Invent. math. 139, 509–523 (2000)
Digital Object Identifier (DOI) 10.1007/s002229900036

Inventiones mathematicae

ABC implies no "Siegel zeros" for L-functions of characters with negative discriminant

Andrew Granville[1,⋆], **H.M. Stark**[2]

[1] Department of Mathematics, University of Georgia, Athens, GA 30602, USA
(e-mail: andrew@math.uga.edu)
[2] Department of Mathematics, University of California, San Diego, CA, USA
(e-mail: hmstark@ucsd.edu)

Oblatum: 15-IX-1998 & 30-VII-1999 / Published online: 29 November 1999

1. Introduction

Oesterlé and Masser's *abc-conjecture* asserts that for any given $\varepsilon > 0$, if a, b and c are coprime positive integers satisfying $a + b = c$ then $c \ll_\varepsilon N(a, b, c)^{1+\varepsilon}$, where $N(a, b, c)$ is the product of the distinct primes dividing abc. Their conjecture has a wide variety of interesting, sometimes surprising, consequences (such as Fermat's Last Theorem, other than perhaps finitely many examples). Vojta [14, page 84] showed how to formulate the *abc*-conjecture in arbitrary number fields (from which Elkies [5] elegantly deduced Faltings's Theorem). We will describe a version of this conjecture after introducing the basic notation.

Given a number field K we define $\Delta_K := |D_K|^{1/[K:\mathbb{Q}]}$, where D_K is the discriminant for the field extension $K/\mathbb{Q}$. For non-zero numbers $a_1, a_2, \dots, a_n \in K$ we define the (absolute) *height* and the *conductor* of $\{a_1, a_2, \dots, a_n\}$ to be

$$H(a_1, a_2, \dots, a_n) = \prod_v \max\left(\|a_1\|_v, \|a_2\|_v, \dots, \|a_n\|_v\right) \quad \text{and}$$

$$N(a_1, a_2, \dots, a_n) = N_K(a_1, a_2, \dots, a_n) = \prod_{\mathfrak{p}\in I} \|\mathfrak{p}\|_{\mathfrak{p}}^{-1},$$

respectively, where v ranges over all the normalized valuations of K, and I is the set of prime ideals $\mathfrak{p}$ of K for which $\|a_1\|_{\mathfrak{p}}, \|a_2\|_{\mathfrak{p}}, \dots, \|a_n\|_{\mathfrak{p}}$ are not all equal. More precisely, v ranges over the prime ideals $\mathfrak{p}$ of K with $\|\mathfrak{p}\|_{\mathfrak{p}} = \mathrm{Norm}_{K/\mathbb{Q}}(\mathfrak{p})^{-1/[K:\mathbb{Q}]}$; and over all of the embeddings $v : K \to \mathbb{C}$ with $\|a\|_v = |a^v|^{1/[K:\mathbb{Q}]}$.

⋆ The first author is a Presidential Faculty Fellow, supported, in part, by the National Science Foundation.

We will now describe Elkies' reformulation of Vojta's conjecture (though one should note that he did not require uniformity over different number fields in the proof in [5]). Note that the conjecture that we state does follow from Vojta's "General Conjecture" 5.2.6 in [14] under the additional assumption that $[K : \mathbb{Q}]$ is bounded.

The uniform *abc*-conjecture for number fields. *For any given $\varepsilon > 0$, if $a + b + c = 0$, where a, b and c are algebraic numbers in some number field K then*

$$H(a, b, c) \ll_\varepsilon (\Delta_K N(a, b, c))^{1+\varepsilon}. \tag{1}$$

Remark. We stress that the value of H is independent of the field K. On the other hand both terms on the right side of (1) are dependent on K and one might ask whether the conjecture over some high degree field extension could possibly imply a stronger criterion for an equation defined in a subfield, than the conjecture in that subfield. In fact if L is a number field containing K, then $\Delta_K \le \Delta_L$ whereas $N_K(a, b, c) \ge N_L(a, b, c)$. However $\Delta_K N_K(a, b, c) \le \Delta_L N_L(a, b, c)$, so (1) is most stringent when K is the field of definition of a and b.

In this paper we will apply the uniform *abc*-conjecture to the very large solutions of Diophantine equations that arise from modular functions and deduce a lower bound for the class number of imaginary quadratic fields. This extends an idea of Chowla [1,2] who indicated, via a conjecture of Hall, how unlikely it is that $\mathbb{Q}(\sqrt{-p})$ has class number one, since Weber [15] showed that there would then be an enormous solution in integers to $x^3 - py^2 = -1728$ (in fact where x is the integer nearest to $e^{\pi\sqrt{p}/3}$).

Theorem 1. *The uniform abc-conjecture for number fields implies that*

$$h(-d) \ge \left\{\frac{\pi}{3} + o(1)\right\} \frac{\sqrt{d}}{\log d} \sum_{\substack{(a,b,c) \\ \text{reduced}}} \frac{1}{a}, \tag{2}$$

for any fundamental discriminant $-d < 0$ (that is, an integer which is not divisible by the square of an odd prime, with $-d \equiv 1 \pmod 4$, or 8 or $12 \pmod{16}$). The sum is over quadratic forms (a, b, c) of discriminant $-d = b^2 - 4ac$, with $-a < b \le a < c$ or $0 \le b \le a = c$ (that is, reduced).

