# VALUE DISTRIBUTION OF $L(1, \chi_d)$

MANFRED PETER

Notes to a talk given at the DMV seminar "The Riemann Zeta Function and Random Matrix Theory" in Oberwolfach, October 15 - 21, 2000.

## 1. Motivation

The following situation was already investigated by C.F. Gauß [12]: Let $d$, not a perfect square, be a *discriminant* (i.e. $d \equiv 0, 1\,(4)$). There are infinitely many *primitive* quadratic forms $f(x, y) = ax^2 + bxy + cy^2$ (i.e. $\gcd(a, b, c) = 1$) with integer coefficients and discriminant $d = b^2 - 4ac$ but they fall in only finitely many *equivalence classes*. Here two forms $f_1$ and $f_2$ are called *equivalent* iff $f_1(x, y) = f_2(\alpha x + \beta y, \gamma x + \delta y)$ for some matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Let $h(d)$ be the number of equivalence classes to a given discriminant $d$ (where in case $d < 0$ only classes of positive definite forms are counted).

Here is one reason why the size of $h(d)$ is of interest. In general no explicit formulae are known for the number of representations of integers by an individual quadratic form (where this number must be suitably defined in case $d > 0$ because then the forms are indefinite). Such formulae are only known for the sum of the representation numbers over a complete system of forms representing all equivalence classes to a given discriminant. So for $h(d) = 1$ we have the rare case that there are explicit formulae for individual representation numbers.

In the case of a *fundamental discriminant* $d$ (i.e. $d$ is squarefree if $d \equiv 1\,(4)$ and $d/4$ is squarefree and $\equiv 2, 3\,(4)$ if $d \equiv 0\,(4)$) $h(d)$ can also be interpreted as the ideal class number in the narrow sense of the quadratic field $\mathbb{Q}(\sqrt{d})$. If $d < 0$ it equals the ideal class number $h^*(d)$ in the wider sense. If $d > 0$ then $h(d) = h^*(d)$ if $u^2 - dv^2 = -4$ has a solution in integers $u, v$ and $h(d) = 2h^*(d)$ otherwise. In this interpretation the size of $h(d)$ is of interest again. The integers in the field $\mathbb{Q}(\sqrt{d})$ form a unique factorization domain iff $h^*(d) = 1$. So in general $h^*(d)$ measures how far away they are from having unique factorization.

What is the behaviour of $h(d)$ as $d$ varies? Gauß gave several conjectures some of which were subsequently proved. In many proofs a key ingredient is Dirichlet's class number formula. Let $\chi_d$ be the *Jacobi character* associated to $d$. This Dirichlet character modulo

*Date*: December 14, 2000.

$|d|$ is defined to be completely multiplicative on $\mathbb{Z}$ with

$$\chi_d(p) := \left(\frac{d}{p}\right) := \left\{ \begin{array}{rl} 1, & p \not| d,\ x^2 \equiv d\,(p) \text{ is solvable,} \\ -1, & x^2 \equiv d\,(p) \text{ is unsolvable,} \\ 0, & p|d, \end{array} \right\} \text{ for } p > 2 \text{ prime,} \quad (1.1)$$

$$\chi_d(2) := \left\{ \begin{array}{rl} 1, & d \equiv 1\,(8), \\ -1, & d \equiv 5\,(8), \\ 0, & d \equiv 0\,(4), \end{array} \right\} \quad \text{and} \quad \chi_d(-1) := \left\{ \begin{array}{rl} 1, & d > 0, \\ -1, & d < 0. \end{array} \right\} \quad (1.2)$$

Define the *Dirichlet L-series*

$$L(s, \chi_d) := \sum_{n \geq 1} \frac{\chi_d(n)}{n^s}, \quad \Re s > 1,$$

associated to $\chi_d$. In fact, since $d$ is not a square the series converges for $\Re s > 0$, and it can be extended to an entire function with respect to $s$. In case $d > 0$, let $(u_d, v_d)$ be the *fundamental solution* of Pell's equation $u^2 - dv^2 = 4$ (i.e. the solution with $u, v > 0$ and $v$ minimal; it is a non-trivial fact that a solution with $v > 0$ exists). Define $\epsilon_d := (u_d + v_d\sqrt{d})/2$. In case $d < 0$, define

$$w_d := \left\{ \begin{array}{ll} 6, & d = -3, \\ 4, & d = -4, \\ 2, & d < -4. \end{array} \right\}$$

Then the fundamental relation (*Dirichlet's class number formula*)

$$h(d) = \left\{ \begin{array}{ll} (2\pi)^{-1} w_d \sqrt{|d|} L(1, \chi_d), & d < 0, \\ (\log \epsilon_d)^{-1} \sqrt{d} L(1, \chi_d), & d > 0, \end{array} \right\} \quad (1.3)$$

holds which reduces the number theoretic/algebraic quantity $h(d)$ to the analytic quantity $L(1, \chi_d)$.

For all this, see [4], [9], [24].

