1970

УДК–511

THE DISTRIBUTION OF THE QUADRATIC CLASS NUMBER

P.D.T.A. Elliott

In memory of M. B. Barban

1. Introduction. For each positive integer D which exceeds 1, and which satisfies one of the conditions $-D \equiv 0,1 \pmod 4$, let $h(-D)$ denote the number of classes of primitive binary quadratic forms whose discriminant is $-D$.

For each pair of real numbers z, $x>0$ let $N(x, z)$ denote the number of integers D not exceeding x for which the inequality $h(-D) < \frac{2}{\pi} \sqrt{Dz}$ is satisfied. It was shown by Chowla and Erdös [2] that the limit

$$\lim_{x\to\infty} x^{-1} N(x, z)$$

exists for all real values of z, and is a continous distribution function. Some years later Barban [1] used the inequality of the Large Sieve of U. V. Linnik, and a mehod somewhat different from that of Chowla and Erdös, to compute the moments ot the function $D^{-\frac{1}{2}} h(-D)$. In this way he recovered their result. He proved, moreover, that the characteristic function of this limiting distribution assumed the form

$$\sum_{k=0}^{\infty} \alpha_k \frac{(it)^k}{k!}$$

n a certain range $|t| < t_0$, $t_0 > 0$. In this expression $\alpha_0 = 1$, and

$$\alpha_k = \sum_{\substack{n=1 \\ (n,\ 2)=1}}^{\infty} \frac{\Phi(n)\, \tau_k(n^2)}{n^3}, \qquad (k = 1,\ 2,\ \ldots)$$

where $\Phi(n)$ denotes Euler's totient function, and $\tau_k(m)$ denotes the number o ways of expressing the integer m as the product of k integers.

In the present paper we shall show that the study of the distribution of the values of $D^{-\frac{1}{2}} \cdot h(-D)$ can be reduced to the consideration of sums of independent random variables defined on a finite probability space. The appropriate characteristic functions take on a simple form, so that it proves possible to largely determine the nature of the limiting distribution. In particular we shall prove that it has a probability density, and that it is analytically continuable into a complex half-plane. We shall also measure the rate of convergence of the frequencies $x^{-1} N(x, z)$ to the limiting distribution as $x \to \infty$.

For each value of $x>0$ we write

$$\nu_x\left(D;\ h(-D)<\frac{2}{\pi}\sqrt{D}\,e^z\right)=\left(\frac{1}{2}\ x\right)^{-1}N(x,\ e^z).$$

Theorem. *There is a distribution function $F(z)$ with the following properties*

(i) *The estimate*

$$\nu_x\left(D;\ h(-D)<\frac{2}{\pi}\sqrt{D}\,e^z\right)=F(z)+O\left(\sqrt{\frac{\log\log x}{\log x}}\right),\qquad (x\geqslant 3),$$

holds uniformly for all real values of z.

(ii) *The characteristic function of $F(z)$ takes the form*

$$\frac{1}{4}\left[2+2^{it}+\left(\frac{3}{2}\right)^{it}\right]\prod_{p\geqslant 3}\left(\frac{1}{p}+\frac{1}{2}\left(1-\frac{1}{p}\right)^{1-it}+\frac{1}{2}\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p}\right)^{it}\right)$$

(iii) *The function $F(z)$ has a probability density, and can be analytically continued into a half-plane* $\mathrm{Im}(z)>-c,\ c>0$.

The presence of the untidy factor in the product representation of the characteristic function of $F(z)$ is due to the irregular behaviour of the prime 2, which is characteristic of problems of a quadratic nature. We shall prove that this function is even Riemann integrable over the whole line. The statement in (iii) that $F(z)$ is analytically continuable into a complex half-plane is to be interpreted in the sense that it coincides for all real values of t with a function which is analytic in a half-plane. In particular we recover the assertion of Chowla and Erdös [2] that $F(z)$ is continuous at all finite real points z.

The starting point of all of these investigations is a classical result of Dirichlet. For our purposes this states that for $D>4$,

$$h(-D)=\frac{\sqrt{D}}{\pi}\,L(1,\ \chi_D)$$

where

$$L(1,\ \chi_D)=\sum_{n=1}^{\infty}\chi_D(n)\,n^{-1},\quad \chi_D(n)=\left(\frac{-D}{n}\right),$$

and χ_D denotes a Kronecker symbol. We shall preserve this notation for the duration of this paper.