Mahler [11] showed that if (2) holds then the Dirichlet L-function $L(s, \chi_d)$, where $\chi_d := \left(\frac{-d}{\cdot}\right)$, has no real zero in the interval $1 - c/\log d < s \le 1$, for some sufficiently small constant $c > 0$ (actually Mahler showed a little less than this but it is not hard to suitably modify his proof). We will refer to such zeros as "Siegel zeros". We can thus deduce:

Theorem 2. *The uniform abc-conjecture for number fields implies that there are no "Siegel zeros" of Dirichlet L-functions for characters $\left(\frac{-d}{\cdot}\right)$ with $-d < 0$.*

Our proof provides no insight into the question of "Siegel zeros" of Dirichlet L-functions for characters $\left(\frac{d}{\cdot}\right)$ with $d > 0$. Indeed there is no suitable analogous theory of modular functions for positive discriminants.

If we knew that the uniform *abc*-conjecture holds with an explicit constant, then our estimate in Theorem 1 could be given explicitly. As a consequence we would be able to solve several outstanding problems about quadratic fields. For example, one would be able to determine all of Euler's "convenient numbers" (*numeri idonei*), which are those d for which there is just one ideal class per genus, in the ideal class group of $\mathbb{Q}(\sqrt{-d})$.

Using a result of Selberg and Chowla [12], we will obtain, in Sect. 3, an unconditional asymptotic formula relating the quantities in (2):

Theorem 3. *For any fundamental discriminant $-d < 0$ we have*

(3)
$$h(-d) = \left\{\frac{\pi}{3} + O\left(\frac{\log\log d}{\log d}\right)\right\}\left(1 + \frac{2}{\log d}\frac{L'(1,\chi_d)}{L(1,\chi_d))}\right)^{-1}\frac{\sqrt{d}}{\log d}\sum_{\substack{(a,b,c)\\ \text{reduced}}}\frac{1}{a},$$

The estimate given in (2) is asymptotically the same as the lower bound of (3), if $L'(1,\chi_d)\big/L(1,\chi_d) = o(\log d)$. Indeed this quantity *is* $O(\log\log d)$ if the Riemann Hypothesis is true for $L(s,\chi_d)$ (as we will prove at the beginning of Sect. 3.1). Thus, under this assumption, we have an infinite sequence of "best possible examples" in the uniform *abc*-conjecture, running through a sequence of number fields with rapidly growing degree.

We note that if we were to replace Δ_K by Δ_K^A, for some $A > 1$, in the conjectural estimate (1), then we can obtain analogous, though slightly weaker, results.

Elkies suggested to us that one might apply our same methods to other Diophantine equations arising from modular functions. In Sect. 4 we examine one other example and this leads to another proof of Theorems 1 and 2, and another infinite sequence of "best possible examples" in the uniform *abc*-conjecture, running through a sequence of number fields with rapidly growing degree. Perhaps if one takes algebraic points on any given modular curve, which arise from modular functions (for example, Heegner points), and then map those points to $\mathbb{P}^1$ using a Belyi map as in [5], one obtains other such "best possible examples" in the uniform *abc*-conjecture.

Zagier suggested to us that one might apply similar methods to other differences of singular moduli, using the beautiful and restrictive formulae of [9] to obtain bounds. We have not succeeded, as yet, in so producing any new examples, though this does seem to be another good avenue to pursue. We will discuss this further in Sect. 4.

Remark on Notation: Throughout $\varepsilon > 0$ will be assumed to be an arbitrarily small fixed constant. However it may be a *different* ε from one line to the next.

As usual $A \ll B$ and $B \gg A$ both mean that there exists a constant $c > 0$ such that $A < cB$ for all such A and B. We write $f \asymp g$ when $f \ll g$ and $g \ll f$.

The value of an infinite sum $\sum_\rho$, over zeros of an L-function, should be understood to mean $\lim_{T\to\infty} \sum_{\rho:\ |\mathrm{Im}(\rho)|\le T}$.

2. Complex multiplication

For τ in the upper half plane, set $q = e^{2i\pi\tau}$. As explained by Weber [15], the classical theory of complex multiplication tells us about special values of the *j-invariant*,

$$j(\tau) := \frac{\left(1 + 240\sum_{n\ge1}\left(\sum_{d|n} d^3\right)q^n\right)^3}{q\prod_{n\ge1}(1-q^n)^{24}} = \frac{1}{q} + 744 + 196884q + \dots \tag{4}$$

and the functions $\gamma_2(\tau)$ and $\gamma_3(\tau)$ related to $j(\tau)$ by

$$j(\tau) = \gamma_2(\tau)^3 = \gamma_3(\tau)^2 + 1728. \tag{5}$$

Suppose that $-d$ is a fundamental discriminant. The value of $j(\tau)$ at $\tau = \frac{-1+\sqrt{-d}}{2}$ or $\frac{\sqrt{-d}}{2}$ (as $-d \equiv 1$ or $0 \pmod 4$) is an algebraic integer whose conjugates are the numbers $j(\tau^*)$, where τ^* runs through the values $\frac{-b+\sqrt{-d}}{2a}$ as $ax^2 + bxy + cy^2$ runs through a complete set of representative quadratic forms from each equivalence class of positive definite binary quadratic forms of discriminant $-d$.

Weber [15] notes that if the class number $h(-d) = 1$ then $\gamma_2(\tau)$ and $\gamma_3(\tau)/\sqrt{-d}$ are both integers, and indeed very large integers by (4). Chowla [1,2] observed that this is very unlikely to happen in view of the relation (5). Indeed Chowla so deduced that there are only finitely many d with $h(-d) = 1$ by applying Hall's conjecture to the Diophantine equation emerging from (5). Hall's conjecture is a consequence of the original *abc*-conjecture, and it is an easy exercise to make the same deduction directly from the original *abc*-conjecture. The uniform *abc*-conjecture for number fields allows us to extend Chowla's observation to the general case, by working in a field containing both $\gamma_2(\tau)$ and $\gamma_3(\tau)$.

