References


MATHMATICS DEPARTMENT
UNIVERSITY OF ARIZONA
TUSCON, ARIZONA

Received on 24.1.1970

On the least integers represented by the genera of binary quadratic forms

by

A. Baker (Cambridge) and A. Schinzel (Warszawa)

In memory of our teacher and friend
Harold Davenport

1. Introduction. It is well known that every class of primitive binary quadratic forms with discriminant \( D \) represents a positive integer \( \leq c(e) |D|^{1/4 + \varepsilon} \) for any \( \varepsilon > 0 \), where \( c(e) \) denotes a number depending only on \( e \).

It seems likely that in fact every genus represents a positive integer \( \leq c(e) |D|^e \) for any \( e > 0 \) but we are unable to establish this conjecture with our present arguments. Also our proof does not allow the number \( c(e) \) mentioned in the theorem to be effectively computed when \( \varepsilon < 1/8 \); an effective estimate, on the other hand, would enable one to determine all the "numerical idonei" of Euler. This result at once from the following:

**Theorem 1.** Every genus of primitive binary quadratic forms with discriminant \( D \) represents a positive integer \( \leq c(e) |D|^e \) for any \( e > 0 \), where \( c(e) \) is effectively computable in terms of \( e \) for any \( e < 1/8 \).

The proof of Theorem 1 involves an argument similar to that used by Ljung and Vinogradov in their paper \( [8] \) on the least prime quadratic residue; thus we shall appeal to the well-known result of Burgess \( [3] \) on character sums and also to Siegel's fundamental theorem \( [12] \) on \( L \)-functions. The proof of Theorem 2 is based on the work of \( [2] \). As demonstrated there, for negative fundamental discriminants \( D \) with class number 1 or 2, where, in the latter case, \( D \) is assumed to be even, the integers referred to in Theorem 1 can be given explicitly and so, by
virtue of Theorem 2, all such \( D \) can be completely determined. For odd
discriminants \( D \) with class number 2 (whence \( D \equiv 5 \pmod{8} \) for \( D < -15 \))
we do not know an explicit construction for the relevant integer represented
by the non-principal genus; if, however, this integer could be determined
then a bound for \(|D|\) would follow at once from an old lemma of Heilbronn [5]
without appeal to Theorem 2 (4). A result related to Theorem 2 has recently
been published by Anferteva and Čudakov [1]. Their theorem allows one to
give a bound in terms of \( c(e) \) for all negative discriminants
with one class in every genus and a prescribed class number.

2. Lemmata. For the proof of Theorem 1 we shall need two lemmata.

**LEMMA 1.** We have

\[
\frac{1}{2\pi i} \int_{c-\infty}^{c+\infty} \frac{s^y}{s(s+1)(s+2)(s+3)} ds = \begin{cases} 
0 & \text{if } 0 < y \leq 1, \\
\frac{1}{6} \left( 1 - \frac{1}{y} \right)^3 & \text{if } y > 1.
\end{cases}
\]

**Proof.** This is a special case given by \( k = 3, c = 2 \) of Theorem B on page 31 of Ingham [6].

**LEMMA 2.** For any non-principal character \( \chi \) with modulus \( k \) and
conductor \( f \) and any \( e > 0 \) we have

\[
(i) \quad \left| \sum_{d \leq \sqrt{x}} \chi(d) d^{-s} \right| \ll k^f f^{1/4} |s| \log x,
\]

\[
(ii) \quad \left| \sum_{d \leq e} \chi(d) d^{-s} \right| \ll k^f f^{1/4} |s|,
\]

\[
(iii) \quad \left| \sum_{d \leq e} \chi(d) d^{-1} \right| \ll k^f f^{1/4} x^{-1/2},
\]

where \( s = \frac{1}{2} + it \) and the constants implied by \( \ll \) depend only on \( e \).

**Proof.** Let \( \chi \) be induced by the character \( \chi_f \) with modulus \( f \); and set

\[
S(M, N) = \sum_{n \leq M} \chi(n).
\]

The sum can be expressed alternatively in the form

\[
\sum_{n \leq M} \frac{X_f(n)}{d(n)} = \sum_{d \leq k} \mu(d) X_f(d) \sum_{M/d < \alpha < M/d} \chi_f(n).
\]

(4) It will be seen from the proof of Theorem 1 that the integers referred to
there can be chosen relatively prime to \( D \).