Now a fundamental question about class numbers for large discriminants can be answered with Siegel's theorem: For $\epsilon > 0$ there is some $C(\epsilon) > 0$ such that $L(1, \chi_d) \geq C(\epsilon)|d|^{-\epsilon}$. Together with the easy upper bound $L(1, \chi_d) \ll \log(|d| + 2)$ it follows that, for $d < 0$,

$$|d|^{-\epsilon} \ll_\epsilon \frac{h(d)}{|d|^{1/2}} \ll \log(|d| + 2), \quad (1.4)$$

i.e. $h(d)$ grows roughly like $|d|^{1/2}$ as $d \to -\infty$. Since for no $0 < \epsilon < 1/2$ an explicit value for $C(\epsilon)$ is known this result does not help to find all those fundamental discriminants $d < 0$ with $h(d) = 1$, say. This *class number problem* of Gauß was solved in 1966 by Baker [1] and in 1967 with different methods by Stark [30].

For $d > 0$ we only get

$$d^{-\epsilon} \ll_\epsilon \frac{h(d) \log \epsilon_d}{d^{1/2}} \ll \log(d + 2).$$

Since the size of $\epsilon_d$ fluctuates wildly with $d$ it is not true that $h(d) \to \infty$ as $d \to \infty$. In fact it is conjectured that there are infinitely many fundamental discriminants $d > 0$ with

$h(d) = 1$. For the computational aspects of class numbers and some conjectures, see [7], Chapter 5, and in particular Section 5.10.

For the remainder of this article we will look at the behaviour of $h(d)$ in the mean. The symbols $c$, $c^{(1)}$, $c_1$ and so on will always denote constants (not necessarily always the same) and $\epsilon$ will denote an arbitrary positive real.

## 2. Results about the limit distribution

A function $F : \mathbb{R} \to [0,1]$ will be called a *distribution function* iff $F$ is increasing, continuous from the right, $\lim_{z\to-\infty} F(z) = 0$ and $\lim_{z\to\infty} F(z) = 1$. In 1951 Chowla and Erdös [6] showed that, for real $s > 3/4$, there is a continuous distribution function $F(\cdot, s)$ which strictly increases on $(0,\infty)$ such that

$$\lim_{x\to\infty} \frac{1}{x/2}\, \#\big\{1 \le \pm d \le x \,\big|\, L(s,\chi_d) \le z\big\} = F(z,s), \quad z \in \mathbb{R}. \tag{2.1}$$

For $s = 1$ it follows that

$$\lim_{x\to\infty} \frac{1}{x/2}\, \#\Big\{1 \le -d \le x \,\Big|\, \frac{\pi h(d)}{|d|^{1/2}} \le z\Big\} = F(z,1), \quad z \in \mathbb{R}.$$

This shows in particular that the upper and lower bounds in (1.4) cannot be replaced by positive constants. In any interval $(a,b] \subseteq \mathbb{R}^+$ of positive length there will be, for large $x$, a positive fraction $\approx F(b,1) - F(a,1)$ of all the values $\pi|d|^{-1/2}h(d)$, $1 \le -d \le x$. The method of proof rests on an approximation of $L(s,\chi_d)$ by a suitably choosen subsum of its Dirichlet series representation. This is true for many of the subsequent results but the particular choice of the subsum varies.

Barban [2] proved in 1966 that, for all $k \in \mathbb{N}$,

$$\sum_{1\le\pm d\le x} L(1,\chi_d)^k = r(k)x + O_k\big(x\exp(-c\sqrt{\log x})\big) \qquad (x\to\infty), \tag{2.2}$$

where $c > 0$ is a constant independent of $k$ and

$$r(k) = \sum_{n\ge1} \frac{\varphi(n) d_k(n^2)}{2n^3}. \tag{2.3}$$

Here $d_k(n)$ is the number of ways one can write $n$ as a product of $k$ positive integers. From this theorem part of Chowla and Erdös' result can be derived using the following theorem of Fréchet and Shohat (e.g. [3], Theorems 25.10 and 30.1): Let $(F_n)_{n\ge1}$ be a sequence of distribution functions such that

- for every $k,n \in \mathbb{N}$, the *$k$-th moment* $\alpha_k(n) := \int_{\mathbb{R}} z^k dF_n(z)$ of $F_n$ exists,
- for every $k \in \mathbb{N}$, the limit $\beta_k := \lim_{n\to\infty} \alpha_k(n)$ exists,
- the power series $\Phi(w) := 1 + \sum_{k\ge1} \beta_k w^k/k!$ has a positive radius of convergence.

Then there is a distribution function $F$ such that

- $F_n$ *converges weakly* to $F$, i.e. $\lim_{n\to\infty} F_n(z) = F(z)$ for all points of continuity $z$ of $F$,
- for all $k \in \mathbb{N}$, the value $\beta_k$ is the $k$-th moment of $F$,
- for $t \in \mathbb{R}$ close to 0, the equation $\Phi(it) = \phi(t)$ holds, where $\phi(t) = \int_{\mathbb{R}} e^{itz} dF(z)$ is the *characteristic function* of $F$.