If d is relatively prime to 6 then Weber showed that $\gamma_2(\tau)$ and $\gamma_3(\tau)$ both belong to the field $M = k(j(\tau))$, where $k := \mathbb{Q}(\sqrt{-d})$. This field is the Hilbert class field of k, which is the maximal unramified abelian extension of k; as such we have $\Delta_M = \Delta_k = \sqrt{d}$. We now bound the discriminant of the field containing $\gamma_2(\tau)$ and $\gamma_3(\tau)$, no matter what the value of $\gcd(6, d)$.

Lemma 1. *If $K = k(\gamma_2(\tau), \gamma_3(\tau))$, where τ is as above, then $\Delta_K \le 6\sqrt{d}$.*

Proof. For many of the facts used in this proof see [15] and [13]. Both $\gamma_2(\tau)$ and $\gamma_3(\tau)$ are in the field of modular functions of level 6; that is, both are invariant under $\Gamma(6)$, and the coefficients of the Fourier expansion at each cusp all belong to the field of sixth roots of unity. Thus both $\gamma_2(\tau)$ and $\gamma_3(\tau)$ are in L, the ray class field of k (mod 6), by Shimura's Reciprocity Law (actually these facts were already well-known to Weber [15]). Thus $K \subseteq L$ and $\Delta_K \le \Delta_L$. Now, by the conductor-discriminant formula, the relative discriminant of L/k is given by $D_{L/k} = \prod_\chi f_\chi$, where the product is over all characters χ of the ray class group (mod 6) of k, and f_χ is the conductor of χ. Moreover f_χ divides 6 for all such χ, and so $D_{L/k}$ divides $6^{[L:k]}$, which implies that $\Delta_L/\Delta_k \le 6$. The result follows since $\Delta_k = \sqrt{d}$.

Remark: A more careful analysis would allow us to reduce the factor of 6.

Proof of Theorem 1. Note that for any algebraic integer α one has $N(\alpha, 1) \le H(\alpha, 1)$, so that in a solution to

$$\gamma_3(\tau)^2 - \gamma_2(\tau)^3 + 1728 = 0 \tag{5'}$$

we have

$$\begin{aligned} N_K(\gamma_2(\tau)^3, \gamma_3(\tau)^2, 1728) &\ll N_K(\gamma_2(\tau), 1) N_K(\gamma_3(\tau), 1) \\ &\le H(\gamma_2(\tau), 1) H(\gamma_3(\tau), 1) \\ &= H(\gamma_2(\tau)^3, 1)^{1/3} H(\gamma_3(\tau)^2, 1)^{1/2} \\ &\ll H(\gamma_2(\tau)^3, \gamma_3(\tau)^2, 1728)^{5/6}. \end{aligned}$$

Therefore, by applying the uniform *abc*-conjecture to (5′) in the field K, and using Lemma 1 to bound Δ_K, we deduce that

$$H(\gamma_2(\tau)^3, \gamma_3(\tau)^2, 1728) \ll_\varepsilon d^{3+\varepsilon}.$$

Therefore, by (5), we have

$$H(j(\tau), 1) \le H(j(\tau), j(\tau) - 1728, 1728) \ll_\varepsilon d^{3+\varepsilon}. \tag{6}$$

We now determine a lower bound for $H(j(\tau), 1)$. Gauss [6] showed that there is a representative τ^* of every ideal class with $a \le \sqrt{d/3}$ (and this corresponds to the reduced quadratic form in every equivalence class of binary quadratic forms). Since $|1/q| = e^{\pi\sqrt{d}/a}$, we deduce from the q-expansion for $j(\tau^*)$ in (4) that

$$\max\{|j(\tau^*)|, 1\} \asymp e^{\pi\sqrt{d}/a}.$$

Remembering that in the definition of height the valuations were normalized to take account of the degree of the field extension, we thus have

(7)
$$H(j(\tau),1)=\left(\prod_{\tau^*}\max\{|j(\tau^*)|,1\}\right)^{1/h(-d)}\asymp\exp\left(\frac{1}{h(-d)}\sum_a\frac{\pi\sqrt{d}}{a}\right).$$

Comparing this with (6) implies Theorem 1.

Remark. Assuming the Generalized Riemann Hypothesis for $L(s,\chi)$ (so that we can take $L'(1,\chi)/L(1,\chi)=O(\log\log d)$ in Theorem 3) we deduce that $H(j(\tau),1)=d^3(\log d)^{O(1)}$, from (7). Assuming the uniform *abc*-conjecture, the proof of Theorem 1 implies that $N_K(\gamma_2(\tau),1)N_K(\gamma_3(\tau),1))\gg H(j(\tau),1)^{1-o(1)}/d^{1/2}=d^{5/2+o(1)}$. On the other hand $N_K(\gamma_2(\tau),1)\le H(\gamma_2(\tau),1)=H(j(\tau),1)^{1/3}=d(\log d)^{O(1)}$, and similarly $N_K(\gamma_3(\tau),1)\le H(\gamma_3(\tau),1)\asymp H(j(\tau),1)^{1/2}=d^{3/2}(\log d)^{O(1)}$. Combining these estimates gives

(8)
$$N_K(\gamma_2(\tau),1)=H(\gamma_2(\tau),1)^{1+o(1)}=d^{1+o(1)}$$
$$\text{and}\quad N_K(\gamma_3(\tau),1)=H(\gamma_3(\tau),1)^{1+o(1)}=d^{3/2+o(1)}.$$

The number $J:=\mathrm{Norm}_{K/\mathbb{Q}}(\gamma_2(\tau))$ has many extraordinary algebraic properties, as shown by Deuring [4] and Gross and Zagier [9]. Since $N_K(\gamma_2(\tau),1)=H(\gamma_2(\tau),1)^{1+o(1)}$, one might guess that the height and conductor of J are of roughly the same size. One can show that $H(J)=H(\gamma_2(\tau),1)^{\{1+o(1)\}[K:\mathbb{Q}]}$, and it is evident that the prime divisors of $N(J)$ and $N_K(\gamma_2(\tau),1)^{[K:\mathbb{Q}]}$ must be the same. However $N(J)$ is by definition square-free, whereas the prime factors p of $N_K(\gamma_2(\tau),1)^{[K:\mathbb{Q}]}$ may occur with multiplicity, perhaps even high multiplicity, corresponding to the number of different prime ideals of K lying above p which divide $\gamma_2(\tau)$. Indeed we show in Sect. 5 that $N(J)=H(J)^{o(1)}$ assuming the Generalized Riemann Hypothesis.

