Class Numbers of Indefinite Binary Quadratic Forms*

Peter Sarnak

Courant Institute of Mathematical Sciences, New York University,
251 Mercer Street, New York, New York 10012

Communicated by D. Zagier

Received February 17, 1981

We determine the asymptotic average sizes of the class numbers of indefinite binary quadratic forms when ordered by the sizes of their corresponding fundamental units. The proofs make use of the Selberg trace formula.

0. Introduction

Denote by $\mathcal{D}$ the set of positive ring discriminants, that is, $\{d > 0: d \equiv 0, 1 \pmod{4}, d$ not a square}. To each such $d$ let $h(d)$ denote the number of inequivalent primitive binary quadratic forms of discriminant $d$, and let $(x_d, y_d)$ be the fundamental solution of the Pellian equation

$$x^2 - dy^2 = 4.$$ 

Let

$$\epsilon_d = \frac{x_d + \sqrt{d} y_d}{2},$$

and let $D_p = \{d \in D: p | y_d\}$, where $p \geq 3$ is a prime, or $p = 1$ when $D_1 = D$.

The average behavior of $h(d)$ was noticed by Gauss [2] to be rather erratic, however, he also noticed that

$$\sum_{d \in D, d \leq x} h(d) \log \epsilon_d = \frac{\pi^2 x^{3/2}}{18 \zeta(3)} + O(x \log x).$$

This was confirmed in Siegel [11].

It seems difficult, however, to separate the quantities $h(d)$ and $\log \epsilon_d$ and understanding the function $d \rightarrow h(d)$ remains one of the major outstanding

* The results of this paper constitute part of the author's thesis, Stanford, 1980.
problems in this theory. In this paper we will develop the asymptotic average behavior of $h(d)$ over the exhausting family of sets,

$$D_{p,x} = \{d \in D_p : \varepsilon_d \leq x \} \quad \text{as} \quad x \to \infty.$$ 

The main result is

**Theorem 4.11.** Let $\gamma > \frac{5}{4}$. Then

$$\frac{1}{|D_{p,x}|} \sum_{d \in D_{p,x}} h(d) = \frac{16}{35} c_p \frac{Li(x^2)}{x} + O(x^{\gamma}),$$

where

$$c_p = \begin{cases} 1 & \text{if } p = 1, \\ \frac{1 + p^2}{p(p^2 - 1)} & \text{if } p \geq 3, \end{cases}$$

and $Li(u) = \int_2^u \frac{1}{\log t} \, dt$ as in the ordinary prime number theorem. ($|A|$ denotes the cardinality of the set $A$.)

The theorem shows that in this ordering the size of $h(d)$ is about the size of $h(d)$ over the sets $D_{p,x}$. In Proposition 4.1 the asymptotics for the sizes for the sets $D_{p,x}$ are derived, indeed all of Section 4 is devoted to this.

In deriving the above formula we develop some other asymptotic formulas of interest. In Theorems 3.1 and 3.4 we develop the asymptotics of the sums of $h(d)$ over the sets $D_{p,x}$. In Proposition 4.1 the asymptotics for the sizes for the sets $D_{p,x}$ are derived, indeed all of Section 4 is devoted to this.

Section 1 deals with the connection between the quantities $h(d)$, $\log \varepsilon_d$ and lengths of closed geodesics on the Riemann surfaces $H/\Gamma(p)$, where $H$ is the Lobachevskii plane with its non-Euclidian metric, and $\Gamma(p)$ is the principal congruence subgroup of $PSL(2, \mathbb{Z})$ of level $p$.

In Section 2 we discuss the asymptotic behavior of these lengths via the Selberg trace formula. We mention that the occurrence of class numbers of forms in a different connection with the Selberg trace formula can be found in Selberg’s original paper (Selberg [9]), where these came up in his evaluation of the trace of the Hecke operators.

1. **Quadratic Forms and Arithmetic Subgroups**

The forms we consider should be primitive and indefinite, that is to say, if

$$Q(x, y) = ax^2 + bxy + cy^2,$$
then \((a, b, c) = 1\), and \(b^2 - 4ac = d > 0\). We will use \([a, b, c]\) to denote such a form. Two forms \([a, b, c]\) and \([a', b', c']\) are called equivalent (in the narrow sense) if there is a unimodular transformation \(\gamma\) such that
\[
\begin{pmatrix}
a' & b'/2 \\
b'/2 & c'
\end{pmatrix} = \gamma
\begin{pmatrix}
a & b/2 \\
b/2 & c
\end{pmatrix} \gamma^{-1},
\]
such an equivalence will be denoted by \([a, b, c] \sim [a', b', c']\). This relation partitions our forms into classes, and two forms from the same class clearly have the same discriminant. As was shown by Gauss the number of classes of a given discriminant \(d\) is finite, this number is the class number and is denoted \(h(d)\).

1.1. Automorphs

Let \(Q = [a, b, c]\) be a primitive form. The subgroup of the modular group \(\Gamma = SL(2, \mathbb{Z})\) which fixes \(Q\) (under the action above) is called the group of automorphs of \(Q\). It is well known (see Landau [5]) that this group consists of matrices of the form
\[
\begin{pmatrix}
t - bu & -cu \\
a & t + bu
\end{pmatrix}
\]
where \(t^2 - du^2 = 4\).

