Numerical Evidence for a Conjecture of Hooley

Ian N Petrow

An Undergraduate Thesis
Advised by Peter C Sarnak
in the Department of Mathematics
of Princeton University
in Candidacy for the Degree
of Artium Baccalaureus

© Copyright by Ian N Petrow, 2007.

All Rights Reserved

Acknowledgements

There are many people in my life that have made the writing of this thesis possible. Far too many to be mentioned here. First and foremost is my family, without whose love and encouragement I would not have been able to make it this far. Few are so lucky to have such support in pursing their dreams. Secondly, I am thankful of my roommates and all of my close friends for keeping things interesting, and preventing me from becoming too much of a nerd. Lastly, I want to thank all of my mathematical mentors. Specifically, Ken Gross for his early guidance in my mathematical life, Eli Stein for introducing me to number theory, Giuliana Davidoff for getting me started on class numbers, Christopher Hooley for writing the paper I've spent the past six months grappling with, and especially Peter Sarnak for his guidance, insight, and for keeping me honest throughout this project.

To my father,

who instilled in me

stubbornness and tenacity

at an early age.

Contents

Chapter 1

Introduction

As the simplest nontrivial diophantine equations, quadratic forms represent one of the most well-studied areas of number theory. Although the subject is a relatively old one, open questions persist, and many are concerned with the class number of quadratic forms. The class number of binary quadratic forms of discriminant d is defined classically as the number of inequivalent forms of discriminant d under the action of $SL_2(\mathbf{Z})$ by change of variable. It can also be realized as the order of the ideal class group, however this definition will not arise in this paper.

It was already clear to Gauss that the behavior of class numbers is vastly different for positive and negative discriminants. The theory of class numbers of negative discriminants is by far the easier case, and was essentially worked out in its entirety in the nineteenth century. The case of positive discriminants has proven, however, to be much more difficult. The difference between the cases of positive and negative d arises in differences in the group of automorphs in $SL_2(\mathbf{Z})$ of a form. This group is always finite for a form of negative discriminant, whereas it is infinite cyclic in the case of positive discriminants. Algebraically stated, $|(\mathcal{O}_d)^\times| < \infty$ for $d < 0$ (where $\mathcal{O}_d$ is the ring of integers associated with the number field $\mathbf{Q}(\sqrt{d})$), and $|(\mathcal{O}_d)^\times| = \infty$ for $d > 0$. Analytically stated, we have the two cases of Dirichlet's class number formula

[Dav],

$$h(d) = \begin{cases} \frac{w(d)\sqrt{|d|}L_d(1)}{2\pi} & \text{if } d < 0, \\ \frac{\sqrt{d}L_d(1)}{\log(\epsilon_d)} & \text{if } d > 0, \end{cases} \tag{1.1}$$

where $L_d(s)$ is the Dirichlet L-function associated to $|d|$, $\epsilon_d = \frac{t+u\sqrt{d}}{2}$ is the fundamental solution to Pell's equation $t^2 - du^2 = 4$, or the fundamental unit of $\mathcal{O}_d$, and $w(d)$ is the number of automorphs of a form of discriminant d or $|(\mathcal{O}_d)^\times|$ and is given by

$$w(d) = \begin{cases} 2 & \text{if } d < -4, \\ 4 & \text{if } d = -4, \\ 6 & \text{if } d = -3. \end{cases}$$

The presence of ϵ_d in (1.1) significantly complicates the study of $h(d)$ for $d > 0$, as it fluctuates wildly with d. Generally speaking, the best possible bounds on ϵ_d are given by

$$2\sqrt{d} < \epsilon_d < e^{c\sqrt{d}\log(d)},$$

where the lower bound is given by elementary considerations, and the upper bound is given by considering the genera of binary quadratic forms of discriminant d and the class number formula.

1.1 The Problem and Its History

One would like to know just how big the class numbers are. Because $h(d)$ fluctuates widely from one discriminant to the next, the appropriate question to ask is "how fast do the $h(d)$ grow on average for large $|d|$?" In this paper we are concerned with the following

Problem 1.1.1. *To what function is*

$$s(x) = \sum_{|d|<x} h(d) \quad \sim \ ?$$

This problem has a long history, from Gauss to Dirichlet, Siegel, Sarnak, Hooley and Kwon. In this paper, I will present numerical evidence supporting a conjecture of Hooley's on the positive discriminant case. These numerics are based on a computation of the all positive class numbers of $d \leq 5.2 \times 10^7$ preformed by Lee Kennard, Jennifer Koonz, Katharine Shultis, Haokun Xu and myself at the summer 2006 Mount Holyoke REU in Mathematics under the direction of Giuliana Davidoff. This paper represents the first time that theory on asymptotics of positive discriminant class numbers has been shown to agree with practical computation. My hope is that the confidence that these numerics give to the theoretical work previously done will encourage new interest in this area of the theory of class numbers.