3. Formulae for L-functions — The proof of Theorem 3

Throughout this section we let χ be the character $\chi_d:=\left(\frac{-d}{\cdot}\right)$.

3.1. Evaluating L-functions at $s=1$

For any $y\ge 2$ we have, by partial summation,

(9)
$$\frac{L'(1,\chi)}{L(1,\chi)}=-\sum_{p\text{ prime}}\frac{\chi(p)\log p}{p-\chi(p)}=-\int_y^\infty\frac{\mathrm{d}\psi(t,\chi)}{t}+O(\log y),$$

where, as usual, $\psi(t,\chi):=\sum_{p^m\le t}\chi(p^m)\log p$. If we assume the Generalized Riemann Hypothesis for $L(s,\chi)$ then $\psi(t,\chi)=O(t^{1/2}\log^2(dt))$ (as in Sect. 20 of [3]), and so the right side of (9) is $O(\log\log d)$, when $y=\log^4 d$.

Another approach, using the explicit formula (17) and the functional equation (13) of pages 82 and 83 of [3], bearing in mind that χ is real, so that the zeros of the L-function are symmetric about the lines $\mathrm{Re}(s) = 1/2$ and $\mathrm{Im}(s) = 0$, yields

$$\frac{L'(1,\chi)}{L(1,\chi)} + \frac{1}{2}\log d = \sum_{\rho:\ L(\rho,\chi)=0} \frac{1}{\rho} + c_0$$
$$= \sum_{\rho=\beta\in\mathbb{R}} \frac{1}{\beta} + \sum_{\rho=\beta+i\gamma,\ \gamma\neq 0} \frac{\beta}{\beta^2+\gamma^2} + c_0,$$

where $c_0 = \frac{1}{2}\{\log(\pi)+\gamma_0+(\chi(-1)+1)\log 2\} > 0$ and $\gamma_0 \approx 0.577215665\ldots$ is the Euler-Mascheroni constant. Notice that every term here is positive since $0 < \beta < 1$ and so, by pairing up the zeros $\beta + i\gamma$ and $1 - \beta - i\gamma$, we get

$$\frac{\beta}{\beta^2+\gamma^2} + \frac{1-\beta}{(1-\beta)^2+\gamma^2} \gg \frac{1}{1+\gamma^2}.$$

The number of zeros with $|\gamma| \le T$ is $\{T/\pi + O(1)\}\log(dT/2e\pi)$ (page 101 of [3]), and so there are $\gg \log d$ zeros with $T_0 \le |\gamma| \le 2T_0$, for sufficiently large T_0. Combining these last two estimates we thus deduce that

$$\frac{L'(1,\chi)}{L(1,\chi)} + \frac{1}{2}\log d \gg \log d. \tag{10}$$

This estimate will prove useful in the next subsection when we prove Theorem 3.

Remark 1: By pairing the ρ and $\overline{\rho}$ terms together, we deduce from (4) of page 102 of [3], that $L'(\sigma,\chi)/L(\sigma,\chi) = \sum_\rho \mathrm{Re}(1/(\sigma-\rho)) + O(\log d)$ uniformly for $1 \le \sigma \le 2$. Since $0 < \mathrm{Re}(1/(1-\rho)) \ll \mathrm{Re}(1/(\sigma-\rho))$ when $\sigma = 1 + 1/\log d$ and $|1-\rho| \gg 1/\log d$, we obtain

$$\frac{L'(1,\chi)}{L(1,\chi)} = \frac{1}{1-\beta} + O\left(\log d + \frac{L'(\sigma,\chi)}{L(\sigma,\chi)}\right),$$

where β is the "Siegel zero", if it exists (otherwise there is no term $1/(1-\beta)$ here). However $|L'(\sigma,\chi)/L(\sigma,\chi)| \le |\zeta'(\sigma)/\zeta(\sigma)| \ll \log d$, so we obtain

$$\frac{L'(1,\chi)}{L(1,\chi)} = \frac{1}{1-\beta} + O(\log d).$$

In fact if $(1-\beta)\log d = o(1)$ then one can modify this argument to improve the error term to $o(\log d)$.

This estimate, combined with Theorem 3, implies Mahler's result discussed in the introduction. Related estimates have been obtained in [7] and [8]; and our remark above can be deduced directly from (4.2) in [10].

Remark 2: Under the assumption of the Generalized Riemann Hypothesis we may deduce, from formulae above, that the estimate

$$\sum_{\rho:\ L(\rho,\chi)=0} \frac{1}{\rho} = \frac{1}{2}\log d + O(\log\log d)$$

holds uniformly for all quadratic characters $\chi \pmod d$.