Lemma 1. *Let $\varepsilon>0$ be given. Then we can find a real number A, depending upon ε, and for each value of $x>3$ a (possibly empty) set $E(x)$ with the following properties*:

(i) *Let D be an integer not exceeding x which does not lie in $E(x)$. Then the approximate relations*

$$L(1,\ \chi_D)=\left\{1+O\left(\frac{1}{\log x}\right)\right\}\prod_{p\leqslant H}\left(1-\chi(p)\,p^{-1}\right)^{-1}$$

hold uniformly for all real numbers H which satisfy $H\geqslant(\log x)^A$, and for all primitive characters $\chi\pmod D$.

(ii) *The number of integers belonging to $E(x)$ is at most $O(x^{\varepsilon})$.*

Proof. This lemma can be proved on exactly similar lines to Theorem 1 of the author's paper [3].

Lemma 2. *Let $a_1, a_2 \ldots$ be a sequence of complex numbers, and let x, H be real number satisfying $x \geqslant 1$, $H \geqslant 3$. Then the Kronecker symbol satisfies the inequality*

$$\sum_{D \leqslant x} \Big| \sum_{n \leqslant H} a_n \left(\frac{-D}{n}\right) \Big|^2 \leqslant x \sum\sum_{\substack{mn = t^2,\ 2t^2 \\ m \leqslant H,\ n \leqslant H}} |a_m a_n| + c_1 H \log H \Big(\sum_{m \leqslant H} |a_m| \Big)^2 .$$

In this inequality we tacitly assume that any integer D under consideration in the left hand sum satisfies $-D \equiv 0,1 \pmod 4$, so that the Kronecker symbol is well defined.

Proof. If we expand the sum on the left hand side of this inequality and invert the order of summation we obtain

$$\sum\sum_{\substack{mn = t^2,\ 2t^2 \\ m \leqslant H,\ n \leqslant H}} a_m \bar{a}_n \sum_{D \leqslant x} \left(\frac{-D}{m}\right)\left(\frac{-D}{n}\right) + \sum\sum_{\substack{mn \neq t^2,\ 2t^2 \\ m \leqslant H,\ n \leqslant H}} a_m \bar{a}_n \sum_{D \leqslant x} \left(\frac{-D}{m}\right)\left(\frac{-D}{n}\right).$$

We can estimate the first of these threefold sums trivially. As to the second, for each value of $m > 0$, there is a number $\varepsilon_m = \pm 1$, so that

$$\left(\frac{-D}{m}\right) = \varepsilon_m \left(\frac{D}{m}\right)$$

for all values of D. Moreover if $m = 2^u m_1$, $2 \nmid m_1$, $n = 2^v n_1$, $2 \nmid n_1$, then

$$\left(\frac{-D}{m}\right)\left(\frac{-D}{n}\right) = \varepsilon_m \varepsilon_n \left(\frac{2}{D}\right)^{u+v} \left(\frac{D}{n_1 n_2}\right).$$

Thus if $''$ denotes summation over the reduced residue classes (mod 8), we can write the innermost sum in the form

$$\varepsilon_m \varepsilon_n \sum_j{}'' \left(\frac{2}{j}\right)^{u+v} \sum_{\substack{D \leqslant x \\ D \equiv j \pmod 8}} \left(\frac{D}{m_1 n_1}\right) + \varepsilon_m \varepsilon_n \sum_{4s \leqslant x} \left(\frac{s}{mn}\right)$$

where the last sum (over s) is to be omitted if mn is even. If $mn \neq t^2, 2t^2$, then $m_1 n_1$ is not an integral square, and when $2 \nmid mn$ neither is mn. All of the symbols in these last expressions are then Jacobi symbols, and non-principal characters (mod $m_1 n_1$), or (mod mn). An appeal to the Polya–Vinogradov inequality therefore shows that these sums are $O(H \log H)$, and lemma 2 follows immediately.