1.1.1 Gauss, Siegel

As is not surprising, Gauss was the first to make any statment concerning asymptotics of class numbers. Whereas we now take a binary quadratic form to be any $ax^2+bxy+cy^2$ with $a,b,c \in \mathbf{Z}$, Gauss worked over "classical forms" $ax^2+2bxy+cy^2$, $a,b,c \in \mathbf{Z}$. Correspondingly, we now take the discriminant to be $d = b^2 - 4ac$, whereas Gauss used "classical discriminants" $D = b^2 - ac$. In an attempt to avoid confusion, I will use lower-case d when refereing to standard discriminants, and a capital D to refer to classical discriminants. Note that d falls into only two residue classes modulo 4, namely $d \equiv 0,1(4)$, whereas D may take on any value modulo 4. The case of $d \equiv 0(4)$ corresponds to the case that b is even, and hence for $d \equiv 0(4)$ we have $d = 4D$. Therefore, although the classical discriminants D fall into all four residue classes modulo 4, they can be viewed as only encompassing half of the possible discriminants.

In Gauss's *Disquisitiones Arithmeticae* he proves for classical negative discriminants

$$\sum_{|D|\leq X} h(D) \sim \frac{4\pi}{21\zeta(3)} X^{\frac{3}{2}}.$$

Whereas for positive discriminants he asserts

$$\sum_{D\leq X} h(D)\log(\epsilon_D) \sim \frac{4\pi^2}{21\zeta(3)} X^{\frac{3}{2}}$$

without giving proof [Gau]. The proof of this statement and that of the corresponding asymptotic for standard discriminants

$$\sum_{d\leq x} h(d)\log(\epsilon_d) \sim \frac{\pi^2}{18\zeta(3)} x^{\frac{3}{2}}$$

were not given until almost 150 year later by Siegel in his 1944 paper *The Average Measure of Quadratic Forms with Given Determinant and Signature.* [Sie].

1.1.2 Hooley

In light of Dirichlet's class number formula (1.1) it is usually assumed that for positive d, the quantities $h(d)$ and $\log(\epsilon_d)$ are essential inseparable. Indeed, this was the view that prompted Gauss and later mathematicians to prove asymptotics for $\sum h(d)\log(\epsilon_d)$ as opposed to $\sum h(d)$ itself. Having been viewed as an extremely difficult question, no asymptotic for the later sum was even proposed until 1984. In his paper *On the Pellian equation and the class number of indefinite binary quadratic forms* Christopher Hooley proposed the following

Conjecture 1.1.2. *For the class numbers $h(D)$ of classical positive discriminants D we have*

$$S(X) = \sum_{D\leq X} h(D) \sim \frac{25}{12\pi^2} X\log^2(X). \tag{1.2}$$

[Hoo]

Although Hooley gives a detailed account of his process for arriving at this asymptotic, he makes a key assumption which unfortunately prevents his arguments from being rigorous. Chapter 3 provides a more detailed account of Hooley's method for arriving at this asymptotic.

1.1.3 Kwon

In the academic year 2005-2006 as her undergraduate thesis (also under the direction of Peter Sarnak) Suehyun Kwon undertook to produce numerics on the asymptotic for standard positive discriminants [Kwo]

$$s(x) = \sum_{d \leq x} h(d) \quad = \ ? \tag{1.3}$$

The program she wrote used an algorithm based on the method of grouping reduced forms into chains of equivalent forms which can be found in Dickson's *Introduction to the Theory of Numbers* [Dic]. Anticipating that any asymptotic of the above form would include logarithm terms, and that $\log(x)$ would remain small for values of x she could reasonably hope to achieve, Kwon looked for a fomula of the type

$$s(x) = ax\log^2(x) + bx\log(x) + cx + O(x^\gamma), \ \text{ where } \gamma \in (0, 1).$$

At such small values of x, the lower order terms corresponding to b and c still significantly affect $s(x)$, and one can not hope to get an accurate picture in this restricted data range without their inclusion. After using her program to compute $h(d)$ for $d \leq 3.5 \times 10^6$, she proposes

Conjecture 1.1.3. *For the class numbers $h(d)$ of standard positive discriminants d*

we have

$$s(x) = \sum_{d<x} h(d) = 0.0661...x\log^2(x) - 0.894...x\log(x) + 4.96...x + O(x^\gamma), \quad \textit{where } \gamma \in (0,1). \tag{1.4}$$

where it is understood that the explicit constants displayed are expected to be accurate to the number of decimal places displayed (and have no meaning beyond their truncation).