Now

$$M_n := \#\{1 \le \pm d \le n \,|\, d \ne \square \text{ is a discriminant}\} \sim \frac{n}{2} \qquad (n \to \infty),$$

and the functions

$$F_n(z) := \frac{1}{M_n}\,\#\{1 \le \pm d \le n \,|\, L(1,\chi_d) \le z\}, \quad z \in \mathbb{R},$$

are distribution functions. According to (2.2), their $k$-th moments have the asymptotics

$$\frac{1}{M_n}\sum_{1\le \pm d \le n} L(1,\chi_d)^k \sim 2r(k) =: \beta_k \qquad (x \to \infty).$$

From (2.3) it follows easily that $\beta_k \ll \exp(c\,k\log\log k)$ as $k \to \infty$ with some constant $c > 0$ and thus $\Phi(w)$ has radius of convergence infinity. Fréchet and Shohat's theorem now shows that there is a distribution function $F$ such that (2.1) holds for $s = 1$ and $F(\cdot, 1) = F$ at all points of continuity $z$ of $F$. Furthermore, for $k \in \mathbb{N}$ the $k$-th moment of $F$ exists and equals $2r(k)$. Finally, $F$ has an entire characteristic function given by

$$\phi(t) = 1 + \sum_{k\ge 1}\frac{2r(k)}{k!}(it)^k, \quad t \in \mathbb{R}.$$

The last two pieces of information are not contained in Chowla and Erdös' result.

Subsequently Wolke [32] improved the error term in (2.2) to $O_{k,\epsilon}(x^{3/4+\epsilon})$ and Jutila [17] to $O_k(x^{1/2}(\log x)^{c_k})$.

Elliott ([10]; [11], Chapter 22) proved with ideas from probabilistic number theory that $F(e^z, s)$ is smooth in $z$, gave an explicit rate of convergence in (2.1) and calculated an Euler product representation for the characteristic function: For $\Re s > 1/2$,

$$\frac{1}{x/2}\,\#\{1 \le \pm d \le x \,|\, |L(s,\chi_d)| \le e^z\} = F(e^z, s) + O\left(\sqrt{\frac{\log\log x}{\log x}}\right), \quad x \ge 2,\ z \in \mathbb{R}, \tag{2.4}$$

where $F(\exp(\cdot), s) \in C^\infty(\mathbb{R})$ is a distribution function, all derivatives of which are bounded on $\mathbb{R}$ and which has the characteristic function

$$\phi(t,s) = \prod_p \left(\frac{1}{p} + \frac{1}{2}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^s}\right)^{-it} + \frac{1}{2}\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p^s}\right)^{-it}\right), \quad t \in \mathbb{R}, \tag{2.5}$$

where $p$ runs through the primes.

The following informal argument gives an idea of the proof: Approximate $L(s,\chi_d)$ by a partial product

$$\prod_{p\le P}\left(1 - \frac{\chi_d(p)}{p^s}\right)^{-1}$$

of its Euler product. Then

$$\log L(s,\chi_d) \approx -\sum_{p \le P}\log\left(1 - \frac{\chi_d(p)}{p^s}\right), \tag{2.6}$$

where the summands correspond to stochastically independent random variables in the following sense: For every choice $c_p \in \{0,1,-1\}$, $p \le P$, there is some discriminant $d$ with $\chi_d(p) = c_p$ for $p \le P$. This follows from the Chinese remainder theorem. For $p > 2$, the number of such $d$'s modulo $p$ is 1 in case $c_p = 0$ and $(p-1)/2$ in case $c_p = \pm 1$ since half of the residues $\not\equiv 0 \mod p$ are quadratic residues and the other half are not. Thus the probability of a randomly choosen discriminant $d$ to have $\chi_d(p) = c_p$ is $1/p$ for $c_p = 0$ and $(1-1/p)/2$ for $c_p = \pm 1$. The joint probability that $\chi_d(p) = c_p$ for all $p \le P$ is the product of these individual probabilities which follows again from the Chinese remainder theorem. The characteristic function of the random variable corresponding to the right hand side of (2.6) is therefore

$$\sum_{(c_p)_{p\le P}:\, c_p\in\{0,1,-1\}} \exp\left(it\sum_{p\le P} -\log\left(1-\frac{c_p}{p^s}\right)\right) \prod_{p\le P:\, c_p=0} \frac{1}{p} \prod_{p\le P:\, c_p=\pm 1} \frac{1}{2}\left(1-\frac{1}{p}\right)$$
$$= \prod_{p\le P}\left(\frac{1}{p} + \frac{1}{2}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^s}\right)^{-it} + \frac{1}{2}\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p^s}\right)^{-it}\right),$$

which is a partial product of (2.5). If it can be shown that this product is uniformly convergent on compact subsets of $\mathbb{R}$ then (2.4) without the error estimate follows from Kolmogorov's continuity theorem.

It seems that not much is known about the shape of the density $\delta(z)$ of $F(z,1)$ with respect to Lebesque measure. From the knowledge of the moments of $F(z,1)$ and $F(e^z,1)$ it is easy to conclude that $\delta(z)$ decreases super-exponentially as $z \to \infty$ and that 0 is a zero of infinite order of $\delta(z)$. From numerical calculations in a related situation I would guess that $\delta(z)$ is *unimodal*, i.e. it has exactly one relative maximum to the left of which it strictly increases and to the right of which it strictly decreases.

## 3. Variations

Elliott [10] obtained results similar to (2.4) where $d$ runs through prime discriminants. Barban [2] stated but did not prove that results similar to (2.2) could be proved when $d$ runs through fundamental discriminants.