**A DDED IN PROOF:** The class number 2 problem has recently been resolved
by some new results on linear forms in logarithms; see papers by Baker and Stark
in **Annals of Math.**

By Theorem 2 of Burgess [3] with \( r = 2 \) we have

\[
|\sum_{n=1}^{N} X_f(n)| \ll N^{1/2} f^{3/16 + \varepsilon}
\]

and thus we see that

\[
|S(M, N)| \ll \sum_{d \leq k} \mu(d) N^{1/2} f^{3/16 + \varepsilon} \ll k^f f^{1/4} N^{1/3}.
\]

Similarly from the Pólya–Vinogradov inequality we obtain

\[
|S(M, N)| \ll k^f f^{1/2} \log f \ll k^f f^{1/2}.
\]

Now the identity

\[
\sum_{d=1}^{N} \chi(d) d^{-s} = \sum_{d=M}^{N} \chi(d) d^{-s} = \sum_{d} \sum_{1 \leq m < d} \chi(d/m) d^{-s} = \sum_{d} \sum_{1 \leq m < d} \chi(d/m) d^{-s} + S(M, N) N^{-s}
\]

gives

\[
\sum_{d=M}^{N} \chi(d) d^{-s} \ll |s| N \sum_{d \leq M} |S(M, N)| d^{-s} \ll |s| k^f f^{1/4} N^{-s},
\]

where \( s = \sigma + it \). Applying this with \( M = 0, N = x, \sigma = 1/2 \) and using
(4) we easily verify (i). Further, substituting \( M = x, \sigma = 1 \) and taking
limits as \( N \to \infty \) we obtain (ii). Finally, to prove (ii), we put \( M = f, \sigma = 1/2 \) and appeal to (2); again taking limits as \( N \to \infty \) we get

\[
\left| \sum_{d \leq x} \chi(d) d^{-s} \right| \ll |s| \sum_{f \leq x} k^f f^{1/4} x^{-1/2} \ll |s| k^f f^{1/2},
\]

and (ii) now follows as a consequence of (i) on substituting \( x = f \) in the latter.

3. Proof of Theorem 1. Let \( D = e^2 D_0 \), where \( D_0 \) is a fundamental
discriminant, and let \( \chi_D \) be the principal character mod \( D \). We shall denote
by \( U \) the set of all generic characters for \( D \); the set \( U \) consists of all
Legendre symbols \( \left( \frac{n}{p} \right) \), where \( p \) runs through the odd prime
divisors of \( D \) together with at most two supplementary characters given
by \((1/n), (2/n)\) or their product (4). By \( T \) we shall denote the subset of \( U \n\)
given by the generic characters of \( D_0 \) and, for any subset \( S \) of \( U \), we put

\[
\chi_S = \chi_D \prod_{\chi_S \in S} \chi
\]

by the law of quadratic reciprocity we have then \( \chi_T = (D/n) \).

(4) Cf. Dickson [4], p. 83 or Mathews [9], p. 135.

(4) The generic characters of \( D_0 \) are contained in \( U \) except possibly for the
product of the supplementary characters mentioned above; in the exceptional case,
however, both factors are included in \( U \) and we replace one of them by their product.
Now consider a genus determined by the values \( \varepsilon = \pm 1 \) to be taken by \( \chi \) in \( U \) and thus satisfying
\[
\prod_{\chi \in \mathcal{U}} \varepsilon \chi = 1.
\]

We shall suppose that no positive integer \( \leq x \) and relatively prime to \( D \) is represented by the genus and we shall obtain a contradiction if \( x \gg |D|^{1/2+\varepsilon} \). The supposition implies that for each integer \( n < x \) we have either \( \chi(n) = -\varepsilon \) for some \( \chi \) in \( U \) or, if \( n \) is not represented by any genus, \( \chi_T(p) = -1 \) for some prime \( p \) which occurs in the factorization of \( n \) with odd exponent. On writing
\[
a_n = \sum_{d \mid n} \chi_T(d),
\]
it is easily verified that the second possibility is equivalent to the condition
\[
a_n = 0 \quad (*) \tag{4}
\]
that is
\[
\sum_{n < x} (1-\frac{n}{x})^2 a_n \chi_T(n) \prod_{\chi \in \mathcal{U}} (1+\varepsilon \chi(n)) = 0,
\]
where
\[
E(x, S) = \sum_{n < x} (1-\frac{n}{x})^2 a_n \chi_T(n).
\]