As we will see later, this conjecture is not correct because $s(x)$ does not settle down to an asymptotic until around $x = 8 \times 10^6$, which Kwon could not have predicted given her limited amount of data.

Chapter 2

Numerical Evidence

2.1 Mount Holyoke

2.1.1 Computation

At Peter Sarnak's suggestion, Giuliana Davidoff led the 2006 Mount Holyoke College mathematics REU in attempting to verify and improve upon the results of Kwon. Much of the summer was spent learning the necessary theoretical background to do work in this discipline, but the project succeeded in extending Kwon's data.

Whereas Kwon used her own algorithm based on the classical theory of quadratic forms, the REU employed PARI/GP, a widely used computer algebra system designed for fast computations in number theory developed by Henri Cohen, to compute the class number [PARI]. Running PARI on 15 classroom desktop PCs for 3 weeks produced all $h(d)$ with $d \leq 5.2 \times 10^7$. Excluding the $h(d)$ for $d \equiv 1(4)$ and dividing the discriminants by 4 gave us the corresponding set of $h(D)$ for classical discriminants $D \leq 1.3 \times 10^7$. Thus, a relatively large set of data was produced with which we could check conjectures concerning both standard and classical discriminants.

To compute class numbers, we used PARI's built-in function qfbclassno( ). As opposed to the classical method of computing $h(d)$, PARI computes the class number analytically via the class number formula (1.1). The algorithm, which can be found in Cohen's book *A Course in Computational Algebraic Number Theory* , has two main parts, the computations of $L_d(1)$ and of the regulator $\log(\epsilon_d)$.

Computation of $L_d(1)$ directly from it's definition

$$L_d(1) = \sum_{n \geq 1} \left(\frac{d}{n}\right) \frac{1}{n}$$

is extremely slow because the defining series is only conditionally convergent. PARI's algorithm however, makes use of the functional equation of $L_d(s)$

$$\Lambda_d(1-s) = \Lambda_d(s)$$

where

$$\Lambda_d(s) = (d/\pi)^{s/2} \Gamma(s/2) L_d(s)$$

to obtain an exponentially converging method of computing $L_d(1)$, which is based on the formula

$$h(d)\log(\epsilon_d) = \frac{1}{2} \sum_{n \geq 1} \left(\frac{d}{n}\right) \left(\frac{\sqrt{D}}{n} \text{erfc}\left(n\sqrt{\frac{\pi}{d}}\right) + E_1\left(\frac{\pi n^2}{d}\right)\right),$$

where

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt$$

is the error function, and

$$E_1(x) = \int_x^\infty \frac{e^{-t}}{t}\, dt$$

is the exponential integral[Coh]. Efficient algorithms are known for computing both of these functions.

Secondly, the algorithm needs $\log(\epsilon_d)$ to compute $h(d)$ from the class number formula. The primary concern with this aspect of the algorithm is that ϵ_d quickly becomes extremely large. Accordingly, PARI computes the regulator $\log(\epsilon_d)$ directly to avoid such an overflow. The algorithm for computing the regulator is essentially the standard continued-fraction method of finding solutions to $t^2 - du^2 = \pm 1$. However, as soon as numbers become large, they are converted to logarithmic form and stored as such, avoiding dealing with absurdly large numbers while still keeping the number of steps of computation down [Coh].

For $d > 0$, the algorithm is $O(d^{\frac{1}{2}+\epsilon})$, and runs substantially faster than the classical algorithm from what we observed.