The group is infinite and cyclic with generator
\[
M = M_{[a, b, c]} = \begin{pmatrix}
to - buo & -cuo \\
auo & to + buo
\end{pmatrix},
\]
where \(t_o, u_o > 0\) is the fundamental solution of Pell's equation \(x^2 - dy^2 = 4\). Notice that we choose the generator to have positive trace.

Associated with the primitive form \(Q = [a, b, c]\) are the two real quadratic numbers, \(\theta_1 = (-b + \sqrt{d})/2a\) and \(\theta_2 = (-b - \sqrt{d})/2a\), which are the roots of \(a\theta^2 + b\theta + c = 0\). \(\theta_1 > \theta_2\) if \(a > 0\), and \(\theta_2 > \theta_1\) if \(a < 0\). This gives a correspondence between ordered pairs of real quadratic conjugates and primitive forms.

Furthermore, if \(\gamma = [a, b, c] \in \Gamma\) sends \([a, b, c]\) to \([a', b', c']\), then \(\gamma^{-1} = [c, -b, a]\) sends \((\theta_1, \theta_2)\) to \((\theta'_1, \theta'_2)\) under the linear fractional action \(z \rightarrow (dz - b)/(cz + a)\). From this it is clear that the group of automorphs of \([a, b, c]\) (or of \([-a, -b, -c]\)) is the stabilizer of \(\theta_1\) (and so also of \(\theta_2\)) in \(\Gamma\).
1.3. Primitive Hyperbolic Transformations

Let \( \Delta \) be a discrete subgroup of \( \text{PSL}(2, \mathbb{R}) \). An element \( \delta \in \Delta \) is called hyperbolic if as a linear fractional transformation its fixed points are real and distinct. Such a transformation can be brought into the form \((\begin{array}{cc} 1 & \delta \\ 0 & 1 \end{array})\), \( i > 1 \), by conjugation by an element of \( \text{PSL}(2, \mathbb{R}) \). Clearly \( \text{Trace}(\gamma) = t + t^{-1} \) and we call the number \( t^2 \) the norm of \( \gamma \) denoted \( N(\gamma) \). We say that \( \delta \) is primitive if it is never a nontrivial power of another element of \( \Delta \).

If we let \( H \) denote the upper half plane with its non-Euclidian geometry, so that \( \text{PSL}(2, \mathbb{R}) \) acts as isometries of \( H \), then a primitive hyperbolic \( \delta \in \Delta \) will take the geodesic joining the two fixed points of \( \delta \) back to itself. It will give rise to a closed geodesic of length \( \log N(\gamma) \) on the quotient \( H/\Delta \). Further, conjugates of \( \delta \) in \( \Delta \) gives rise to the same closed geodesics on the quotient \( H/\Delta \). All closed geodesics on \( H/\Delta \) are accounted for in this way. This gives a nice geometric interpretation of the numbers \( \log N(\gamma) \) as \( \gamma \) varies over conjugacy classes of primitive hyperbolic transformations of \( \Delta \).

Returning to 1.2 we see that \( M_{[a,b,c]} \) is hyperbolic and its norm is \( \varepsilon_2^2 \). It is also clearly primitive since \( M \) is the generator of the stabilizer of \( \theta_1 \).

**Proposition 1.4.** Define the map \( \phi \) by \( \phi([a, b, c]) = M_{[a, b, c]} \); then

(i) \( \phi \) is a one-to-one map of primitive quadratic forms onto primitive hyperbolic elements of \( \Gamma \);

(ii) \( \phi \) commutes with the action of \( \text{PSL}(2, \mathbb{Z}) \) so that \([a, b, c] \sim [a', b', c']\) iff \( M_{[a,b,c]} \) is conjugate to \( M_{[a',b',c']} \).

This proposition tells us that the primitive hyperbolic elements of the modular group are exactly the fundamental automorphs of indefinite primitive quadratic forms (conjugacy classes correspond to classes of forms).

**Proof.** Let \( P \) be the primitive hyperbolic transformation \((\begin{array}{cc} a & b \\ c & d \end{array})\). \( P \) has two real fixed points \( \theta_1, \theta_2 \) which are roots of \( \gamma \theta^2 + (\delta - \alpha) \theta - \beta = 0 \). Let \( \Gamma_{\theta_i} \) be the stabilizer of \( \theta_i \) in \( \Gamma \). \( \Gamma_{\theta_i} \) is generated by \( P \) since \( P \) is primitive. By an earlier remark \((\theta_1, \theta_2) \) corresponds to a primitive form \([a, b, c]\), either \( M_{[a,b,c]} = P \) or \( P^{-1} \). We see from this that \( \phi \) is surjective.

However, if \( M = M' \) with \( M = M_{[a,b,c]} \) and \( M' = M_{[a',b',c']} \), then \( M \) and \( M' \) have the same fixed points, so that \((\theta_1, \theta_2) = (\theta'_1, \theta'_2) \) or \((\theta'_1, \theta'_2) \). Now \( M = (a^* b^*) - M' = (a'^* b'^*) \) so that \( a - a' \), \((\theta_1, \theta_2) = (\theta'_1, \theta'_2) \) and \( [a, b, c] \sim [a', b', c'] \). One easily verifies that if \( \gamma \in \Gamma \) is such that \([a, b, c] \sim [a', b', c']\), then

\[
\gamma^{-1} M_{[a,b,c]} \gamma = M_{[a',b',c']}. 
\]

We have established the following geometric realization of the numbers \( h(d) \) and \( \log \varepsilon_d \).
COROLLARY 1.5. The norms of the conjugacy classes of primitive hyperbolic transformations of $\Gamma$ are $\varepsilon_d^2$ where $d \in C$, with multiplicity $h(d)$. Or, put another way, the lengths of the closed geodesics on $H/\Gamma$ are the numbers $2 \log \varepsilon_d$ with multiplicity $h(d)$.