The behaviour of $L(s,\chi_d)$ changes drastically if $\Re s = 1/2$. Jutila [18] showed in 1981 that

$$\sum_{1\le \pm d\le x:\, d \text{ a fund. discr.}} L\Big(\frac{1}{2},\chi_d\Big) = c^{(1)} x\log x + c^{(1)}_{\pm} x + O_\epsilon(x^{3/4+\epsilon}), \tag{3.1}$$
$$\sum_{1\le \pm d\le x:\, d \text{ a fund. discr.}} L\Big(\frac{1}{2},\chi_d\Big)^2 = c^{(2)} x(\log x)^3 + O_\epsilon(x(\log x)^{5/2+\epsilon}). \tag{3.2}$$

He also got a result similar to (3.1) for prime discriminants. $L(1/2,\chi_d)$ does not have a mean value but, according to (3.1), it is of order $\log d$ in the mean. On the other hand, from (3.2) it follows that in the quadratic mean it is of order $(\log d)^{3/2}$. Thus it is not a priori clear what the right normalization of $L(1/2,\chi_d)$ would be in order to get results on limit distributions. In fact nothing is known in this direction.

Jutila's main tools are very sharp estimates for quadratic character sums. In 1985 Goldfeld and Hoffstein [14] sharpened (3.1) by proving that

$$\sum_{1\le \pm d\le x:\, d\,\text{a fund. discr.}} L(s,\chi_d)$$
$$= \left\{ \begin{array}{ll} c(s)x + O(x^{1/2+\epsilon}) & , \quad \Re s \ge 1, \\ c(s)x + c_\pm(s)x^{3/2-s} + O_\epsilon(x^{\Theta(s)+\epsilon}) & , \quad 1/2 \le \Re s < 1,\ s \ne 1/2, \\ c(1/2)x\log x + c_\pm(1/2)x + O_\epsilon(x^{19/32+\epsilon}), & \quad s = 1/2, \end{array} \right\}$$

with explicitly given $\Theta(s)$. Besides the small improvement of the error term for $s = 1/2$ their approach is of principal interest. They use the fact that the Fourier coefficients of real analytic Eisenstein series for the metaplectic group involve the values $L(s,\chi_d)$. In later papers, with a considerably expanded version of this method, they investigate the mean values for quadratic twists $L(s,f,\chi_d)$ of Dirichlet series $L(s,f)$ associated to cusp forms $f$ of weight $k$. Of particular interest is the case where $s$ is the centre $k/2$ of the critical strip. For $L(1/2,\chi_d)$ no arithmetical meaning is known. In contrast to this if $k=2$ and $f$ comes from a modular elliptic curve $E$ then according to the Birch-Swinnerton-Dyer conjecture $L(1,f,\chi_d)$ should contain information on the group of rational points on the Elliptic curve which is obtained from $E$ by twisting (see [5] and [13]).

Besides (3.1) and (3.2) no further formulae for moments of $L(1/2,\chi_d)$ have been proved. There are conjectural formulae for all moments which include predictions of the values of certain coefficients therein on the basis of Random Matrix Theory (see [8]).

So far only mean values for class numbers were considered which are ordered according to the size of the discriminant. For positive discriminants a fundamentally different ordering is possible. In 1982 Sarnak [28] proved that

$$\sum_{\epsilon_d\le x} h(d)\log\epsilon_d = \frac{x^2}{2} + O(x^{3/2}(\log x)^3) \qquad (x\to\infty). \tag{3.3}$$

In sums of type (2.2) the quantities $h(d)$ and $\log\epsilon_d$, which are contained in $L(1,\chi_d)$, cannot be separated. Here a simple partial summation gives

$$\sum_{\epsilon_d\le x} h(d) = \mathrm{Li}\,(x^2) + O(x^{3/2}(\log x)^2)$$

where $\mathrm{Li}\,(x) := \int_2^x (\log t)^{-1}dt$, $x\ge 2$. The error term in (3.3) was improved by Iwaniec [16] to $O_\epsilon(x^{35/24+\epsilon})$ with the help of estimates for sums of Kloosterman sums. A further improvement to $O(x^{7/10+\epsilon})$ was obtained by Luo and Sarnak [21]. The main tool for the good error term in (3.3) is Selberg's trace formula. This unfortunately makes the method rather inflexible. For example, no higher moments can be treated, and the factor $\sqrt{d}$ in $h(d)\log\epsilon_d$ cannot be removed to get results on $L(1,\chi_d)$. With a method similar to Barban's it was proved in [25] that

$$\sum_{\epsilon_d\le x} (h(d)\log\epsilon_d)^k = c_k x^{k+1} + O_{\epsilon,k}(x^{k+\rho+\epsilon})$$

where $0<\rho<1$ is explicitly given and independent of $k$. The same method gives

$$\sum_{\epsilon_d\le x} L(1,\chi_d)^k = c_k^* x + O_{\epsilon,k}(x^{\rho+\epsilon})$$

and consequently

$$\lim_{x\to\infty}\frac{1}{\#\{\epsilon_d\le x\}}\,\#\bigl\{\epsilon_d\le x \,\big|\, L(1,\chi_d)\le z\bigr\} = F(z)$$

for all points of continuity $z$ of $F$ where $F$ is some distribution function. In fact it can be proved with tools similar to Elliott's that $F$ is smooth. The method also plays a role in the proof of the following arithmetical result which is of interest in connection with Farey fraction spin chains. Let

$$A := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and let $\Phi(n)$ be the number of matrices $C$ with $\operatorname{tr}(C)=n$ which can be written as products of $A$'s and $B$'s which both occur at least once. Kleban and Özlük [20] conjectured that $\Phi(n)\sim c\cdot n\log n$ as $n\to\infty$. In [19] it was proved that

$$\sum_{n\le x}\Phi(n) = \frac{6}{\pi^2}x^2\log x + O(x^2\log\log x).$$

In [26] the conjecture was proved to be false by showing that $\Phi(n)/(n\log n)$ has a smooth limit distribution.