We have
\[
a_n \chi_T(n) = a_n \sum_{d \mid n} \chi_T(d) = \sum_{d \mid n} \chi_T(n/d) \chi_{T-S}(d),
\]
where \( T \sim S \) denotes the symmetric difference, that is the set of elements in precisely one of \( S \) and \( T \), and thus it is clear that
\[
E(x, S) = E(x, T \sim S).
\]
Furthermore we have
\[
E(x, S) = \sum_{d \mid n} \chi_{T-S}(d) \sum_{m \mid n/d} (1-md/x)^3 \chi_S(m)
\]
and, by virtue of Lemma 1, the last sum can be expressed in the form
\[
\sum_{m=1}^{x} \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{6\alpha^e}{s(s+1)(s+2)(s+3)} \chi_S(m) \frac{ds}{md} \, ds.
\]

Hence we obtain
\[
E(x, S) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} F(x, S, s) \, ds,
\]
where
\[
F(x, S, s) = \frac{6\alpha^e}{s(s+1)(s+2)(s+3)} L(s, \chi_T) \sum_{d \mid \infty} \chi_{T-S}(d) \frac{d^{-s}}{d}.
\]

On moving the line of integration from \( \sigma = 2 \) to \( \sigma = \frac{1}{2} \) we get
\[
E(x, S) = \frac{1}{2\pi i} \int_{1-\infty}^{1+\infty} F(x, S, s) \, ds + B(S),
\]

where \( E(S) = 0 \) if \( S \) is not \( \emptyset \), the empty set, and
\[
B(S) = \frac{x}{4} \cdot \frac{\psi(|D|)}{|D|} \sum_{d \mid \infty} \chi_{T}(d) \frac{d}{d}.
\]
In view of Lemma 2 and a property of \( \zeta(s) \) we have, on the line \( \sigma = \frac{1}{2} \),
\[
|L(s, \chi_{T})| \ll |s|^{n} |D|^{1/2},
\]
\[
\begin{align*}
\left| \sum_{d \mid \infty} \chi_{T-S}(d) \frac{d^{-s}}{d} \right| & \ll |s| |D|^{1/2} \log |D| \quad (S \neq T),
\end{align*}
\]
where \( f_{\chi} \) denotes the conductor of \( \chi_S \). It follows that, if \( S \neq T \), the above integral is
\[
\ll x^{1/2} \log |D| \frac{1}{2} (f_{\chi_f} f_{\chi_T})^{1/2}.
\]
Also from Lemma 2 we see that, if \( T \neq \emptyset \),
\[
\sum_{d \mid \infty} \chi_{T}(d) \frac{d}{d} = L(1, \chi_T) + O(|D|^{1/2+\varepsilon} x^{-1/2}),
\]
and obviously the left-hand side is at least 1 if \( T = \emptyset \). Further we note that
\[
|D|^{1/2} (f_{\chi_f} f_{\chi_T})^{1/2}.
\]
But now, on observing that the number of elements in \( U \) is \( \ll 2^{\nu(|D|)} \ll |D|^{1/2} \), where \( \nu(|D|) \) denotes the number of distinct prime divisors of \( |D| \), and recalling also that \( E(x, T) = E(x, \emptyset) \), we readily verify from (3), (4) and (6) that
\[
L(1, \chi_T) \ll x^{-1/2} \log x |D|^{1/8} (f_{\chi_f} f_{\chi_T})^{1/16} + |D|^{3/16+\varepsilon} x^{-1/2}.
\]

(*) Here essential use is made of the four factors \((s+j); s(s+1)\) as occurs in \([8]\) does not suffice.
if $T \neq \emptyset$, and that the same holds if $T = \emptyset$ with $L(1, \chi_T)$ replaced by 1. Since finally $f_{S_0} = f_{S_0}f^2$, where $f$ denotes the conductor of $\mathbb{Z}_S$ with $S'$ defined as the set of elements in $S$ but not in $T$, and $f_{S_0}$ divide $D$, $c$ respectively, we clearly have a contradiction to Siegel's estimate [12] for $L(1, \chi)$ if $x \gg |D|^{3/13+\varepsilon}$. The contradiction proves the theorem.