3.2. The Selberg-Chowla formula

In this subsection we will prove Theorem 3. In [12] Selberg and Chowla give a highly convergent expansion for Epstein's zeta-function which can be deduced as a consequence of Kronecker's limit formula (when summed over all ideal classes). They deduce from this an identity, given on the last line of page 109 in [12] (which contains an important typographical error where "$e^{2\pi n i(b_j+i\sqrt{|d|})/2a_j}$" appears incorrectly as "$e^{2\pi n i\,\frac{b_j+i\sqrt{|d|}}{2a_j}}$"). Now $a \le \sqrt{d/3}$ for any reduced binary quadratic form (a, b, c) of discriminant $-d$, so that $|\exp(2i\pi n(b+i\sqrt{d})/2a)| = \exp(-\pi n\sqrt{d}/a) \le 1/C^n$ where $C = \exp(\pi\sqrt{3})$, and therefore

$$\left|\sum_{n\ge1}\left(\sum_{d|n} d\right) e^{2i\pi n\frac{(b+i\sqrt{d})}{2a}}\right| \le \sum_{n\ge1}\left(\sum_{d|n} d\right) C^{-n}$$
$$= \sum_{d\ge1}\frac{d}{C^d-1} = .004390084081\ldots$$

Thus one can deduce from Selberg and Chowla's identity that

$$L'(1,\chi) = \frac{\pi^2}{6}\sum_{\substack{(a,b,c)\\ \text{reduced}}}\frac{1}{a} + \frac{\pi}{\sqrt{d}}\sum_{\substack{(a,b,c)\\ \text{reduced}}}\log(a/d) + O\left(\frac{h(-d)}{\sqrt{d}}\right).$$

By adding $\pi h(-d)\log d/2\sqrt{d}$ to both sides, and noting that each $\log(\sqrt{d}/a) \gg 1$, we obtain

(11)

$$\frac{\pi h(-d)}{\sqrt{d}}\left(\frac{L'(1,\chi)}{L(1,\chi)} + \frac{1}{2}\log d\right) = \frac{\pi^2}{6}\sum_{\substack{(a,b,c)\\ \text{reduced}}}\frac{1}{a} + O\left(\frac{1}{\sqrt{d}}\sum_{\substack{(a,b,c)\\ \text{reduced}}}\log(\sqrt{d}/a)\right),$$

using Dirichlet's class number formula, $L(1,\chi) = \pi h(-d)/\sqrt{d}$ for $d > 4$.

We now bound the error term in (11). Let $\rho(a)$ denote the number of reduced binary quadratic forms (a, b, c) of discriminant $-d$, for some integers b and c. Let $\rho_1(a)$ denote the number of distinct solutions $b \bmod 2a$

to $b^2 \equiv -d \pmod{4a}$. By the Chinese Remainder Theorem we have $\rho_1(a) \asymp \prod_{p|4a}\{1 + (-d/p)\}$, and, in fact $\rho(a) \le r(a)$, the number of distinct divisors of a. Now, b is, by definition, the least residue in absolute value from a residue class mod $2a$ of solutions to $b^2 \equiv -d \pmod{4a}$; therefore $\rho(a) \le \rho_1(a)$. On the other hand if $|b| \le a$ and $b^2 \equiv -d \pmod{4a}$ with $a < \sqrt{d}/2$ then when we define $c := (b^2+d)/4a$ we get $c > a$ so that (a, b, c) is a reduced form. Therefore if $a < \sqrt{d}/2$ then $\rho(a) = \rho_1(a)$.

Since $\log(\sqrt{d}/a)$ is a decreasing function in a, we thus deduce that

$$\sum_{\substack{(a,b,c)\\ \text{reduced}}} \log(\sqrt{d}/a) = \sum_{a\ge 1} \rho(a)\log(\sqrt{d}/a) \le \sum_{a\le A} r(a)\log(\sqrt{d}/a),$$

where A is chosen to be the smallest integer for which $\sum_{a\le A} r(a) \ge h(-d)$. Dirichlet showed that $\sum_{a\le A} r(a) \sim A\log A$ so that $A \sim h(-d)/\log(2h(-d))$. Therefore, by partial summation, using Dirichlet's estimate, the error term in (11) is

$$\ll \frac{1}{\sqrt{d}} \sum_{a\le A} r(a)\log(\sqrt{d}/a) \ll \frac{A\log A}{\sqrt{d}} \log(\sqrt{d}/A)$$

$$\ll \frac{h(-d)}{\sqrt{d}} \log\left(\frac{\sqrt{d}}{h(-d)}\log h(-d)\right) \ll \frac{\log\log d}{\log d} \max\left\{1, \frac{h(-d)\log d}{\sqrt{d}}\right\}.$$

However, by (10), the left side of (11) is $\gg h(-d)\log d/\sqrt{d}$; and, since there is always the principal form, with $a = 1$, the main term on the right side of (11) is $\gg 1$. Therefore the above estimate for the error term in (11) does imply (3) after suitable re-arrangement, and thus we have proved Theorem 3.

4. Some suggested generalizations

4.1. The λ-function

In an email dated August 22nd, 1994, Elkies suggested that one might try the same approach with the λ-function. Any elliptic curve E can be written in the form $y^2 = 4x(x-1)(x-\lambda)$. There are, generically, six choices for λ, the roots of the equation

$$f(x) := 256(x^2 - x + 1)^3 - j(E)(x^2 - x)^2, \tag{12}$$

where $j(E)$ is the j-invariant of the elliptic curve E. If λ is a root of (12) then the six roots are

$$\lambda,\ 1/\lambda,\ 1-\lambda,\ 1/(1-\lambda),\ (\lambda-1)/\lambda,\ \lambda/(\lambda-1).$$

Let λ be the root of (12) which is largest in absolute value. If $|\lambda| \le 2$ then all of the roots of (12) have size between $1/2$ and 2, and so $\prod_{f(\alpha)=0} \max\{1, |\alpha|\} \asymp 1$. If $|\lambda| > 2$ then $\max\{1, |\lambda|\}$, $\max\{1, |1-\lambda|\} \asymp |\lambda|$, whereas $\max\{1, |\alpha|\} \asymp 1$ for $\alpha = 1/\lambda,\ 1/(1-\lambda),\ (\lambda-1)/\lambda$, or $\lambda/(\lambda-1)$. From (12) we have $j \asymp |\lambda|^2$, so that

$$\prod_{f(\alpha)=0} \max\{1, |\alpha|\} \asymp \max\{1, |j|\}. \tag{13}$$

Now λ satisfies the equation $\lambda + (1-\lambda) = 1$. Moreover λ and $1-\lambda$ are evidently 2-units, by (12), so that $N(\lambda, 1-\lambda, 1) \asymp 1$. Therefore the uniform *abc*-conjecture implies that $H(\lambda, 1-\lambda, 1) \ll_\varepsilon \Delta_L^{1+\varepsilon}$, where $L = M(\lambda)$.