Another variation was considered by Sarnak [29] who proved a mean value formula for class numbers on certain thin subsets of $\mathbb{N}$. He showed that for $\nu\in\mathbb{N}$,

$$\sum_{2<n\le x} h(n^{2\nu}-4) \sim c_\nu\frac{x^{\nu+1}}{\log x}.$$

So far a fixed Dirichlet series (e.g. the Riemann zeta function or an L-series attached to a cusp form) was twisted by quadratic characters. In the averaging process the set of quadratic characters was exhausted in a suitable way. But it is also possible to leave the Dirichlet series untwisted and change the underlying cusp form instead. Luo [22] considered the values $L(1,\operatorname{sym}^2 f)$ of symmetric square L-functions at 1 where $f$ runs through an orthonormal set of Maaß wave forms for the group $\mathrm{SL}_2(\mathbb{Z})$ which are simultaneous eigenforms of all Hecke operators. The forms are ordered according to the sizes of their Laplacian eigenvalues. Royer [27] investigated the analoguous situation for newforms of fixed even weight $k\in\mathbb{N}$ and arbitrary group $\Gamma_0(N)$, $N\in\mathbb{N}$. Here $f$ runs through normalized newform eigenforms for all Hecke operators. They are ordered according to the sizes of their levels N.

## 4. The method of Barban

The proof of (2.2) is done in three steps.

4.1. **Approximation of the Dirichlet series.** The value $L(1,\chi_d)^k$ can be approximated by a partial sum or a partial Euler product using partial summation and the Polya-Vinogradov inequality or an approximate functional equation. Here we use a smoothing factor $e^{-n/N}$ in the Dirichlet series expansion which leads to a short and elegant proof but does not give the best possible error term.

Let $1/2<\gamma<1$, $0<\alpha<1$. These parameters will be fixed later and are independent of $x$. For $1/2\le\sigma\le 1$, $T\ge 1$, define the rectangle

$$R(\sigma,T):=\big\{s\in\mathbb{C}\,\big|\,\sigma\le\Re s\le 2,\ |\Im s|\le T\big\}.$$

Let $x\ge 2$ and set $N:=x^\alpha$. Let $1\le|d|\le x$. Cauchy's theorem gives, for $T\ge 1$,

$$\frac{1}{2\pi i}\int_{\partial R(\gamma,T)} L(s,\chi_d)^k\Gamma(s-1)N^{s-1}ds=L(1,\chi_d)^k. \tag{4.1}$$

Stirling's formula ([9], Chapter 10) shows that for $c_1>0$ there is some $c_2>0$ such that

$$\Gamma(s)\ll e^{-\pi|\Im s|/2}|\Im s|^{c_2}$$

for $|\Im s|\ge 1$, $|\Re s|\le c_1$. Furthermore,

$$L(s,\chi_d)\ll|ds|$$

for $\Re s\ge 1/2$ ([9], Chapter 12). Thus

$$\int_{\gamma\pm iT}^{2\pm iT} L(s,\chi_d)^k\Gamma(s-1)N^{s-1}ds\ll(|d|T)^kT^{c_2}e^{-\pi T/2}N.$$

Letting $T\to\infty$ in (4.1) gives therefore

$$L(1,\chi_d)^k=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} L(s,\chi_d)^k\Gamma(s-1)N^{s-1}ds-I(d,N), \tag{4.2}$$

where

$$I(d,N):=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} L(s,\chi_d)^k\Gamma(s-1)N^{s-1}ds.$$

For $\Re s=2$,

$$L(s,\chi_d)^k=\sum_{n\ge 1}\frac{\chi_d(n)d_k(n)}{n^s}.$$

Plugging this into the integral in (4.2) and using Mellin's inversion formula

$$\frac{1}{2\pi i}\int_{\beta-i\infty}^{\beta+i\infty}\Gamma(s)y^{-s}ds=e^{-y},\quad y,\beta>0,$$

gives for the integral in (4.2) the value

$$\sum_{n\ge 1}\frac{\chi_d(n)d_k(n)}{n}e^{-n/N}. \tag{4.3}$$

This is, up to the smoothing factor $e^{-n/N}$, the Dirichlet series representation of $L(1,\chi_d)^k$. When we remove the smoothing factor in the end it will work as if it were 1 for $n \le N$ and 0 for $n > N$.