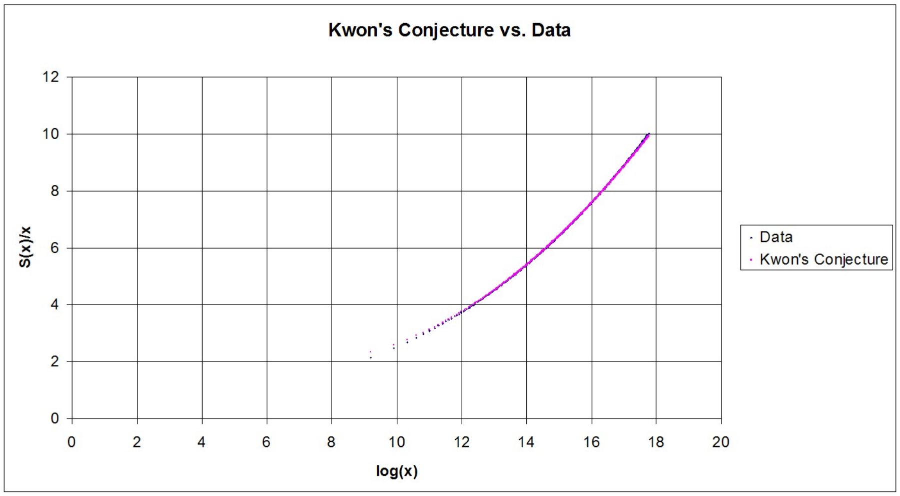
2.1.2 Naïve Data Analysis

The Mount Holyoke REU used its data for two purposes. First, to check if Kwon's conjecture (Conjecture 1.1.3 above) continued to hold above $x = 3.5 \times 10^6$. Secondly, to check the Hooley's conjecture (Conjecture 1.1.2 above) numerically. In the first aim the REU was partially successful, but at the time results were inconclusive as to the second aim.

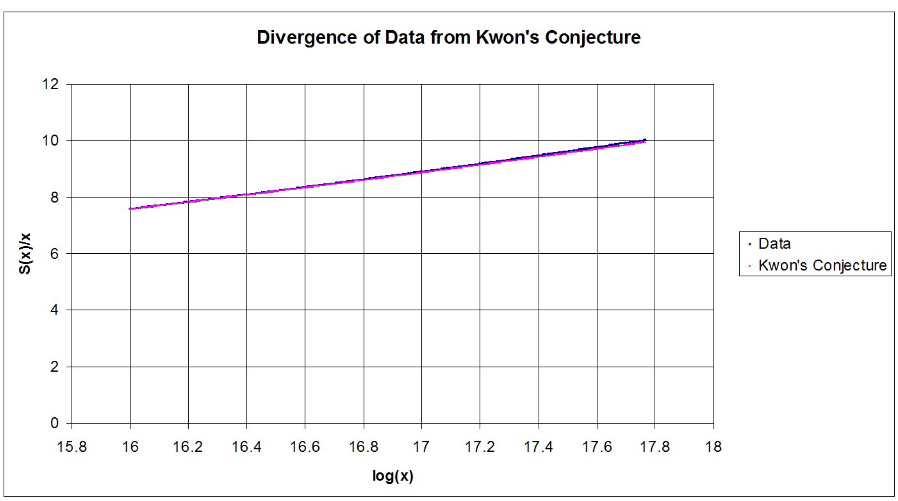
All of the subtlety in the following analysis flows from the fact that logarithms arise in the curves we are trying to predict. As we have already mentioned above, and will below show, $s(x)/x$ or $S(X)/X$ (as defined in either (1.2) or (1.3)) does not settle down to an asymptotic until $x > 8 \times 10^6$ for standard discriminants ($X > 2 \times 10^6$ for classical discriminants). We also are restrained by $x \leq 5.2 \times 10^7$ for standard discriminants ($X \leq 1.3 \times 10^7$ for classical discriminants). These considerations restrain us to looking at the data in the restricted domains $15.89 < \log(x) < 17.77$, or $14.51 < \log(X) < 16.38$. Hence, it is useless to merely consider the $ax\log^2(x)$ term in the asymptotics for $s(x)$ and $S(X)$. We must look at the next two lower order terms $bx\log(x)$ and cx as well.

In the summer of 2006 it was not yet known when $s(x)$ would settle down. We will ignore the lower restrictions on x from the previous paragraph for the remainder of this section in order to illustrate what was known by the end of the summer of 2006.

First, one would like to check Kwon's prediction with knowledge of $s(x)$ to $x = 5.2 \times 10^7$. The below graph plots Kwon's prediction and the actual data on the same axes.

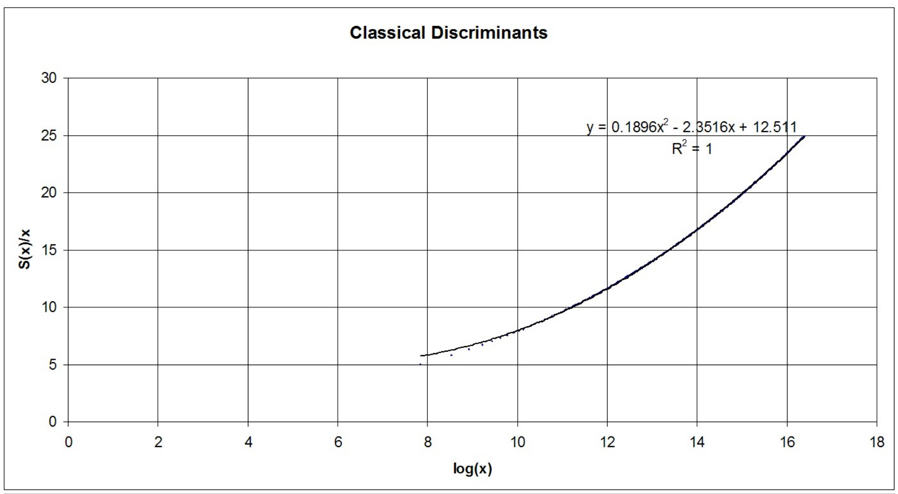


At first glace, Kwon's curve seems to agree with the data. However, the astute reader will notice that the blue data line seems to just barely overtake the pink prediction line at the large end of the available data. Expanding this region of the graph,