We now make a slight detour to discuss the behavior of the lengths of the closed geodesics on Riemann surfaces carrying the Poincaré metric.

2. PRIME GEODESIC THEOREMS

Let $M$ be a complete Riemannian manifold. We let $CP$ denote the set of oriented closed geodesics on $M$. For $\gamma \in CP$ we let $\tau(\gamma)$ denote the length of $\gamma$ and for $x \geq 0$ define $\pi(x)$ by

$$\pi(x) = \#\{\gamma \in CP : \tau(\gamma) \leq x\}.$$ 

By a prime geodesic theorem we mean an asymptotic formula for $\pi(x)$ as $x \to \infty$. It is remarkable that for compact manifolds of negative sectional curvature such a formula exists, see Margulis [6]. One of the applications of the Selberg trace formula is such prime geodesic theorems, with sharp error terms, for manifolds which are quotients of various symmetric spaces.

As is apparent from Section 1 we would like to use such prime geodesic theorems for manifolds $H/\Delta$, which are of finite area. We should mention that $H/\Gamma$ is not smooth since $\Gamma$ has transformations with fixed points, however, in proving such theorems by use of the Selberg trace formula one works with the primitive hyperbolic transformations as described earlier, so that there is never any problem with elliptic transformations (i.e., those with fixed points in $H$).

In Sarnak and Woo [8] we give a simple direct proof of the following Theorem 2.1 for noncompact finite volume surfaces. The corresponding theorem for compact surfaces appears in various places, e.g., Huber [4] and Hejhal [3]. The possibility of proving such a theorem by using his $\zeta$ function was indicated by Selberg in [9].

Let $M = H/\Gamma_1$, where $\Gamma_1$ is a discrete subgroup of $PSL(2, \mathbb{R})$, which acts discontinuously on $H$ and for which $M$ has finite area. Let $\Delta$ be the Laplace–Beltrami operator on $M$, and let $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k \leq 3/16$ be the first few discrete eigenvalues of $\Delta$.

**Theorem 2.1.** With the above notation, set $it_1 = \sqrt{\lambda_1 - \frac{1}{4}}, \ldots, it_k = \sqrt{\lambda_k - \frac{1}{4}}$, so that $\frac{1}{2} \leq t_k < \frac{1}{2}$; then

$$\pi(x) = Li(e^x) + Li(e^{(1/2 + t_1)x}) \ldots + Li(e^{(1/2 + t_k)x}) + O(e^{3/4x}x^2)$$

as $x \to \infty$. 


Before returning to our number theoretic considerations we develop an interesting analogy with algebraic number theory. This will be needed later for our discussion of the surfaces \( y_p = H/\Gamma(p) \), \( \Gamma(p) \) the congruence subgroup of \( \Gamma \) of level \( p \).

2.2.

Let \( S \) and \( W \) be two Riemann surfaces of hyperbolic type, with \( W \) a finite regular cover of \( S \). Thus the group of cover transformations \( G \) is transitive. We may realize \( W \) and \( S \) as \( H/\Gamma(W) \) and \( H/\Gamma(S) \), respectively, where \( \Gamma(W) \) and \( \Gamma(S) \) are discrete subgroups of \( \text{PSL}(2, \mathbb{R}) \). \( \Gamma(W) \) is then a normal subgroup of \( \Gamma(S) \) with \( G \cong \Gamma(S)/\Gamma(W) \).

By abuse of language, we will call an oriented primitive closed geodesic on any of the surfaces (relative to the Poincaré metric) a prime. Relative to the metric, \( G \) acts a group of isometries while the natural projection \( \pi: W \to S \) is a local isometry.

A prime \( \gamma \) in \( W \) is said to lie over a prime \( \gamma \) in \( S \), if \( \pi(\gamma) = \gamma \) (or we may say \( \gamma \) lifts to \( \gamma \)). In such a case it is clear that \( \tau(\gamma) = m\tau(\gamma) \) for some \( m \in \mathbb{Z} \).

Above each prime \( \gamma \) in \( S \) are a finite number of primes \( \gamma_1, \gamma_2, \ldots, \gamma_k \) (which are disjoint) in \( W \). Call their lengths \( m_1, m_2, \ldots, m_k \), respectively. On the assumption that \( G \) acts transitively on the sheets one sees that \( m_1 = m_2 = \cdots = m \), say, and \( mk = |G| = n \), where \( n \) is the order of \( G \). We will say that \( \gamma \) splits completely if above \( \gamma \) lie \( n \) distinct primes (all will then be of the same length as \( \gamma \)). The situation is then analogous to that of the splitting of prime ideals in an extension of a number field. In fact, our group \( G \) will correspond to the Galois group. We wish to develop an analogue of the Chabotarev density theorem.

We first translate the above lifting ideas to the group \( \Gamma(S) \) and \( \Gamma(W) \). Each closed geodesic \( \gamma \) is associated to a conjugacy class \( \{P_\gamma\} \). With each \( \{P_\gamma\} \) we associate an integer which is the order of \( P_\gamma \) (or any of its conjugates) in \( G = \Gamma(S)/\Gamma(W) \). Now \( \{P_\gamma^m\} \) is a subset of \( \Gamma(W) \) and splits into a disjoint union of conjugacy classes in \( \Gamma(W) \).

