(1,D)}{L(1,D)}=\pm\frac{1}{2}\log\log x+o(\log\log x).$$

At least one term must exceed this average, proving Theorem A. □

REFERENCES

[1] G. MOORE, *Arithmetic and Attractors*, hep-th/9807087; *Attractors and Arithmetic*, hep-th/9807056.
[2] A. STROMINGER AND C. VAFA, *Microscopic origin of the Bekenstein-Hawking entropy*, hep-th/9601029; Phys. Lett., B379 (1996), pp. 99–104.
[3] S. FERRARA, R. KALLOSH, AND A. STROMINGER, *N=2 extremal black holes*, hep-th/9508072.
[4] A. STROMINGER, *Macroscopic entropy of N=2 extremal black holes*, hep-th/9602111.
[5] S. FERRARA AND R. KALLOSH, *Universality of sypersymmetric attractors*, hep-th/9603090; *Supersymmetry and attractors*, hep-th/9602136; S. FERRARA, *Bertotti-Robinson geometry and supersymmetry*, hep-th/9701163.
[6] K. BEHRNDT, G. LOPES CARDOSO, B. DE WIT, R. KALLOSH, D. LÜST, AND T. MOHAUPT, *Classical and quantum N=2 supersymmetric black holes*, hep-th/9610105.
[7] J. MALDACENA, A. STROMINGER, AND E. WITTEN, *Black hole entropy in M-theory*, hep-th/9711053.
[8] BERNARD DE WIT, GABRIEL LOPES CARDOSO, DIETER LÜST, THOMAS MOHAUPT, AND SOO-JONG REY, *Higher-order gravitational couplings and modular forms in $N=2, D=4$ heterotic string compactifications*, hep-th/9607184; G. L. CARDOSO, B. DE WIT, AND T. MOHAUPT, *Corrections to macroscopic supersymmetric black hole entropy*, hep-th/9812082.
[9] R. DIJKGRAAF, E. VERLINDE, AND H. VERLINDE, *Counting dyons in N=4 string theory*, hep-th/9607026; Nucl. Phys., B484 (1997), pp. 543–561.
[10] J. SILVERMAN, *Arithmetic Geometry*, G. Cornell and J. H. Silverman, ed., Springer Verlag, 1986.
[11] E. WITTEN, *Phases of $N=2$ theories in two dimensions*, hep-th/9301042; Nucl. Phys., B403 (1993), pp. 159–222.
[12] R. KALLOSH AND B. KOL, *E(7) symmetric area of the black hole horizon*, hep-th/9602014.
[13] L. ANDRIANOPOLI, R. D'AURIA, S. FERRARA, P. FRE', AND M. TRIGIANTE, *$E_{7(7)}$ duality, BPS black-hole evolution and fixed scalars*, hep-th/9707087; Nucl. Phys., B509 (1998), pp. 463–518.
[14] S. FERRARA AND J. MALDACENA, *Branes, central charges and U-duality invariant BPS conditions*, hep-th/9706097.
[15] D. A. COX, *Primes of the Form x^2+ny^2*, John Wiley, 1989.
[16] J. W. S. CASSELS, *Lectures on Elliptic Curves*, Cambridge University Press, 1995.
[17] D. HUSEMOLLER, *Elliptic Curves*, Springer-Verlag, GTM 111.
[18] J. SILVERMAN, *The Arithmetic of Elliptic Curves*, Springer Verlag, GTM 106, 1986.
[19] G. FALTINGS, *Calculus on arithmetic surfaces*, Ann. Math., 119 (1984), p. 387–424.
[20] S. LANG, *Introduction to Arakelov Geometry*, Springer-Verlag, 1988.
[21] S. BLOCH, *The proof of the Mordell conjecture*, Math. Intell., 6 (1984), p. 41–47.
[22] B. MAZUR, *Arithmetic on curves*, Bull. Amer. Math. Soc., 14(1986), p. 207–259.
[23] L. ALVAREZ-GAUMÉ, J.-B. BOST, G. MOORE, P. NELSON, AND C. VAFA, *Bosonization on higher genus Riemann surfaces*, Commun. Math. Phys., 112 (1987), p. 503–552.
[24] J. E. CREMONA, *Algorithms for Modular Elliptic Curves*, Cambridge University Press, 1992.
[25] D. ROHRLICH, *Elliptic curves with good reduction everywhere*, J. London Math. Soc., 25 (1982), p. 216–222.
[26] B. GROSS, *Arithmetic on Elliptic Curves with Complex Multiplication*, Springer-Verlag, LNM 776.

[27] B. GROSS AND D. ZAGIER, *On singular moduli*, J. reine angew. Math., 355 (1985), p. 191–220.
[28] W. DUKE, *Hyperbolic distribution problems and half-integral weight Maass forms*, Invent. Math., 92 (1988), p. 73–90.
[29] H. DAVENPORT, *Multiplicative Number Theory*, Second edition, Springer-Verlag, GTM 74.
[30] J. E. LITTLEWOOD, *On the class number of the corpus* $P(\sqrt{-k})$, Proc. Lond. Math. Soc., Ser. 2, 27 (1928), Part 5.
[31] S. D. MILLER, *Large values of* $\frac{L'}{L}(1,-p)$, appendix to this paper.
[32] A. WEIL, *Elliptic Functions According to Eisenstein and Kronecker*, Springer-Verlag, 1976.
[33] S. LANG, *Elliptic Functions*, Addison-Wesley, 1973.
[34] T. TATUZAWA, *On a theorem of Siegel*, Jap. J. Math., 21 (1951), p. 163–178.
[35] A. GRANVILLE AND H. STARK, *ABC implies no 'Siegel zeroes'*, preprint.
[36] P. T. BATEMAN, S. CHOWLA, AND P. ERDÖS, *Remarks on the size of* $L(1,\chi)$, Publ. Math. Debrecen, 1 (1950), pp. 165–182.
[37] P. D. T. A. ELLIOT, *Probablistic Number Theory II: Central Limit Theorems*, Grundlehren der mathematischen Wissenschaften 240, Springer, New York, 1980.