we see that the two lines indeed diverge. Therefore, postponing a detailed analysis, we might guess that Kwon's prediction is close to correct but the data diverges slightly at the end, perhaps due to natural fluctuation in the data. At the very least it suggests that we must pay closer attention to what happens in this region and perhaps revise our prediction.

Turning now to the case of classical discriminants and Hooley's conjecture, we preform the same operation. This time we use a simple least-squares regression instead of checking against a conjecture with three terms as above.



Least squares predicts an asymptotic of

$$S(X) \sim 0.1896...X\log^2(X)$$

for the classical discriminant case, which does not closely resemble Hooley's theoretical prediction

$$S(X) \sim \frac{25}{12\pi^2}X\log^2(X) = .2111...X\log^2(X).$$

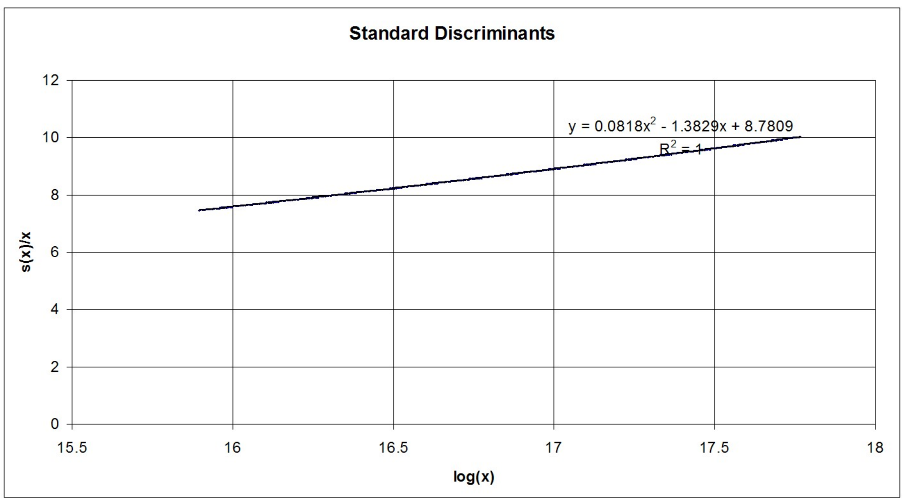
Again, it seems that something is amiss with either our numerics or with Hooley's prediction.

2.2 Princeton

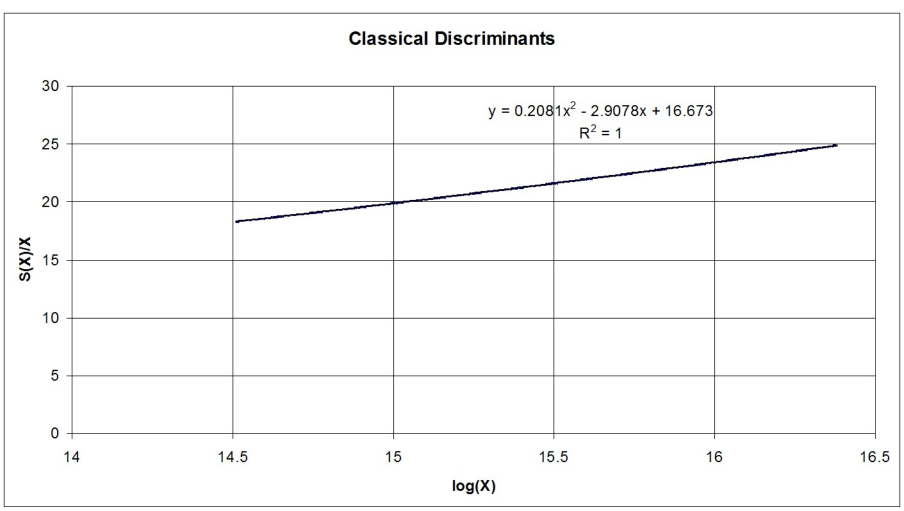
In light of the inadequacies of the above analysis of the data at hand, a closer look was necessary after the program at Mount Holyoke had ended.