**Proposition 2.3.** For \( z = P_\gamma^m \in \Gamma(W) \), the number of conjugacy classes into which \( \{Z\}_S \) splits is \( k \), where \( k = n/m \). These classes correspond to the previous \( \gamma_1, \gamma_2, \ldots, \gamma_k \) of lengths \( m\tau(\gamma) \).

**Proof.** First, we observe the following. For \( z \in \Gamma(W) \) the number of classes into which \( \{Z\}_S \) splits is the number of double cosets

\[ c(z)\Gamma(S)/\Gamma(W), \]

where \( c(z) \) is the centralizer of \( z \) in \( \Gamma(S) \). To see this, the double coset relation is

\[ t \sim t' \quad \text{iff} \quad t' = htk, \quad h \in c(z), \quad k \in \Gamma(W). \]
So we define \( t \equiv t' \) iff \( t^{-1}zt \) and \( (t')^{-1}zt' \) are conjugate in \( \Gamma(W) \), which is true

\[
\text{iff } \mu^{-1}t^{-1}zt\mu = (t')^{-1}zt', \quad \mu \in \Gamma(W),
\]

\[
\text{iff } t'\mu^{-1}t^{-1} \in C(z),
\]

\[
\text{iff } t' = h\mu \quad \text{with} \quad h \in C(z), \mu \in \Gamma(W),
\]

which proves the claim.

To complete the proof of the proposition we use the double coset counting identity

\[
[(S): \Gamma(W)] = \sum_{i=1}^{k} [C(z): C(z) \cap t_i \Gamma(W) t_i^{-1}],
\]

where

\[
\Gamma(S) = \bigcup_{i=1}^{k} C(z) t_i \Gamma(W)
\]

is the double coset decomposition.

However, \( \Gamma(W) \) is a normal subgroup of \( \Gamma(S) \), and so if \( z = p^m \) so that \( C(z) = \langle P \rangle \) (i.e., \( C(z) \) is generated by \( P \)) then

\[
n = \sum_{i=1}^{k} [C(z): C(z) \cap \Gamma(W)]
\]

\[
= \sum_{i=1}^{k} [G\langle P \rangle: \langle P^m \rangle] = mk.
\]

We now associate to each prime \( \gamma \) in \( S \) a conjugacy class \( B_\gamma \) of \( G \). (One can use the cover transformation that \( \gamma \) induces to do so, or proceed as we do here.) The map \( \{z\}_S \mapsto \{z\Gamma(W)\}_G \) is well defined. Define \( B_\gamma \) by

\[
\gamma \mapsto \{P_\gamma \Gamma(W)\}_G.
\]

\( B_\gamma \) tells us how \( \gamma \) splits in \( W \), and in particular tells us the numbers \( k \) and \( m \). Notice that \( \gamma \) splits completely iff \( B_\gamma = \\{\text{identity}\} \), iff \( P_\gamma \in \Gamma(W) \).

Let \( C \) be a fixed conjugacy class of \( G \), define

\[
\pi_C(x) = \sum_{\gamma \in C P, \left|\tau(\gamma)\right| \leq x} 1
\]

which counts the number of primes whose length is less than \( x \), which split in a certain way.
One may use the Selberg trace formula as it applies for vector functions and unitary finite dimensional representations (see Selberg [9]) to prove the following analogue of the Dirichlet–Chabotarev density theorem.

**Theorem 2.4.** Suppose that $\Delta$ on $W$ has no eigenvalues in $(0, 3/16)$ (the formula is modified in a fashion similar to Theorem 2.1 in the case that such eigenvalues do occur). Then

$$\pi_c(x) = \frac{|C|}{|G|} Li(e^x) + O(e^{(1/4)x}x^2).$$

The proof of this theorem may be found in Sarnak [7]. However, in this paper we will only have occasion to use the above for the case of $C = \{\text{identity}\}$. In this situation Theorem 2.4 is easily proved as follows. To count primes in $W$ one counts the primes in $S$ which split at various levels; however, it is easy to see that the number of those which have length less than $x$ and do not split completely is $O(e^{x/2})$. So that

$$\pi_w(x) = |G| \pi_{\{\text{id}\}}(x) + O(e^{x/2}).$$

Now use the asymptotics of lengths for $W$, from Theorem 2.1.

3.

We now return to the arithmetic groups $\Gamma$ and more generally to

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma: \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N \right\}.$$

We will need to know that for such groups there are no eigenvalues of $\Delta$ in $(0, 3/16]$. This was proved by Selberg [10] by remarkably connecting this question with estimates on Kloosterman sums which are known to be true. It is as yet an unsolved conjecture of Selberg that for these groups, $\Delta$ has no eigenvalues in $(0, 1/4)$. (It is not difficult to show that the latter holds for small $N$.)

Applying the prime geodesic Theorem 2.1 to the surface $H/\Gamma$, together with the realization of the lengths as in Cor. 1.5, we have

**Theorem 3.1.**

$$\sum_{\{d \in D: e_d < x\}} h(d) = Li(x^2) + O(x^{3/2}(\log x)^2) \quad \text{as} \quad x \to \infty.$$
Proof. Theorem 2.1 gives

\[ \sum_{d \in D: \log e_d \leq x} h(d) = \text{Li}(e^x) + O(e^{(3/4) x x^2}). \]

To compare this with Siegel's result mentioned in the introduction, we may easily sum the last by parts to obtain

**Corollary 3.2.**

\[ \sum_{d \in D: \log e_d \leq x} h(d) \log e_d = \frac{x^2}{2} + O(x^{3/2} (\log x)^3). \]

Our aim, however, has been to obtain averages of \( h(d) \) alone. To obtain the averages rather than sums of Theorem 3.1 we will develop the asymptotics of \( |\{d \in D: e_d \leq x\}| \) in Section 4.