Lemma 3. *Let η, B be real numbers satisfying $B > 0$, $0 < \eta < 1$. For each $x \geqslant 2$ there is a set $G(x)$ possessing the following two properties*

(i) *If D satisfies $D \leqslant x$, $D \notin G(x)$, $-D \equiv 0,1 \pmod 4$, then the Kronecker symbol satisfies*

$$L(1, \chi_D) = \left\{ 1 + O\left(\sqrt{\frac{\log\log x}{\log x}} \right) \right\} \prod_{p \leqslant \eta \log x} \left(1 - \chi_D(p) p^{-1}\right)^{-1}$$

(ii) *The number of integers contained in $G(x)$ is at most $0\left(x (\log x)^{-B}\right)$.*

Proof. We first prove this result under the addition restriction that the integers $-D$ involved only run through the fundamental discriminants. When this is the case the Kronecker symbol becomes a primitive character (mod D). We therefore assume that $G(x)$ contains the set $E(x)$ of lemma 1, defined with $\varepsilon = 1/2$.

Set $U=\eta \log x$, $H=(\log x)^A$, $P^2=F\log x\,(\log\log x)^{-1}$. For each positive integer k we define a sequence of real numbers $a_1, a_2,$ by

$$T_k=\sum_{D\leqslant x}{}' \Big[\sum_{U<p\leqslant H}\chi_D(p)\Big]^{2k}=\sum_{D\leqslant x}{}' \Big[\sum_{m\leqslant H^k} a_m\chi_D(m)\Big]^2$$

where 'indicates summation over those integers D for which $-D$ is a fundamental discriminant. Applying lemma 4 we deduce that

$$T_k\leqslant x\sum\sum_{mn=l^2,\ 2\nmid l} a_m a_n m^{-1}n^{-1}+c_1 H^k\log H^k\Big(\sum_{U<p\leqslant H}\frac{1}{p}\Big)^{2k}.$$

Consider the expression

$$L_k=\sum_{p_1}\cdots\sum_{\substack{U<p_j\leqslant H\\ p_{2k}=l^2}}\frac{1}{p_1}\qquad \frac{1}{p_{2k}}=\sum\sum_{mn=l^2,\ 2\nmid l} a_m a_n m^{-1}n^{-1}.$$

The condition $p_1\ldots p_{2k}=l^2$ ensures that there is a value of j satisfying $1\leqslant j\leqslant 2k-1$ so that $p_j=p_{2k}$. By considering each possibility in turn we see that

$$L_k\leqslant(2k-1)\sum_{U<p\leqslant H}\frac{1}{p^2}\cdot L_{k-1}\leqslant\qquad\leqslant 2^k k!\Big(\sum_{p>U}\frac{1}{p^2}\Big)^k$$

Moreover, an elementary estimate shows that

$$\sum_{U<p\leqslant H}\frac{1}{p}=\log\Big(\frac{\log H}{\log U}\Big)+O\Big(\frac{1}{\log U}\Big)\leqslant c_2,$$

so that

$$T_k\leqslant x2^k k!\Big(\frac{2}{U}\Big)^k+c_1(c_2^2H)^k\log H^k.$$

The number of integers D not exceeding x for which the estimate

$$\Big|\sum_{U<p\leqslant H}\chi_D(p)p^{-1}\Big|>p^{-1}$$

holds is then at most

$$p^{2k}T_k=O\left(x(\log x)^{-B}\right)$$

provided that F and so k is chosen suitably. For the remaining values of D we see that

$$L(1,\chi_D)\prod_{p\leqslant U}\left(1-\chi_D(p)p^{-1}\right)=\left(1+O\Big(\frac{1}{\log x}\Big)\right)\times$$

$$\times\exp\left\{O(p^{-1})+O\Big(\sum_{U<p\leqslant H}\frac{1}{p^2}\Big)\right\}=1+O(p^{-1}).$$

Call the set of integers $D\leqslant x$ which are exceptional in the above sense $J(x)$.

We now use the fact that any integer D can be uniquely represented in the form $D_1^2D_2$, where $\mu^2(D_2)=1$. We define $G(x)$ to consist of all those integers D not exceeding x which satisfy any of the following conditions:

(i) $D\leqslant x^{\frac{3}{4}}$

(ii) D *has more than* $(1+2B)\log\log x$ *distinct prime divisors.*

(iii) There is an integer m satisfying $\frac{1}{2}x^{\frac{1}{4}} \leqslant 2^m \leqslant 4x$ *so that* D_2 *lies in* $J(2^m)$.