4. Proof of Theorem 2. Let $k$ be a positive fundamental discriminant relatively prime to $d$ and let $\chi(n) = (k/n)$. By $h(k)$ and $a$ we denote respectively the class number and fundamental unit in the quadratic field $\mathbb{Q}(\sqrt{k})$. Further we set $D = -d$, denote by $h(kd)$ the class number of quadratic forms with discriminant $kd$ and suppose that

$$f = ax^2 + bxy + cy^2$$

runs through a complete set of inequivalent quadratic forms with discriminant $D^*(4)$. We have (see [21])

$$h(k)h(kd)\log a = \frac{1}{\pi} \sum_{n} \frac{\chi(n)}{|a|} \prod_{l \neq p} (1 - p^{-1}) + B_0 + \sum_{r \neq 0} \sum_{n \to \infty} B_n e^{\pi r(n/d)}$$

where

$$B_0 = -\log p \sum_{n \to \infty} \chi(n)$$

if $k$ is the power of a prime $p$, $B_0 = 0$ otherwise and, for $r \neq 0$$^1$

$$|B_r| \leq k^{|r|/e^{-\pi r^2(d/a)}}.$$ 

By Theorem 1 every class represents a positive integer $< c(e) d^{3/13+\varepsilon}$ and, after dividing by a common factor if necessary, we can assume that the representation is proper. Thus we can suppose that $a < c(e) d^{3/13+\varepsilon}$ for every $f$. We have

$$\sum_{r \neq 0} \sum_{n \to \infty} B_n e^{\pi r(n/d)} \leq h(d) \sum_{r \neq 0} \sum_{n \to \infty} e^{\pi r^2(d/a)} = 2h(d) \eta (1 - \eta)^{-1},$$

where $h(d)$ denotes the number of $f$ and $\eta = e^{\pi r^2(d/a)}$. Assuming that $k < d^2$, where $d = (d/e)$, it is easily seen that $\eta < e^{-c_1d^2}$, where $c_1$, like $c_2, c_3, \ldots$, signifies a number effectively computable in terms of

$$(^1) \text{It would suffice to consider fundamental discriminants since by a result of Grube 1874, recently rediscovered by Grosswald (see Acta 8, pp. 205-308) any nonfundamental discriminant $D < -315$ with one class in every genus differs from a fundamental discriminant only by a factor 4.}$$

$$(^2) \text{Though it is unimportant, a factor 2 which mistakenly entered in the work of [2] via the definition of $A_r$, has been omitted here.}$$

If $d > c_3$ then $\eta < 1/2$ and, recalling that $h(d)$ is also the number of genera, it follows that the expression on the right of (7) is at most

$$8h(d) \eta \leq 4^{16} d^{3/13+\varepsilon}.\eta < e^{-c_2d^2}.$$ 

Now there exist four distinct primes $p, q, q', p''$, say, not dividing $d$ and less than $d^{1/13+\varepsilon}$, for there are $< d^{1/13+\varepsilon}$ prime divisors of $d$ and Chebychev's estimates, for instance, show that there are $> d^{1/13+\varepsilon}$ primes below $d^{1/13+\varepsilon}$. Let $k = pp', k = pp''$ or $q = q''$ so that, for $k = k_1$ or $k_2$ we have $k \equiv 1 \mod 4$, $(k, d) = 1$ and $B_0 = 0$. We introduce the suffix $j = 1$ or $2$ to distinguish the quantities defined above corresponding to $k_1$ or $k_2$. We write

$$b_1 = h(k_1)h(k_1d)k_2 P_1 \sum_{a} \chi_3(a)/a;$$

where

$$P_1 = \prod_{p \neq 2} p, \quad Q_2 = \prod_{p \neq 2} (p^2 - 1),$$

and we define $b_2$ similarly by interchanging the suffixes 1 and 2. Then clearly $b_1, b_2$ are rational integers and from (6) and (7) we obtain

$$|b_1| < c_1 d^{1/13+\varepsilon} < 2dH e^{-c_2d^2},$$

where $H$ is some number exceeding the maximum of $|b_1|$, $|b_2|$. Further, if $d > c_4$, we have