Now let \( S = H/\Gamma \) and \( W = H/\Gamma(P) \); then by 2.4 we know the asymptotics \( \pi_{\{d\}}(x) \) for the primes which split completely. The following will allow us to express everything solely in number theoretic quantities.

**Proposition 3.3.** Let \( p \geq 3 \) be prime. A fundamental automorph \( V \) of a primitive quadratic form of discriminant \( d \) lies in \( \Gamma(p) \) iff \( d \in D_p \) (see introduction for the set \( D_p \)).

**Proof.**

\[
V = \left[ \begin{array}{cc} \frac{t - bu}{2} & au \\ -cu & \frac{t + bu}{2} \end{array} \right],
\]

\[ V = [a, b, c], \quad i^2 - du^2 = 4. \]

\( V \equiv \pm I \pmod p \) iff either

\[
au \equiv 0 \pmod p, \quad cu \equiv 0, \quad \frac{t - bu}{2} \equiv 1, \quad \frac{t + bu}{2} \equiv 1
\]

or

\[
au \equiv 0, \quad cu \equiv 0, \quad \frac{t - bu}{2} \equiv -1, \quad \frac{t + bu}{2} \equiv -1.
\]

In either case one has

\[ au \equiv 0, \quad bu \equiv 0, \quad cu \equiv 0; \]

since \( V \) is primitive, it follows that \( p \mid u \) and so \( d \in D_p \).
Conversely, if \( p \mid u \) (i.e., if \( d \in D_p \)) and \( V \) is primitive of discriminant \( d \), then let
\[
\frac{t - bu}{2} = m, \quad \frac{t + bu}{2} = n \quad (t = m + n, 0 = m - n),
\]
and
\[
t^2 = 4 \pmod{p}, \quad t = 2 \pmod{p},
\]
therefore
\[
2m \equiv 2(p), \quad \text{or} \quad 2m \equiv -2(p),
\]
therefore
\[
m \equiv \pm 1(p), \quad n \equiv +1(p), \quad \Rightarrow V \in \Gamma(p).
\]
Thus we either have all classes of a given discriminant splitting completely, or none do.

Applying Theorem 2.4 we obtain the following sums of \( h(d) \) over the sets \( D_p \).

\textbf{Theorem 3.4.}

\[
\sum_{\{d: d \in D_p, \nu_d < x\}} h(d) = \frac{2}{p(p^2 - 1)} Li(x^2) + O(x^{3/2}(\log x)^2) \quad \text{as} \quad x \to \infty.
\]

\textbf{Proof.}

\[
\pi_{id}(x) = \sum_{\{d: d \in D_p, 2 \log \nu_d < x\}} h(d) = \frac{1}{|G|} Li(e^x) + O(e^{1/4} x^{4/3}).
\]

Now \( G \cong \text{PSL}(2, \mathbb{F}/p\mathbb{F}) \), from which
\[
|G| = (p^2 - 1)p/2
\]
follows.

4.

To complete our analysis of the previous formulas, we need to know the
sizes of the sets $D_p(x) = \{ d \in D_p : \varepsilon_d \leq x \}$. It should be noted that $\varepsilon_d$ itself is rather irregular, and not much more than $\sqrt{d} \leq \varepsilon_d \leq \varepsilon_d^d$ is known.

**Proposition 4.1.** For $\gamma > 2/3$ we have

$$|D_p(x)| = \frac{35}{8(1 + p^2)} x + O(x^\gamma), \quad p = 1 \text{ or a prime.}$$

We will give the proof in detail when $p = 1$, and indicate the modifications needed to do the general case.

We begin by reducing the problem to one of counting solutions to a diophantine inequality.

**Lemma 4.2.** Let $N(x) = \text{the number of solutions of } m^2 - n^2d = 4 \text{ with } m, n > 0, d \in D, \text{ and } m \leq x$. Then

$$N(x) = |D_1(x)| + O(x^{1/2}).$$

**Proof:** For $d \in D$ and $k \geq 1$ write

$$\varepsilon_d^k = \frac{x_{d,k} + \sqrt{d}y_{d,k}}{2}, \quad x_{d,k}, y_{d,k} \in \mathbb{Z}.$$  

So

$$x_{d,k}^2 - dy_{d,k}^2 = 4.$$  

Now let $\psi_1(x) = |\{(d, k): \varepsilon_d^k \leq x\}|$; and let $\psi(x) = |D_1(x)|$. It is clear that

$$\psi_1(x) = \psi(x) + \psi(x^{1/2}) + \psi(x^{1/3})...,$$

the series being finite. We will later see that $\psi_1(x) = O(x)$. Combining this with $\psi(x^{1/k}) = 0$ if $x^{1/k} < 2$ or $k > \log x / \log 2$, we have

$$\psi(x^{1/3}) + \psi(x^{1/4}) ... = O(x^{1/3} \log x),$$

and therefore

$$\psi_1(x) = \psi(x) + O(x^{1/2}).$$

Now

$$\varepsilon_d^k = \frac{x_{d,k} + \sqrt{x_{d,k}^2 - 4}}{2} = x_{d,k} + O \left( \frac{1}{x_{d,k}} \right),$$
so that essentially
\[ \psi_1(x) = \left| \{(d, k) : x_{d, k} \leq x\} \right| \]
for large \( x \).