4.2. **Estimation of the remainder under GRH.** If we assume the Generalized Riemann Hypothesis then $L(s,\chi_d)$ has no zeros in $\Re s > 1/2$. By a standard procedure we can conclude that the *Generalized Lindelöf Hypothesis* is true, i.e.

$$L(s,\chi_d) \ll_\epsilon (|d|(1+|\Im s|))^\epsilon, \quad 1/2+\epsilon \le \Re s \le 2$$

(see [31], Theorem 14.2, for the Riemann zeta function). Therefore we get

$$I(d,N) \ll \int_{\mathbb{R}} (|d|(1+|t|))^\epsilon e^{-\pi|t|/2}(1+|t|)^{c_2} N^{\gamma-1} dt \ll |d|^\epsilon N^{\gamma-1}. \tag{4.4}$$

Collecting (4.2)-(4.4) gives

$$\Sigma(x) := \sum_{1\le \pm d\le x} L(1,\chi_d)^k = \sum_{n\ge 1} \frac{d_k(n)}{n} e^{-n/N} \sum_{1\le \pm d \le x} \chi_d(n) + O\big(x^{1+\epsilon}N^{\gamma-1}\big). \tag{4.5}$$

Originally, $d$ was assumed to run through non-squares. If we drop this condition the additional error is

$$O\Bigg(\sum_{n\ge 1} \frac{1}{n^{1-\epsilon}} e^{-n/N} x^{1/2}\Bigg) = O(N^\epsilon x^{1/2})$$

which fits into the error in (4.5).

4.3. **Evaluation of the main term.** For $n \in \mathbb{N}$, $x \ge 2$, set

$$S_n(x) := \sum_{1\le \pm d\le x} \chi_d(n).$$

We are going to prove now that

$$S_n(x) = \left\{ \begin{array}{ll} x\varphi(l)/(2l) + O(n^{1/2}), & n = l^2,\ l \in \mathbb{N}, \\ O(n^{1/2}\log(n+2)) \quad , & \text{otherwise.} \end{array} \right\} \tag{4.6}$$

If $n$ is a perfect square, $n = l^2$ with $l \in \mathbb{N}$, then $\chi_d(n) = 1$ for $\gcd(d,l) = 1$ and 0 otherwise. Thus

$$S_n(x) = \#\big\{1 \le \pm d \le x \,\big|\, d \equiv 0,1\,(4),\ \gcd(d,l) = 1\big\}.$$

Let $l = 2^e l'$ where $2\not| l'$ and $e \in \mathbb{N}_0$. The Chinese remainder theorem gives

$$\#\big\{d \bmod 4l' \,\big|\, d \equiv 0,1\,(4),\ \gcd(d,l) = 1\big\} = \left\{ \begin{array}{ll} 2\varphi(l'), & e = 0, \\ \varphi(l') \ , & e \ge 1. \end{array} \right\}$$

Dividing the range from $\pm 1$ to $\pm x$ into complete residue systems modulo $4l'$ gives

$$S_n(x) = \Big[\frac{x}{4l'}\Big] \cdot \left\{ \begin{array}{ll} 2\varphi(l'), & e = 0 \\ \varphi(l') \ , & e \ge 1 \end{array} \right\} + O(l') = \frac{x}{2}\frac{\varphi(l)}{l} + O(n^{1/2}).$$

This gives the first case in (4.6).

If $n$ is not a perfect square we will show that $S_n(x)$ does not contribute to the main term. To this end we must realize the function $d \mapsto \chi_d(n)$ as a linear combination of non-principal characters. Define

$$\epsilon_8'(m) := \left\{ \begin{array}{ll} 1 \ , & m \equiv 1,3\,(8), \\ -1, & m \equiv 5,7\,(8), \\ 0 \ , & \text{otherwise}, \end{array} \right\}$$

$$\epsilon_4(m) := \left\{ \begin{array}{ll} 1, & 2 \not| m, \\ 0, & 2|m, \end{array} \right\} \quad \text{and} \quad \epsilon_4'(m) := \left\{ \begin{array}{ll} 1 \ , & m \equiv 1\,(4), \\ -1, & m \equiv 3\,(4), \\ 0 \ , & \text{otherwise}. \end{array} \right\}$$

Then $\epsilon_4$ and $\epsilon_4'$ are Dirichlet characters modulo 4 and $\epsilon_8'$ is a Dirichlet character modulo 8. From the definition (1.1) and (1.2) it follows that for a discriminant $d$,

$$\chi_d(n) = \epsilon_8'(d)^{e_0} \left(\frac{d}{p_1}\right)^{e_1} \cdots \left(\frac{d}{p_r}\right)^{e_r},$$

where $n = 2^{e_0} p_1^{e_1} \cdots p_r^{e_r}$ is the prime factorization of $n$. Here one of the exponents $e_0, \ldots, e_r$ is odd since $n$ is not a perfect square. Furthermore,

$$\frac{1}{2}(\epsilon_4(m) + \epsilon_4'(m)) = \left\{ \begin{array}{ll} 1, & m \equiv 1\,(4), \\ 0, & \text{otherwise}. \end{array} \right\}$$

Thus

$$\begin{aligned} S_n(x) \ &= \sum_{1 \le \pm 4m \le x} \left(\frac{m}{p_1}\right)^{e_1} \cdots \left(\frac{m}{p_r}\right)^{e_r} \cdot \left\{ \begin{array}{ll} 1, & e_0 = 0 \\ 0, & e_0 \ge 1 \end{array} \right\} \\ &\quad + \sum_{1 \le \pm m \le x} \frac{1}{2}(\epsilon_4(m) + \epsilon_4'(m)) \epsilon_8'(m)^{e_0} \left(\frac{m}{p_1}\right)^{e_1} \cdots \left(\frac{m}{p_r}\right)^{e_r}. \end{aligned}$$

The second sum can be divided into two sums the summands of which are characters modulo $4n$ in $m$ which are non-principal by the Chinese remainder theorem. If $e_0 = 0$ then one of the $e_j$, $1 \le j \le r$, is odd, and the function in the first sum is a non-principal character modulo $n$ in $m$. Thus the Polya-Vinogradov inequality gives

$$S_n(x) \ll n^{1/2} \log(n+2),$$

which is the second case in (4.6).