Now, λ always belongs to the ray class field (mod 2), which is of degree ≤ 3 over M, so that $[L : M] \le 3$. It can be shown that $[L : M]$ equals each of $1, 2$ and 3 infinitely often. In the case that $[L : M] = 3$ then by studying automorphisms we find that the the equation in (12) splits into two cubics, the first with roots $\lambda, (\lambda-1)/\lambda, 1/(1-\lambda)$, the other with roots $1/\lambda, \lambda/(\lambda-1), 1-\lambda$. Thus $H(\lambda, 1-\lambda, 1) \asymp H(j(\tau), 1)^{1/6}$. Combining the last two estimates thus gives

$$H(j(\tau), 1) \ll_\varepsilon \Delta_L^{6+\varepsilon}.$$

This implies (6) since $\Delta_L \ll d^{1/2}$ (as in Lemma 1).

We will return to this, and the cases where $[L : M] = 1$ or 2 in a subsequent paper.

4.2. Differences of singular moduli

In an email dated December 2nd, 1994, Zagier asked us whether similar methods might be applied to other differences of singular moduli (the value of j at a quadratic imaginary number τ). In [9], Gross and Zagier showed that the norm of differences $j(\tau_1) - j(\tau_2)$ have only small prime factors and then often to quite high powers (indeed they show how to determine to what exact power each prime appears), when the discriminants of τ_1 and τ_2 are relatively prime (see the remark after the proof of Theorem 1, above). Our proof of Theorem 1 may be viewed as applying the uniform *abc*-conjecture to the equation

$$(j(\tau) - j(i)) + (j(i) - j(\omega)) = (j(\tau) - j(\omega)),$$

where $i^2 = -1$ and $\omega^2 + \omega + 1 = 0$, since $j(i) = 1728$ and $j(\omega) = 0$. Thus, in general one might look at

$$(j(\tau_1) - j(\tau_2)) + (j(\tau_2) - j(\tau_3)) = (j(\tau_1) - j(\tau_3)), \tag{14}$$

especially if the corresponding discriminants are pairwise coprime, since we then have a considerable amount of information available from [9]: moreover if $K = \mathbb{Q}(j(\tau_1), j(\tau_2), j(\tau_3), \sqrt{-d_1}, \sqrt{-d_2}, \sqrt{-d_3})$ then $\Delta_K \asymp \sqrt{d_1 d_2 d_3}$ and $[K : \mathbb{Q}] = 8h(-d_1)h(-d_2)h(-d_3)$.

As yet we have been unable to succeed with this strategy, perhaps because we used the fact, in the proof of Theorem 1, that the relative discriminant for the field extension containing $j(\tau)^{1/3}$ and $(j(\tau)-1728)^{1/2}$ over M is absolutely bounded (see Lemma 1), whereas we have not determined an analagous property in general.

Another possibility would be to consider (14) when τ_1, τ_2, τ_3 are all unequal but have the same discriminant, since Gross and Zagier [9] also give formulae to describe the discriminant of the minimum polynomial for $j(\tau)$.

5. Estimates for the norm of $j(\tau)$

Let $A := \left|\mathrm{Norm}_{\mathbb{Q}(j(\tau))/\mathbb{Q}}(j(\tau))\right|$, and $J := \mathrm{Norm}_{K/\mathbb{Q}}(\gamma_2(\tau))$ as above.

5.1. The height of J

We will show that $H(A) = H(j(\tau),1)^{\{1+o(1)\}h(-d)}$ assuming the Generalized Riemann Hypothesis, from which it follows that $H(J) = H(\gamma_2(\tau),1)^{\{1+o(1)\}[K:\mathbb{Q}]}$ by appropriate scaling. Now, from (7) we have that

$$\begin{aligned}\left|\mathrm{Norm}_{\mathbb{Q}(j(\tau))/\mathbb{Q}}(j(\tau))\right| &= \prod_{\tau^*} \max\{|j(\tau^*)|,1\}\min\{|j(\tau^*)|,1\}\\ &= \exp\left(\pi\sqrt{d}\sum_a \frac{1}{a} + O(h(-d))\right)\prod_{\tau^*}\min\{|j(\tau^*)|,1\}.\end{aligned}$$

We thus need to understand "small values" of $j(\tau)$: The only zero of $j(\tau)$ occurs at ω, a primitive cube root of unity inside the fundamental domain for $\mathrm{SL}(2,\mathbb{Z})$. In fact j has a zero of order three there so that $|j(\tau)| \asymp |\tau-\omega|^3$ in a small ball around ω. Thus if $\tau^* = (-b+\sqrt{-d})/2a$ and $|j(\tau^*)|$ is sufficiently small then, as $|b| \le a \le \sqrt{d/3}$ for a reduced form,

$$\begin{aligned}|j(\tau^*)| &\asymp |1-|b|/a|^3 + |\sqrt{d/3}/a - 1|^3 \asymp ((\sqrt{d/3}-|b|)/a)^3\\ &\asymp \left(1-\frac{|b|}{\sqrt{d/3}}\right)^3 \asymp \left(1-\frac{3b^2}{d}\right)^3 \gg \left(1-\frac{3a^2}{d}\right)^3 \gg \frac{1}{d^3}\end{aligned}$$

since we must have $|b|, a \asymp \sqrt{d/3}$ for this to be small.