The mapping \( (d, k) \mapsto (d, x_{d, k}, y_{d, k}) \) gives a correspondence between solutions of
\[ m^2 - n^2d = 4, \quad m, n > 0, \quad d \in D, \]
and pairs \( (d, k), \ d \in D, \ k \geq 1 \). The last shows that \( N(x) = \psi_1(x) \), which proves the lemma.

We are led to study
\[ N(x) = \# \{(m, n, d) : m \leq x, m^2 - n^2 = 4, \ m, n > 0, \ d \in D\}. \tag{4.3} \]

First, we notice that in counting solutions to (4.3) we may include the case of \( d \) being a perfect square without altering the behavior of \( N(x) \). So for these purposes we may think of \( D \) as being all positive integers congruent to 0 or 1 (mod 4).

Let \( S(n) \) denote the number of solutions of (4.3) in the variables \( m \) and \( d \), for fixed \( n \).

\[
N(t) = \sum_{n=1}^{t} S(n) = \sum_{n \leq t^{1/2}} S(n) + \sum_{t^{1/2} < n \leq t} S(n) = A + B.
\]

We first establish that
\[ B = O(t^{2/3 + \epsilon}). \tag{4.4} \]

To see this, let
\[ S^*(n) = \# \{(m, k) : m \leq t, m^2 - kn^2 = 4\} \]
so that
\[ B \leq \sum_{t^{1/2} < n \leq t} S^*(n). \]

Let \( T^*(n) = \) the number of residue class solutions of
\[ m^2 \equiv 4 \pmod{n^2}. \]

\( T^* \) is multiplicative and one easily checks that \( T^*(n) = O(n^\epsilon) \) for any \( \epsilon > 0 \).
For \( t^{1/2} < n \), \( S^*(n) \leq T^*(n) \). Now

\[
B \leq \sum_{t^{1/2} < n < t^{2/3}} S^*(n) + \sum_{t^{2/3} < n < t} S^*(n) = O(t^{1/2} + \varepsilon) + \sum_{t^{2/3} < n < t} S^*(n).
\]

The latter term is the number of solutions of

\[
m^2 - n^2k = 4, \quad m \leq t, \quad t^{2/3} < n \leq t,
\]

or

\[
n^2k - (m - 2)(m + 2), \quad m \leq t, \quad t^{2/3} < n < t.
\]

So we may write

\[
n = yz \quad \text{with} \quad y^2 \mid (m + 2), \quad z^2 \mid (m - 2),
\]

or

\[
n = 2yz \quad \text{with} \quad 2y^2 \mid (m + 2), \quad z^2 \mid (m - 2), \quad y^2 \mid (m + 2), \quad 2z^2 \mid (m - 2).
\]

In any case, one of \( y \) or \( z > t^{1/3}/2 \).

Thus the \( m \)'s which are solutions of (4.5) are among those \( m \)'s satisfying

\[
m \equiv \pm 2 \pmod{v^2}, \quad \text{where} \quad t^{1/3} \leq v \leq t^{1/2}.
\]

Given such an \( m \) and \( v \), there are clearly at most \( d_2(m \pm 2) \) solutions of (4.5) in \( n \) and \( k \), where \( d_2(l) \) is the number of square divisors of \( l \). Now \( d_2(l) = O(l^\varepsilon) \) for any \( \varepsilon > 0 \).

Therefore the number of solutions of (4.5) is at most

\[
t^\varepsilon \sum_{t^{1/3} < v < t^{1/2}} \# \{ m \leq t: m \equiv \pm 2 \pmod{v^2} \}
\]

\[
= t^\varepsilon \sum_{t^{1/3} < v < t^{1/2}} \left\{ \frac{t}{v^2} + O(1) \right\}
\]

\[
= \frac{t^{1 + \varepsilon}}{t^{1/3}} + O(t^{1/2} + \varepsilon) = O(t^{1/2} + \varepsilon),
\]

proving (4.4).
We return to the term $A$ above (4.4).

$$A = \sum_{n=1}^{t^{1/2}} S(n).$$

To calculate $S(n)$, let $S_1(n)$ be the number of solutions of

$$m^2 - 4kn^2 = 4, \quad k \geq 1, \quad m \leq t,$$

and $S_2(n)$ the number of solutions of

$$m^2 - (4k+1)n^2 = 4, \quad k \geq 1, \quad m \leq t.$$  \hfill (4.7)

So that $S(n) = S_1(n) + S_2(n)$; we evaluate each in turn.