It is now an easy exercise to prove that this set satisfies all of the properties required in the lemma.

Construction of the finite probability spaces

Let P, x be real numbers, and let $q_1, q_2,$ be a sequence of k prime numbers which are constrained by

$$2 \leqslant q_1 \leqslant q_2 \leqslant \ldots \leqslant q_k \leqslant P, \qquad q_1 \qquad q_k = R.$$

We form an algebra of sets $\{E\}$ by taking as a typical member the union of residue classes (mod R). If A is any finite set of distinct positive integers, then the collection $\{A \cap E\}$, where $A \cap E$ is to be interpreted as those members of A which belong to any of the residue classes (mod R) which are represented by E is also an algebra. For each class l (mod R), let $A(x, R, l)$ denote the number of members a_i of A which do not exceed x, and which satisfy $a_i \equiv l \pmod{R}$. Let $A(x)$ denote the total number of a_i not exceeding x. Further let there be R numbers $\lambda(l, R)$, so that the asymptotic estimate

$$\sum_{l=1}^{R} |A(x, R, l) - \lambda(l, R) A(x)| = o\left(A(x)\right), \qquad (x \to \infty).$$

This hypothesis is to be interpreted in the sense that the set A, the modulus R and the numbers $\lambda(l, R)$ all may depend upon x.

We define a measure on this last algebra as follows. Let E represent the classes $l_j \pmod{R}$, $(j=1, \ldots, m)$. Then we set

$$\mu(A \cap E) = \sum_{j=1}^{m} \lambda(l_j, R).$$

It is clear that

$$\sum_{l=1}^{R} \lambda(l, R) = 1$$

and that the pair $(\{A \cap E\}, \mu)$ is a finite probability space. Moreover, if we denote by $|B$ the number of integers in the set B, then

$$B| = \{1 + o(1)\}\, \mu B \cdot A(x).$$

We form two models M_1, M_2, by taking A to be

$$\{D;\ D \leqslant x,\ -D \equiv d_j \pmod{4}\}, \qquad (j=1, 2),$$

with $d_1 = 0$, $d_2 = 1$ respectively. In both models we set $P = \frac{1}{2}\log x$. In the first model we set $q_2 = 3$, $q_3 = s$, so that the q_j are the first $\pi\left(\frac{1}{2}\log x\right) - 1$ distinct odd primes. In the second we set $q_1 = 2$, $q_2 = 2$, $q_3 = 2$, $q_4 = 3$, $q_5 = 5$, ..., and so on. Thus the respective values of R are $p_2 \quad p_k$, and $4p_1 \ldots p_k$, where p_j deno-

tes the j^{th} rational prime. In both cases $A(x)=\frac{1}{4}x+O(1)$. For M_1 we set $\lambda(l, R)=R^{-1}$ for each value of l, and for M_2 we set

$$\lambda(l, R)=\begin{cases}\frac{4}{R} & \text{if } l\equiv 1 \text{ or } 5 \pmod 8,\\ 0 & \text{otherwise.}\end{cases}$$

Denote the respective measures so defined on the spaces M_j by μ_j, $(j=1, 2)$. In either model, by virtue of our choice of R we shall have the estimate

$$|B|=\mu B\cdot A(x)+O(x^{\frac{1}{2}}).$$

In these spaces we define independent random variables X_j for $j=2, \ldots, k$, and $j=1, \ldots, k$, respectively, by

$$X_j(D)=-\log\left(1-\left(\frac{D}{q_j}\right)q_j^{-1}\right).$$

Finally, for $i=1, 2$ define

$$\nu_x^i\left(D;\ L(1, \chi_D)<e^z\right)=\frac{4}{x}\{D;\ D\leqslant x, -D\equiv d_i \pmod 4,\ L(1, \chi_D)<e^z\}.$$

Lemma 4. *For any fixed value of B, there is an absolute constant c_3 so that the estimate*

$$\nu_x\left(D;\ h(-D)<\frac{2}{\pi}\sqrt{D}e^z\right)=\frac{1}{2}\sum_{j=1}^{2}\mu_j(x_1+\ldots+x_m<z)+$$

$$+\Theta\sum_{j=1}^{2}\mu_j\left(|x_1+\ldots+x_m-z|\leqslant c_3\sqrt{\frac{\log\log x}{\log x}}\right)+O\left((\log x)^{-B}\right),\quad (|\Theta|\leqslant 1),$$

holds uniformly for all real values of z.