Therefore if $|j(\tau^*)| < 1/\log^9 d$ then $\sqrt{d/3} - a \ll \sqrt{d}/\log^3 d$ and so the number of such forms is $\leq \sum_{0\leq\sqrt{d/3}-a\ll\sqrt{d}/\log^3 d} \rho(a) \leq \sum_{0\leq\sqrt{d/3}-a\ll\sqrt{d}/\log^3 d} r(a) \ll \sqrt{d}/\log^2 d$; and so their total contribution to $\prod_{\tau^*} \min\{|j(\tau^*)|, 1\}$ is $e^{O(\sqrt{d}/\log d)}$. The contribution of the remaining forms is $\geq 1/\log^{9h(-d)} d$. Thus we have proved that

$$A = \exp\left(\pi\sqrt{d}\left(\sum_a \frac{1}{a} + O\left(\frac{1}{\log d}\right)\right) + O(h(-d)\log\log d)\right)$$

$$= \exp\left(\pi\sqrt{d}\left(\sum_a \frac{1}{a}\left\{1 + O\left(\frac{\log\log d}{\log d}\right)\right\}\right)\right), \tag{15}$$

the last error term obtained by substituting (10) into Theorem 3 to get the upper bound $h(-d) \ll (\sqrt{d}/\log d)\sum_a \frac{1}{a}$. More accurately, Theorem 3 under the assumption of the Generalized Riemann Hypothesis becomes

$$3h(-d)(\log d + O(\log\log d)) = \pi\sqrt{d}\sum_a \frac{1}{a}.$$

Using this formula to estimate the main term in (15) gives

$$H(A) = \left(d(\log d)^{O(1)}\right)^{3h(-d)} = H(j(\tau), 1)^{\{1+o(1)\}h(-d)}, \tag{16}$$

by (8), as desired.

5.2. The conductor of J

For simplicity suppose that d is prime and $-d \equiv 1 \pmod 6$. By the remarks between (1.5) and (1.6) in [9] we see that $N(J)$ is the product of those primes $\ell \equiv 2 \pmod 3$ for which $3d$ can be written as $3d = x^2 + \ell y^2 + 3\ell z^2$ where x, y and z are integers with $y + z$ even. In other words ℓ is the unique prime $\equiv 2 \pmod 3$ which divides $3d - x^2$ to an odd power. We deduce that $\ell \leq 3d/4$.

We will get an upper bound for the number, ν, of distinct prime factors of $N(J)$ as follows: For given small $\kappa > 0$, let L be the set of primes $\ell \leq d^\kappa$ for which $\ell \equiv 2 \pmod 3$ and $(-d/\ell) = -1$. Then

(17)

$$\nu \leq |L| + \#\{x < \sqrt{3d} : \text{ Either } \ell \nmid 3d - x^2 \text{ or } \ell^2 | 3d - x^2 \text{ for all } \ell \in L\}$$
$$\ll |L| + \sqrt{d}\prod_{\ell\in L}\left(1 - \frac{2}{\ell}\right) \ll \sqrt{d}\prod_{\ell\in L}\left(1 - \frac{1}{\ell}\right)^2$$

by the fundamental Lemma of the small sieve. If $\ell \nmid 3d$ and $\ell \leq d^\kappa$ then

$$\left(1 - \frac{1}{\ell}\right)\left(1 + \frac{(-3/\ell)}{\ell}\right)\left(1 + \frac{(-d/\ell)}{\ell}\right)\left(1 - \frac{(3d/\ell)}{\ell}\right) = \begin{cases}(1 - 1/\ell)^4 \text{ if } \ell \in L\\ 1 \text{ if } \ell \notin L\end{cases}$$

Thus, since d is prime,

$$\prod_{\ell\in L}\left(1-\frac{1}{\ell}\right)^2 \asymp \left\{\prod_{\ell\le d^\kappa}\left(1-\frac{1}{\ell}\right)\left(1+\frac{(-d/\ell)}{\ell}\right)\left(1-\frac{(3d/\ell)}{\ell}\right)\right\}^{1/2}$$
$$\asymp \left(\frac{1}{\log d}\prod_{\ell\le d}\left(1+\frac{(-d/\ell)}{\ell}\right)\left(1-\frac{(3d/\ell)}{\ell}\right)\right)^{1/2}.$$

Inserting this into (17) we obtain

$$\log N(J) \le \nu\log d$$
$$\ll \sqrt{d\log d}\left(\prod_{\ell\le d}\left(1+\frac{(-d/\ell)}{\ell}\right)\left(1-\frac{(3d/\ell)}{\ell}\right)\right)^{1/2}. \tag{18}$$

Below we will prove that

$$\sum_a \frac{1}{a} \asymp \prod_{p\le\sqrt{d}}\left(1+\frac{1}{p}\right)\left(1+\frac{(-d/p)}{p}\right)$$
$$\asymp \log d\prod_{p\le d}\left(1+\frac{(-d/p)}{p}\right), \tag{19}$$

so that, by (15),

$$\log H(J) \asymp \log H(A) \asymp \sqrt{d}\sum_a\frac{1}{a} \asymp \sqrt{d}\log d\prod_{p\le d}\left(1+\frac{(-d/p)}{p}\right).$$