A solution $(m, k, n)$ of (4.6) must have $m$ even, $m = 2m'$, say, and so we are looking at

$$(m')^2 - kn^2 = 1, \quad m' \leq t/2.$$

Let $T_1(n)$ be the number of residue class solutions of $(m')^2 \equiv 1 \pmod{n^2}$. Clearly,

$$S_1(n) = \frac{t T_1(n)}{2n^2} + O(T_1(n))$$

$$= \frac{t}{2} \frac{T_1(n)}{n^2} + O(n^\epsilon) \quad \text{for any } \epsilon > 0. \quad (4.8)$$

Let $T_2(n)$ be the number of residue class solutions of

$$m^2 \equiv (n^2 + 4) \pmod{4n^2}.$$

Then from (4.7)

$$S_2(n) = \frac{t}{4n^2} T_2(n) + O(T_2(n)),$$

which again

$$- \frac{t}{4n^2} T_2(n) + O(n^\epsilon) \quad \text{for any } \epsilon > 0.$$
Therefore,

\[
A = \sum_{n \leq \sqrt{t}} \left\{ \frac{tT_1(n)}{2n^2} + \frac{tT_2(n)}{4n^2} \right\} + O(n^\epsilon)
\]

\[
= t \left( \sum_{n=1}^{\infty} \frac{T_1(n)}{2n^2} + \frac{T_2(n)}{4n^2} \right) + O(t^{(1/2)+\epsilon}).
\]

So it remains to evaluate the above infinite series.

\( T_1(n) \) is multiplicative with \( T_1(2) = 2 \), and \( T_1(2^j) = 4 \) for \( j \geq 2 \).

\[ T_1(p^e) = 2 \quad \text{for} \quad 1 \geq 1, \; p \geq 3 \; \text{a prime}. \]

Thus

\[
\sum_{n=1}^{\infty} \frac{T_1(n)}{n^2} = \prod_{\rho \leq 3} \left( 1 + T_1(\rho) \rho^{-2} + T_1(\rho^2) / (\rho^{-2})^2 \cdots \right)
\]

\[
= (1 + 2(2^{-2}) + 4(2^{-2})^2 \cdots) \prod_{\rho \geq 3} (1 + 2\rho^{-2} + 2(\rho^{-2})^2 \cdots)
\]

\[
= \left( 1 + \frac{1}{2} + \frac{4 \cdot 2^{-4}}{1 - 2^{-2}} \right) \prod_{\rho \geq 3} \left( 1 + \frac{2\rho^{-2}}{1 - \rho^{-2}} \right)
\]

\[
= \left( \frac{3}{2} + \frac{1}{3} \right) \prod_{\rho \geq 3} \left( \frac{1 + \rho^{-2}}{1 - \rho^{-2}} \right)
\]

\[
= \left( \frac{11}{6} \right) \left( \frac{1 - 2^{-2}}{1 + 2^{-2}} \right) \prod_{\rho} \left( \frac{1 + \rho^{-2}}{1 - \rho^{-2}} \right)
\]

\[
= \left( \frac{11}{6} \right) \left( \frac{3}{4} \right) \left( \frac{4}{5} \right) \left( \frac{\zeta(2)^2}{\zeta(4)} \right)
\]

\[
= \frac{11}{4}.
\]

\( T_2(n) \) is not multiplicative, indeed, if \( n = 2^\beta p_1^{e_1} \cdots p_k^{e_k}, p_i \geq 3 \), or we write it as \( p_0^{e_0} \cdots p_k^{e_k} \) with \( p_0 = 2 \), \( e_0 = \beta \). Then

\( T_2(n) = 2^k \phi(\beta), \)

where

\[
\phi(0) = 2, \; \phi(1) = 0, \; \phi(2) = 0, \; \phi(k) = 8, \quad k \geq 3. \quad (4.9)
\]

To see this notice that we are counting solutions of

\[
m^2 \equiv (n^2 + 4) \mod (p_0^{2e_0 + 2} p_1^{2e_1} \cdots p_k^{2e_k}).
\]
The number of these is the product of residue class solutions to

\[ m^2 \equiv (n^2 + 4) \mod (p_i^k), \quad i = 0, \ldots, k, \]

\[ \gamma_i = 2e_i \quad \text{for} \quad i = 1, \ldots, k; \quad \gamma_0 = 2e_0 + 2. \]

If \( p_i \) is odd then \( p_i^k \mid n^2 \) and \( m^2 \equiv 4 \mod (p_i^k) \) has exactly two solutions. So we are left with

\[ m^2 \equiv (n^2 + 4) \mod (2^{2\beta+2}), \quad n^2 = 2^{2\beta}y^2, \]

where \( y \) is odd.

\[ m^2 \equiv (2^{2\beta}y^2 + 4) \mod 2^{2\beta+2} \]

when \( \beta = 0 \) gives

\[ m^2 \equiv (y^2 + 4) \mod 4, \]

i.e., two solutions. \( \beta = 1 \) gives

\[ m^2 \equiv (4y^2 + 4) \mod 16, \]

which has no solutions. \( \beta = 2 \) gives

\[ m^2 \equiv (16y^2 + 4) \mod 2^8, \]

which has no solutions. \( \beta \geq 3 \) gives

\[ m^2 \equiv (2^{2\beta}y^2 + 4) \mod 2^{2\beta+2}, \]

\[ m = 2m', \quad 0 < m' < 2^{2\beta+1}. \]

In this case \( 2^{2\beta-2}y^2 \equiv 1 \mod 8 \), so there are four solutions in \( \{0, 1, \ldots, 2^{2\beta} - 1\} \), and so eight in \( \{0, 1, \ldots, 2^{2\beta+1}\} \).

This proves (4.9).