2.2.1 A Somewhat Less Naïve Approach

We now take into account a knowledge of the lower bounds $8 \times 10^6 < x$ and $2 \times 10^6 < X$ mentioned at the beginning of the previous section. For the case of standard discriminants we have



which suggests an asymptotic quite different from that predicted by Kwon. The case of classical discriminants however, provides more confidence,



as the predicted asymptotic

$$S(X) \sim 0.2081...X \log^2(X)$$

is quite close to Hooley's theoretical prediction.

However, we cannot trust these two calculations too strongly. If one chooses modestly different lower bounds on our ranges for x and X, then the coefficients predicted vary significantly. For example, if one chooses instead $1\times10^7 < x < 5.2\times10^7$ ($\log(1\times10^7) = 16.12$), then the first coefficent rises from .0818 to .0839 in the standard discriminant case. The following section will show more clearly the behavior of $s(x)$ and $S(X)$, and show from whence we derive the bounds of $x > 8\times10^6$ and $X > 2\times10^6$.

2.2.2 Method of Derivatives

We will use the fact that we expect $s(x)$ and $S(X)$ to be of the form

$$ax\log^2(x) + bx\log(x) + cx + O(x^\gamma), \quad \text{where } \gamma \in (0,1)$$

to obtain a second method of finding a, b and c. Let

$$f''_h(u) = \frac{f(u-h) - 2f(u) + f(u+h)}{h^2} \tag{2.1}$$

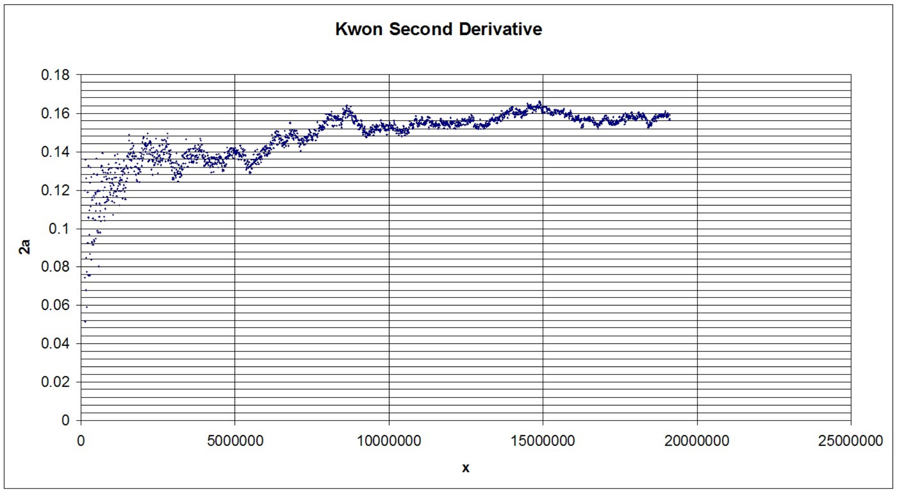
be the second symmetric derivative of $f(x)$, so named because it resembles the usual definition of the derivative, taken twice. For $f \in C^2(\mathbf{R})$,

$$\lim_{h\to0} f''_h(u) = f''(u),$$

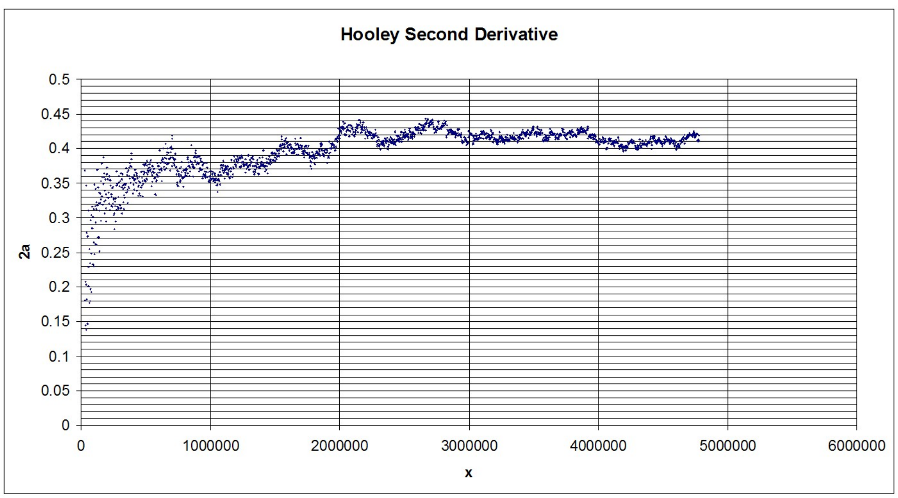
the usual second derivative. However, if f is a quadratic polynomial in u, a simple calculation yields