(4.5) and (4.6) now give

$$\Sigma(x) = \sum_{l \ge 1} \frac{d_k(l^2)}{l^2} e^{-l^2/N} \frac{x}{2} \frac{\varphi(l)}{l} + O\Big( \sum_{n \ge 1} \frac{d_k(n)}{n} e^{-n/N} n^{1/2+\epsilon} \Big) + O\big(x^{1+\epsilon} N^{\gamma-1}\big).$$

The first error term is $O(N^{1/2+2\epsilon})$. In the main term, use $e^{-l^2/N} = 1 + O(l^2/N)$ for $l \le N^{1/2}$ and $e^{-l^2/N} = O(1)$ for $l > N^{1/2}$. This gives for the main term

$$\frac{x}{2} \sum_{l \le N^{1/2}} \frac{d_k(l^2)\varphi(l)}{l^3} + O(x\, N^{-1/2+\epsilon}).$$

Expanding the range of summation to infinity gives an additional error $O(x\, N^{-1/2+\epsilon})$. Thus

$$\Sigma(x) = \frac{x}{2} \sum_{l \geq 1} \frac{d_k(l^2)\varphi(l)}{l^3} + O(x\, N^{-1/2+\epsilon} + N^{1/2+2\epsilon} + x^{1+\epsilon} N^{\gamma-1}).$$

Since $1/2 < \gamma < 1$ and $0 < \alpha < 1$ the error is $O(x^{1-\delta})$ with some $\delta > 0$. This ends the proof of Barban's result under GRH. □

4.4. **Estimation of the remainder without GRH.** If no unproven hypothesis is assumed we must distinguish two cases.

**Case 1:** The function $L(s, \chi_d)$ has no zeros in $R(\gamma - \epsilon, (\log x)^2)$. Using the Borel-Carathéodory theorem and Hadamard's Three Circles Theorem gives

$$L(s, \chi_d) \ll_\epsilon (|d|(1 + |\Im s|))^\epsilon, \quad |\Im s| \leq \frac{1}{2}(\log x)^2,\ \gamma \leq \Re s \leq 2$$

as in [31], Theorem 14.2. Thus

$$\begin{aligned} I(d, N) \ &\ll_\epsilon \ \int_{|t| \leq (\log x)^2/2} (|d|(1 + |t|))^\epsilon e^{-\pi|t|/2} (1 + |t|)^{c_2} N^{\gamma-1} dt \\ &\quad + \int_{|t| \geq (\log x)^2/2} (|d|(1 + |t|))^k e^{-\pi|t|/2} (1 + |t|)^{c_2} N^{\gamma-1} dt \\ &\ll \ |d|^\epsilon N^{\gamma-1}. \end{aligned} \tag{4.7}$$

This estimate is of the same quality as (4.4).

**Case 2:** The function $L(s, \chi_d)$ has a zero in $R(\gamma - \epsilon, (\log x)^2)$. Here we must estimate trivially. From (4.2) and (4.3) it follows that

$$I(d, N) \ll \sum_{n \geq 1} \frac{1}{n^{1-\epsilon}} e^{-n/N} + |L(1, \chi_d)|^k \ll N^\epsilon + |d|^\epsilon. \tag{4.8}$$

We used $L(1, \chi_d) \ll \log(|d| + 2)$ which follows easily from the orthogonality relation for characters using partial summation.

Now we must show that Case 2 cannot occur too often. This is done with zero density estimates for L-series. From [23], Theorem 12.2, it follows that, for $\gamma > 4/5$,

$$\begin{aligned} &\#\big\{1 \leq |d_0| \leq x \,\big|\, d_0 \text{ a fundamental discriminant}, \\ &\qquad L(s, \chi_{d_0}) \text{ has a zero in } R(\gamma - \epsilon, (\log x)^2)\big\} \ll x^{4(1-\gamma)/\gamma + \epsilon}. \end{aligned}$$

Here some accuracy is lost since we only consider quadratic characters but use an estimate which takes into account also non-quadratic characters. Using [15] instead would allow to reduce the factor 4 in the exponent somewhat. Every discriminant $d$ can be written as $d_0 r^2$ with a uniquely determined fundamental discriminant $d_0$ and $r \in \mathbb{N}$. Since $L(s, \chi_{d_0})$ and

$$L(s, \chi_d) = L(s, \chi_{d_0}) \prod_{p|r} \left(1 - \frac{\chi_{d_0}(p)}{p^s}\right)$$

have the same zeros in $\Re s > 0$ it follows that

$$\#\big\{1 \leq |d| \leq x \,\big|\, L(s,\chi_d) \text{ has a zero in } R(\gamma-\epsilon,(\log x)^2)\big\} \\ \ll \sum_{1\leq r\leq x^{1/2}} \Big(\frac{x}{r^2}\Big)^{4(1-\gamma)/\gamma+\epsilon} \ll x^{\max\{1/2,4(1-\gamma)/\gamma\}+\epsilon}. \tag{4.9}$$