Now if \( W_z(n) = T_z(n)/2 \), then \( W_z(1) = 1, \quad W_z(2) = 0, \quad W_z(2^j) = 0, \quad W_z(2^k) = 4 \) for \( k \geq 3 \), while \( W_z(p^j) = 2 \) for \( j \geq 1, \quad p \geq 3 \) a prime. Therefore

\[ \sum_{n=1}^{\infty} \frac{W_z(2)}{n^z} = \prod_p (1 + W_z(p)p^{-z} + \cdots) \]

\[ = (1 + 4(2^{-2}) + 4 \cdot (2^{-2})^4 \cdots) \prod_{p \geq 3} (1 + 2p^{-2} + 2(p^{-2})^2 \cdots) \]

\[ = \left(1 + \frac{4 \cdot 2^{-6}}{1 - 2^{-2}}\right) \prod_{p \geq 3} \left(1 + \frac{2p^{-2}}{1 - p^{-2}}\right) \]
\[
\left(1 + \frac{1}{12}\right) \prod_{p \geq 3} \left(\frac{1+p^{2}}{1-p^{2}}\right)
= \left(1 + \frac{1}{12}\right) \left(\frac{1-2^{2}}{1+2^{2}}\right) \prod_{p} \left(\frac{1+p^{2}}{1-p^{2}}\right)
= \frac{13}{12} \cdot \frac{3}{5} \left(\frac{\zeta(2)}{\zeta(4)}\right)
= \frac{13}{8}.
\]

Thus
\[
\sum_{n=1}^{\infty} \frac{T_2(n)}{n^2} = \frac{13}{4}.
\]

Finally,
\[
A = t \left(\frac{11}{8} + \frac{13}{16}\right) + O(t^{1/2+\epsilon})
= \frac{35}{16} t + O(t^{1/2+\epsilon}).
\]

Therefore, by (4.4),
\[
N(t) = \frac{35}{16} t + O(t^{2/3+\epsilon})
\]
proving Theorem 4.1 for \(p = 1\).

To do the case \(p \geq 3\) a prime we let
\[
\psi_p(x) = \|\{d: d \in D_p, \varepsilon_d \leq x\}\| = |D_{p,x}|
\]
and let
\[
N_p(x) = \|\{(t, u, d): t \leq x, t^2 - du^2 = 4, t, u \geq 0, d \in D, u \equiv 0 \pmod{p}\}\|.
\]

**Lemma 4.10.**
\[
\psi_p(x) = N_p(x) + O(x^{1/2}).
\]

**Proof.** First, \(N_p(x) \geq \psi_p(x)\) since every \(d \in D_p\) gives rise to a \((t_d, u_d, d)\) with \(t_d \leq x\) (essentially), and \(u_d \equiv 0 \pmod{p}\). So we need to estimate the difference
\[
N_p(x) - \psi_p(x).
\]
This will be dominated by the number of solutions \((t, u, d)\) of \(t^2 - du^2 = 4, t \leq x, u \equiv 0 \pmod{p}\), for which \((t, u)\) is not a fundamental solution of the Pell equation. The latter is dominated by the number of solutions \((d, k)\) of \(\varepsilon_d^k \leq x, d \in D, k \geq 2\), which we have seen before is \(O(x^{1/2})\). This proves the lemma, and notice the implied constant is independent of \(p\).

We can now complete the proof of Theorem 4.1 for the case \(p \geq 3\). We may proceed as we did for \(p = 1\), the only difference now is that \(u\) is divisible by \(p\). The term corresponding to \(\sum_{x^{1/2} < n < x, S^*(n)\) is again \(O(t^{2/3} + \epsilon)\), since it is dominated by the \(p = 1\) term.

We have

\[
N_p(t) = t \left[ \frac{1}{2} \sum_{p | n} \frac{T_1(n)}{n^2} + \frac{1}{4} \sum_{p | n} \frac{T_2(n)}{n^2} \right].
\]

Similar Euler product calculations to the previous ones give

\[
\sum_{p | n} \frac{T_1(n)}{n^2} = \frac{11}{2(1 + p^2)}
\]

and

\[
\sum_{p | n} \frac{T_2(n)}{n^2} = \frac{13}{2(1 + p^2)}.
\]

Therefore

\[
N_p(t) = \frac{35}{8(1 + p^2)} t + O(t^{2/3} + \epsilon),
\]

which completes the proof of Theorem 4.1.

We now combine Theorems 3.1 and 3.4 and Proposition 4.1 to obtain our main result.

**Theorem 4.11.** For \(\gamma > \frac{2}{3}\),

\[
\frac{1}{|D_{p,x}|} \sum_{d \in D_{p,x}} h(d) = \frac{16}{35} c_p \frac{\text{Li}(x^2)}{x} + O(x^{\epsilon}),
\]

where

\[
c_p = 1 \quad \text{if} \quad p = 1,
\]

\[
-\frac{1 + p^2}{p(p^2 - 1)} \quad \text{if} \quad p \geq 3.
\]
As a final remark, we point out that the above tells us that the average of $h$ over the sets $D_{p,x}$ decreases as $p \to \infty$ (since $c_p \to 0$ as $p \to \infty$). This decrease in the size of $h$ is not surprising in view of the work of Siegel (see Davenport [1]), which, though ineffective, shows that $h(d) \log \epsilon_d$ is approximately of the order of $d^{1/2}$. Now on the sets $D_p$ we are forcing the fundamental unit to be larger, and so we expect the average of $h$ to be smaller.

**ACKNOWLEDGMENTS**

I would like to thank my advisor Professor Paul Cohen, as well as Alex Woo, for many fruitful discussions. I would also like to thank Drs. Lagarias and Odlyzko for considerably simplifying the proof of (4.1).

**REFERENCES**