In this result, and for the remainder of the proof of our theorem, the symbol x_1 is to be deleted from the terms involving μ_1.

Proof. In terms of the frequency function used in the statement of Theorem 1 we see that

$$\nu_x\left(D;\ h(-D)<\frac{2}{\pi}\sqrt{D}e^z\right)=\frac{1}{2}\sum_{j=1}^{2}\nu_x^j\left(D;\ L(1, \chi_D)<e^z\right).$$

The present representation theorem now follows from lemma 3.

For each value of j set

$$F_x^j(z)=\mu_j(X_1+\ldots\ X_m<z).$$

We shall consider the space M_1 in detail. The model M_2 can be similarly treated. The random variable x_j (in M_1) has the characteristic function

$$\Phi_j(t)=\frac{1}{q_j}+\frac{1}{2}\left(1-\frac{1}{q_j}\right)^{1-it}+\frac{1}{2}\left(1-\frac{1}{q_j}\right)\left(1+\frac{1}{q_j}\right)^{-it}.$$

There is an absolute constant c_4, so that if $t<c_4 q_j$ is satisfied, then

$$\Phi_j(t)=1+\frac{it}{2q_j^2}-\frac{t^2}{q_j^2}+O\left(\frac{|t|+|t|^3}{q_j^3}\right).$$

We define

$$g_1(t)=\prod_{j=1}^{\infty}\Phi_j(t).$$

It is clear that this infinite product converges uniformly in any bounded region of the complex t-plane, thus $g_1(t)$ is an integral function of t.

Lemma 5. *There exist positive constants* c_r, $r=5, \ldots, 9$, *so that*

(i) $|g_1(t)| \leqslant c_5 e^{-c_6|t|}$

for all real values of t,

$$\text{(ii)} \quad \Big| \prod_{j=1}^{\infty} \Phi_j(t) - g_1(t) \Big| \leqslant c_7 \left[\frac{|t|}{(\log x)^2} + \frac{|t|^3}{\log x} \right] |g_1(t)|$$

for all values of $|t| \leqslant c_8 \sqrt{\log x}$, *for all* $x > c_9$.

Proof. For each integer $j \geqslant 2$ and all sufficiently large values of t which satisfy $|t| \leqslant c_8 q_j$, we obtain the inequality

$$|\operatorname{Re} \log \Phi_j(t)| \leqslant 1 - \frac{t^2}{2q_j^2}.$$

Since for real values of t, the characteristic function $\Phi_j(t)$ satisfies $|\Phi_j(t)| \leqslant 1$, it follows that

$$|g_1(t)| \leqslant \exp\left(\frac{t^2}{2} \sum_{q_j > c_8^{-1}|t|} \frac{1}{q_j^2} \right) \leqslant \exp(-c_5|t|).$$

For absolutely small values of t the inequality (i) follows from the continuity of $g_1(t)$.

The second inequality has an equally simple proof.

Lemma 6. *Let* $F(z)$ *be a non-decreasing function, and* $G(z)$ *a function of bounded variation, whose respective Fourier – Stieltjes transforms are* $f(t)$, *and* $g(t)$, *and which satisfy*

(i) $F(\pm\infty) = G(\pm\infty)$,

(ii) $\int_{-\infty}^{\infty} |F(z) - G(z)|\, dz < \infty$,

(iii) $G'(z)$ *exists for all values of* z, *and* $|G'(z)| \leqslant H$. *Let* T *be a positive real number. Then to every* $k > 1$, *there corresponds a further positive number* $c(k)$, *depending only upon* k, *so that*

$$\sup_z |F(z) - G(z)| \leqslant c(k) \left\{ \frac{H}{T} + \frac{k}{2\pi} \int_{-T}^{T} \left| \frac{f(t) - g(t)}{t} \right| dt \right\}.$$

Proof. A proof of this lemma of Esseen, can be found in Gnedenko and Kolmogorov [5], Chapter 8, § 39, or in Esseen's original paper [4].