Dividing this into (18) we obtain

$$\frac{\log N(J)}{\log H(J)} \ll \left(\frac{\prod_{\ell\le d}(1-(-d/\ell)/\ell)\,(1-(3d/\ell)/\ell)}{\log d}\right)^{1/2} \tag{20}$$

unconditionally. Notice that the terms in the Euler product with $(-3/\ell) = -1$ contribute a bounded amount. Thus the Euler product is $\asymp \prod_\ell(1-(-d/\ell)/\ell)^2 \ll \log d$ where the product is over those primes $\ell\le d$ with $\ell\equiv 1 \pmod 3$. Thus (20) is $o(1)$ unless

$$\sum_{\substack{\ell\le d\\(-3/\ell)=(-d/\ell)=1}}\frac{1}{\ell} \ll 1.$$

It can be shown that this never happens under the assumption of the Generalized Riemann Hypothesis. Thus we formally state:

If the Generalized Riemann Hypothesis is true then $N(J) = H(J)^{o(1)}$.

Proof of (19): From Sect. 3.2 we have

$$\sum_{\substack{(a,b,c)\\ \text{reduced}}} \frac{1}{a} = \sum_{a\geq 1} \frac{\rho(a)}{a} \leq \sum_{a\leq\sqrt{d}} \frac{\rho_1(a)}{a} \ll \sum_{a\leq\sqrt{d}} \frac{\prod_{p|a}\{1+(-d/p)\}}{a}$$
$$\leq \prod_{p\leq\sqrt{d}} \left(1+\frac{\{1+(-d/p)\}}{p-1}\right),$$

which implies the upper bound implicit in (19). In the other direction we have

$$\prod_{\substack{p\leq\sqrt{d}\\ (-d/p)=-1}} \left(1+\frac{2}{p-1}\right) \sum_{\substack{(a,b,c)\\ \text{reduced}}} \frac{1}{a} \geq \sum_{\substack{m\leq\sqrt{d}\\ p|m \Longrightarrow (-d/p)=-1}} \frac{2^{\omega(m)}}{m} \sum_{a\leq\sqrt{d}/2} \frac{\rho_1(a)}{a}$$
$$\gg \sum_{n\leq\sqrt{d}/2} \frac{2^{\omega(n)}}{n},$$

since every integer $n < d$ may be written in the form am where $p|a$ if $(-d/p) = 1$, and $p|m$ if $(-d/p) = -1$, and as $\rho_1(a) \gg 2^{\omega(a)}$ for such a. Now $2^{\omega(n)}$ equals the number of squarefree divisors d of n, and so is at least the number of pairs d, r, each $\leq \sqrt{x}$ with $dr = n$ and d squarefree. Therefore

$$\sum_{n\leq x} \frac{2^{\omega(n)}}{n} \geq \sum_{\substack{d\leq\sqrt{x}\\ d \text{ squarefree}}} \frac{1}{d} \sum_{r\leq\sqrt{x}} \frac{1}{r} \gg \log^2 x.$$

Combining these last two displayed equations, with $x = \sqrt{d}/2$ gives, via Mertens' Theorem,

$$\sum_{\substack{(a,b,c)\\ \text{reduced}}} \frac{1}{a} \gg \prod_{\substack{p\leq\sqrt{d}\\ (-d/p)=1}} \left(1+\frac{2}{p-1}\right),$$

which gives the lower bound implicit in (19).

Acknowledgements. We would like to thank Barry Mazur, Hugh Montgomery, Ken Ono, Fernando Rodriguez-Villegas, Jean-Pierre Serre, Paul Vojta, Mark Watkins, the anonymous referee, and particularly Noam Elkies, Robert Rumely and Don Zagier for useful discussions and remarks concerning this paper.

References

1. S. Chowla, On a conjecture of Marshall Hall, Proc. Nat. Acad. Sci. **56** (1966), 417–418
2. S. Chowla, Remarks on class-invariants and related topics, chapter VI of 'Seminar on Complex Multiplication' (eds A. Borel et al.) Lect. Notes Math. **21** (1966)
3. H. Davenport, Multiplicative Number Theory (2nd ed.), Springer, New York, 1980
4. M. Deuring, Teilbarkeitseigenschaften der singulären Modulen der elliptischen Funktionen und die Diskriminante der Klassengleichung, Comm. Math. Helvet. **19** (1946), 74–82
5. N. Elkies, *ABC* implies Mordell, Int. Math. Res. Not. **7** (1991), 99–109
6. C.F. Gauss, Disquisitiones Arithmeticae, Yale Univ. Press, New Haven, 1966
7. D.M. Goldfeld, A. Schinzel, On Siegel's Zero, Ann. Scuola Norm. Sup. Pisa(4), **2** (1975), 571–583
8. D.M. Goldfeld, An asymptotic formula relating the Siegel zero and the class number of quadratic fields, Ann. Scuola Norm. Sup. Pisa(4), **2** (1975), 611–615
9. B.H. Gross, D.B. Zagier, On singular moduli, J. reine angew. Math. **355** (1985), 191–220
10. D.R. Heath-Brown, Prime twins and Siegel zeros, Proc. London Math. Soc. **47** (1983), 193–224
11. K. Mahler, On Hecke's Theorem on the real zeros of the L-functions and the class number of quadratic fields, J. London Math. Soc. **9** (1934), 298–302
12. A. Selberg, S. Chowla, On Epstein's zeta-function, J. reine angew. Math. **227** (1967), 86–110
13. H.M. Stark, L-functions at $s = 1$. IV. First derivatives at $s = 0$, Adv. in Math. **35** (1980), 197–235
14. P. Vojta, Diophantine Approximations and Value Distribution Theory, Lect. Notes Math. **1239** (1987)
15. H. Weber, Lehrbuch der Algebra, vol. 3, 2nd ed., Vieweg, Braunschweig, 1908