$$f''_h(u) = 2a \quad \text{for all } h.$$

This will allow us to compute $2a$ for our data by taking $f(u) = S(e^u)/e^u$ and $u = \log(x)$. For sake of simplicity we take $h = 1$ and plot $f_1''(\log(x))$ for our data (although it seems one might be able to improve these results by allowing h to vary). We have



for standard discriminants and



for classical discriminants. In both cases, the predicted a rises steadily for $x < 8 \times 10^6$ or $X < 2 \times 10^6$, and then appears to level off after that point. It is from the second symmetric derivative that we predict that $s(x)$ and $S(X)$ do not fit well to an asymptotic before these points, and thus we derive the previously stated lower bounds.

We would also like to predict the next two lower order terms $bx\log(x)$ and cx in similar fashion. Taking the average value of $\frac{1}{2}f_1''(\log(x)) = a$ for $s(x)$ and $S(X)$ from 8 million to 19.1 million and 2 million to 4.7 million, respectively, we predict that for standard discriminants $a = .0783$ and for classical discriminants $a = .2088$. For standard discriminants, the standard deviation of these data is $\sigma = .0017$ and for classical discriminants, we have $\sigma = .0045$. Taking

$$s^*(x) = s(x) - ax\log^2(x)$$

and likewise for $S(X)$, we predict that $s^*(x)/x$ and $S^*(X)/X$ are linear in $\log(x)$, of the form

$$b\log(x) + c + O(x^{-\gamma}), \text{ where } \gamma \in (0,1).$$

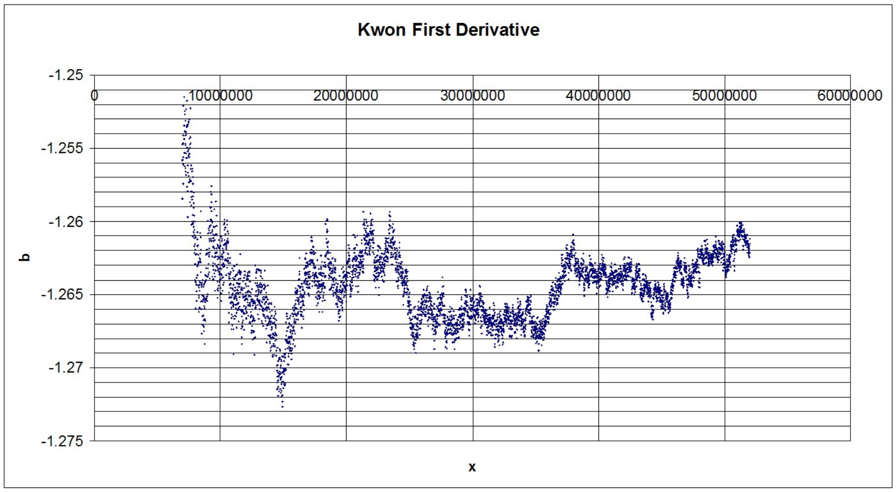
Hence, we may apply the same trick, this time using the first derivative instead. Define as in (2.1)

$$f_h'(u) = \frac{f(u) - f(u-h)}{h} \tag{2.2}$$

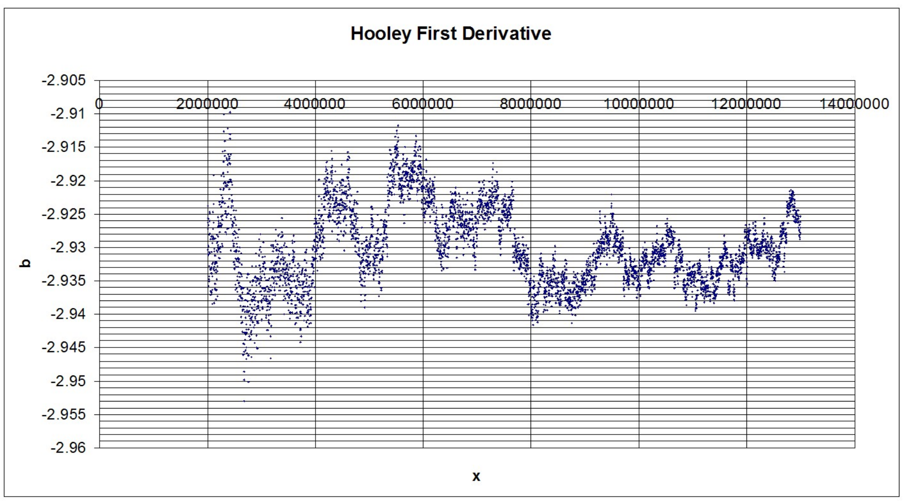
Again, if $f(u)$ is linear in u, then we have that