Proof of the theorem

Since for any real value of t $g_1(t)$ is the limit of characteristic functions, and is continuous at the origin, it is the characteristic function of a distribution function $G^1(z)$, with the property that $F_x^1(z) \to G^1(z)$ as $x \to \infty$.

We apply lemma 6 with $F(z) = F_x^1(z)$, $G(z) = G^1(z)$, $k = 1$, $T = c_8\sqrt{\log x}$. We see from lemma 5 (i) that $g_1(t)$ is integrable over the whole real line, so that $G^1(x)$ has a probability density,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} g_1(t)\, e^{-itx}\, dt.$$

It follows from this representation that

$$\frac{G^1(z+h)-G^1(z)}{h}=\frac{1}{2\pi}\int_{-\infty}^{\infty} g_1(t)\,e^{-itz}\left(\frac{e^{-ith}-1}{-ith}\right)dt\to\frac{1}{2\pi}\int_{-\infty}^{\infty} g_1(t)\,e^{-itz}\,dt,\ (h\to 0),$$

the operation being justified by the dominated convergence theorem for Lebesgue integrals. Indeed it is clear from this argument that $G^1(z)$ is even analytically continuable into a half plane $\mathrm{Im}(z)>-c$, $c>0$.

We deduce that

$$F_x^1(z)=G^1(z)+O\left(\frac{1}{\sqrt{\log x}}\right),\qquad (x>c_0).$$

In a precisely similar way we can consider the space M_2, and obtain functions $g_2(t)$, and $G^2(z)$. We set $G(z)=\frac{1}{2}\left(G^1(z)+G^2(z)\right)$ so that $G(z)$ has the characteristic function

$$\frac{1}{2}\left(g_1(t)+g_2(t)\right)=\frac{1}{4}\left[2+2^{it}+\left(\frac{3}{2}\right)^{it}\right]\prod_{q_j\geqslant 3}\Phi_j(t),$$

and

$$\frac{1}{2}\sum_{j=1}^{2}\mu_j(X_1+\ldots+X_m<z)=G(z)+O\left(\frac{1}{\sqrt{\log x}}\right).$$

The assertion (i) of the theorem now follows from lemma 4, and the fact that the inequality

$$G(z+h)-G(z)=O(|h|)$$

holds uniformly for all real values of z and h. This completes the proof of the theorem.

University of Nottingham

Received May 18, 1969

References

1. М. Б. Барбан, „Большое решето" Ю. В. Линника и предельная теорема для числа классов идеалов мнимого квадратического поля, Изв. АН СССР, сер. матем., 26, 4 (1962), 573–580.
2. S. Chowla and P. Erdős, A theorem on the distribution of values of L-functions, Journ. Ind. Math. Soc., 15, A (1951), 11–18.
3. P. D. T. A. Elliott, On the size of $L(1,\chi)$, Journ. für die reine und ang. Math., 1969.
4. C. G. Esseen, Fourier analysis of distribution functions, A mathematical study of the Laplace–Gaussian law, Acta Math., 77 (1945), 1–125.
5. B. V. Gnedenko and A. N. Kolmogorov, Limit distributions for sums of independent random variables, Addison Wesley, Reading, Massachusetts (1954).

KVADRATINIŲ FORMŲ KLASIŲ SKAIČIAUS PASISKIRSTYMAS

P. D. T. A. Elliott

(*Reziumė*)

Naudojant tikimybinę interpretaciją, straipsnyje įrodoma, kad atitinkamai normuotas teigiamų binarinių kvadratinių formų klasių skaičius turi tolydinę ribinę pasiskirstymo funkciją. Įvertinamas konvergavimo į ribinę funkciją greitis.

РАСПРЕДЕЛЕНИЕ ЧИСЛА КЛАССОВ КВАДРАТИЧНЫХ ФОРМ

П. Д. Т. А. Эллиотт

(*Резюме*)

Используя вероятностную интерпретацию автор доказывает, что соответственно нормированное число классов положительных бинарных квадратичных форм имеет непрерывную предельную функцию распределения.

Оценивается скорость сходности к предельному закону.