(4.7)-(4.9) now show that in the unconditional case there is an additional error

$$O\big(x^{\max\{1/2,4(1-\gamma)/\gamma\}+2\epsilon}\big)$$

in (4.5). Since $4/5 < \gamma < 1$ this is still $O(x^{1-\delta})$ with some $\delta > 0$. The remainder of the proof is as in 4.3. □

## References


[1] A. Baker, Transcendental Number Theory, Cambridge University Press, London 1975
[2] M.B. Barban, The “Large Sieve” method and its applications in the theory of numbers, Russian Math. Surveys 2(1966), 49–103
[3] P. Billingsley, Probability and Measure, John Wiley & Sons, New York 1995
[4] Z.I. Borevich, I.R. Shafarevich, Number Theory, Academic Press, New York 1966
[5] D. Bump, S. Friedberg, J. Hoffstein, On some applications of automorphic forms to number theory, Bull. Am. Math. Soc. 33(1996), 157–175
[6] S. Chowla, P. Erdös, A theorem on the distribution of values of L-series, J. Indian Math. Soc. 15(1951), 11–18
[7] H. Cohen, A Course in Computational Algebraic Number Theory, Graduate Texts in Mathematics 138, Springer, Berlin, Heidelberg 1993
[8] J.B. Conrey, D.W. Farmer, Mean values of L-functions and symmetry, Internat. Math. Res. Notices 2000, No. 17, 883–908
[9] H. Davenport, Multiplicative Number Theory, Graduate Texts in Mathematics 74, Springer, New York, Berlin 1980
[10] P.D.T.A. Elliott, On the distribution of the values of quadratic L-series in the half-plane $\sigma > 1/2$, Invent. math. 21(1973), 319–338
[11] P.D.T.A. Elliott, Probabilistic Number Theory II, Grundlehren der mathematischen Wissenschaften 240, Springer, New York, Berlin 1980
[12] C.F. Gauss, Disquisitiones Arithmeticae, Yale University Press, New Haven 1966
[13] D. Goldfeld, C. Viola, Mean values of L-functions associated to Elliptic, Fermat and other curves at the centre of the critical strip, J. Number Th. 11(1979), 305–320
[14] D. Goldfeld, J. Hoffstein, Eisenstein series of 1/2-integral weight and the mean value of real Dirichlet L-series, Invent. math. 80(1985), 185-208
[15] D.R. Heath-Brown, A mean value estimate for real character sums, Acta Arithm. 72(1995), 235–275
[16] H. Iwaniec, Prime geodesic theorem, J. Reine Angew. Math. 349(1984), 136–159
[17] M. Jutila, On character sums and class numbers, J. Number Th. 5(1973), 203–214
[18] M. Jutila, On the mean value of $L(1/2,\chi_d)$ for real characters, Analysis 1(1981), 149–161
[19] J. Kallies, A. Özlük, M. Peter, C. Snyder, On asymptotic properties of a number theoretic function arising out of a spin chain model in statistical mechanics, preprint
[20] P. Kleban, A. Özlük, A Farey fraction spin chain, Commun. Math. Phys. 203(1999), 635-647
[21] W. Luo, P. Sarnak, Quantum ergodicity of eigenfunctions on $\mathrm{PSL}_2(\mathbb{Z})\setminus\mathbb{H}^2$, Inst. Hautes Etudes Sci. Publ. Math. No. 81 (1995), 207–237
[22] W. Luo, Values of symmetric square L-functions at 1, J. Reine Angew. Math. 506(1999), 215–235
[23] H.L. Montgomery, Topics in Multiplicative Number Theory, Lecture Notes in Mathematics 227, Springer, Berlin, Heidelberg 1971

[24] T. Ono, An Introduction to Algebraic Number Theory, Plenum Press, New York 1990
[25] M. Peter, Momente der Klassenzahlen binärer quadratischer Formen mit ganzalgebraischen Koeffizienten, Acta Arithm. 70(1995), 43–77
[26] M. Peter, The limit distribution of a number theoretic function arising from a problem in statistical mechanics, to appear in J. Number Th.
[27] E. Royer, Statistique de la variable aléatoire $L(\mathrm{sym}^2 f, 1)$, preprint
[28] P. Sarnak, Class numbers of indefinite binary quadratic forms, J. Number Th. 15(1982), 229–247
[29] P. Sarnak, Class numbers of indefinite binary quadratic forms II, J. Number Th. 21(1985), 333–346
[30] H. Stark, A complete determination of the complex quadratic fields of class number one, Michigan Math. J. 14(1967), 1–27
[31] E.C. Titchmarsh, The Theory of the Riemann Zeta-Function, Oxford University Press, Oxford 1986
[32] D. Wolke, Moments of class numbers III, J. Number Th. 4(1972), 523–531

Mathematisches Institut, Albert-Ludwigs-Universität, Eckerstr. 1, D-79104 Freiburg
*E-mail address*: peter@arcade.mathematik.uni-freiburg.de