$$\sum_{a} \chi_3(a)/a \leq (\sqrt{2}d)^{60} = d^{60}$$

and, for $\chi = \chi_1$ or $\chi_2$,

$$|\sum_{a} \chi_3(a)/a| \leq \sum_{a} |a| \leq h(d) < d.$$ 

Since also $h(k) < d^2$, $h(kd) < d^2$, $P_1 < d^{60}$, $Q_2 < d^{60}$ we see that a possible value for $H$ is $d^{61}$. This gives

$$|b_1| < c_1 d^{1/13+\varepsilon} < 2dH e^{-c_2d^2}.$$ 

On the other hand, by Theorem 2 of [10], a version of well-known theorem of Gelfond in which the dependence on the logarithms is specified explicitly, we have

$$|b_1| < c_1 d^{1/13+\varepsilon} < 2dH e^{-c_2d^2}$$

providing that $\log H > a_0 + 1 + a_0$, where

$$a_0 = \log \max \{aD_0^{1/4}, 1, \bar{a}_1, \bar{a}_2, \bar{a}_3, \bar{a}_4\}.$$ 

$$(^3) \text{To apply the results of [10], note that $\bar{a}_4 = \bar{a}_1^{\varepsilon}$, where $\varepsilon = b_1 \log a_1 - b_2 \log a_2$, and that $a_1, a_2$ are multiplicative independent.}$$
and $D_0$ is the discriminant of the field $\mathbb{Q}(\sqrt{k_1}, \sqrt{k_2})$. Furthermore, the classical estimate of fundamental units in real quadratic fields (cf. Schur [11]) gives

$$|a_j| \leq b_j^{1/2} \quad (j = 1, 2),$$

and it is readily verified that

$$|D_0|^{1/2} \leq 64k_1k_2|k_1 - k_2|.$$  

Thus we obtain $a < c_1x^{1/4}\log d$, whence $\log H = c_1^{100\varepsilon}$ satisfies the above condition.

It is clear that the inequalities are consistent only if $d < a$, and this completes the proof of Theorem 2.

We conclude by expressing our thanks to Professor Heilbronn for pointing out a mistake in an earlier draft of this paper.

References


Received on 25. 1. 1970

Subgroups of the modular group generated by parabolic elements of constant amplitude

by

R. A. Rankin (Glasgow)

Dedicated to the memory of Harold Davenport

I. Let $\Gamma(1) = \text{SL}(2, \mathbb{Z})$ denote the modular group of unimodular $2 \times 2$ matrices with rational integral entries and put

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$  

For each positive integer $n$ we write $\Delta(n)$ for the normal closure of $U^n$; it is the subgroup of $\Gamma(1)$ generated by the conjugate parabolic matrices $L_n^*U^nL_n\Gamma(1))$. Further, $\Delta(n)$ is a subgroup of the principal congruence group

$$\Gamma(n) = \{ T \in \Gamma(1) : T \equiv I (\text{mod} n) \}.$$  

The purpose of this paper is to provide a new proof of the well known

**Theorem.** (i) For $1 \leq n \leq 5$, $\Delta(n) = \Gamma(n)$. (ii) For $n \geq 6$, the index $[\Gamma(n) : \Delta(n)] = \infty$.

If $\mathcal{G}$ is any subgroup of $\Gamma(1)$, we denote by $\hat{\mathcal{G}}$ the corresponding inhomogeneous group. Thus $\Gamma(1) = \text{SL}(2, \mathbb{Z})$ and $\hat{\mathcal{G}}$ is the image of $\mathcal{G}$ under the natural mapping from $\Gamma(1)$ to $\Gamma(1)/\Delta$, where $\Delta = \{ I, -I \}$ is the centre of $\Gamma(1)$. A corresponding theorem holds for $\Delta(n)$ and $\hat{\Gamma}(n)$.

As pointed out by Knopp [5], who gave an independent proof of part (ii) of the theorem, the results stated can be excavated by the persevering reader from the first volume of Klein and Fricke's monumental treatise [14], pp. 354-360. An alternative proof of part (i) of the theorem has been given by Brenner [11], pp. 215-217. The quickest and most elegant proof of the theorem is obtained by using the canonical presentation of a Fuchsian group; see Wohlfahrt [10]. For an application of properties of tesselation groups to prove part (ii), see Menonick [6].

The proof of part (i) given below uses elementary properties of indefinite binary quadratic forms. It is, perhaps, worth pointing out that