$$f_h'(u) = b \text{ for all h.}$$

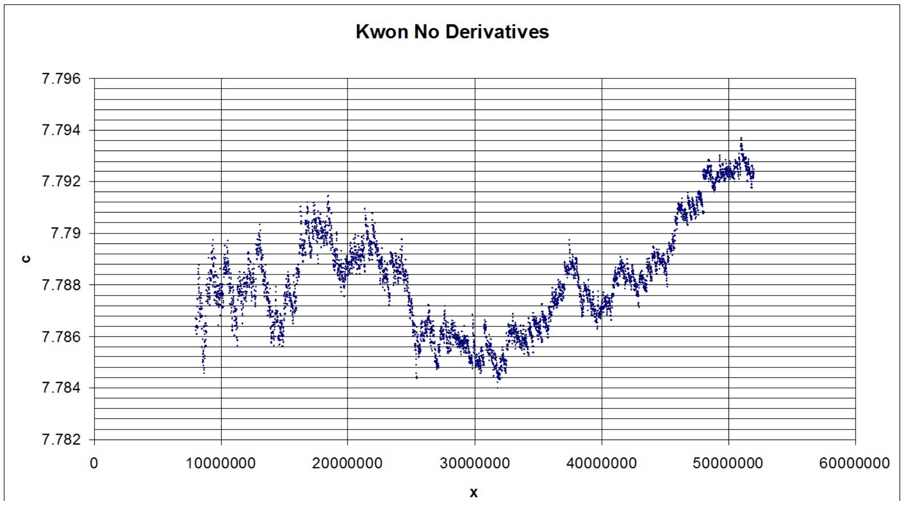
Taking a to be the average value from the previous analysis, we plot the $f_1'(\log(x)) = b$ corresponding to the $s^*(x)$ and $S^*(X)$, and have



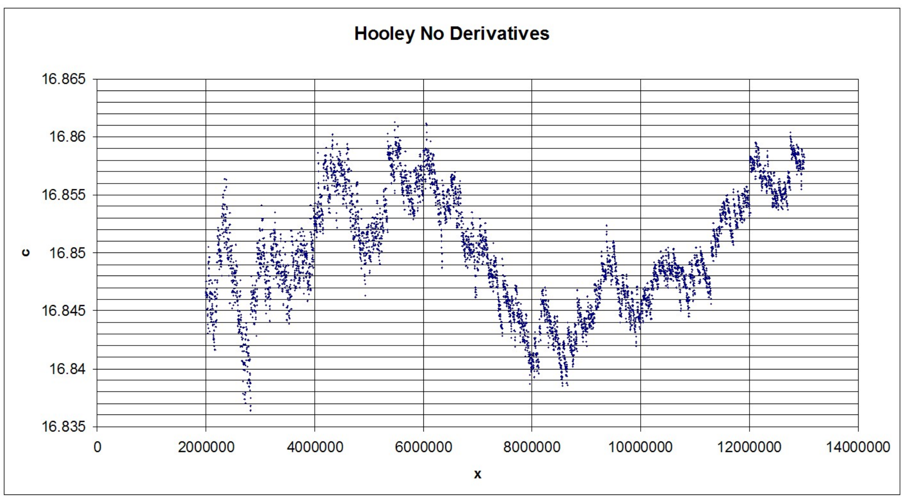
for standard discriminants and



for classical discriminants. These suggest the average values $b = -1.26$ for standard discriminants and $b = -2.93$ for classical discriminants. However, we must take into account that the a values we assumed to create these graphs have some uncertainty given by their standard deviations. Taking the extreme values in the ranges $.0783 \pm .0017$ and $.2088 \pm .0045$ for a, one sees that we can only expect that $b = -1.26 \pm .06$ for standard discriminants and $b = -2.93 \pm .14$ for classical discriminants. Again assuming the average values for b, we may simply subtract these linear terms from $s^*(x)$ and $S^*(X)$, which allows us to plot the predicted constant terms as well. We have



for standard discriminants and



for classical discriminants. By the same process we predict an even greater uncertainty in c. We can nonetheless be confident that that $c = 7.8 \pm .6$ for standard discriminants and $c = 16.9 \pm 1.1$ for classical discriminants.

2.2.3 Predictions and Errors

In light of the above arguments, we make the following two conjectures.

Conjecture 2.2.1. *For the class numbers $h(d)$ of standard positive discriminants d*

we have

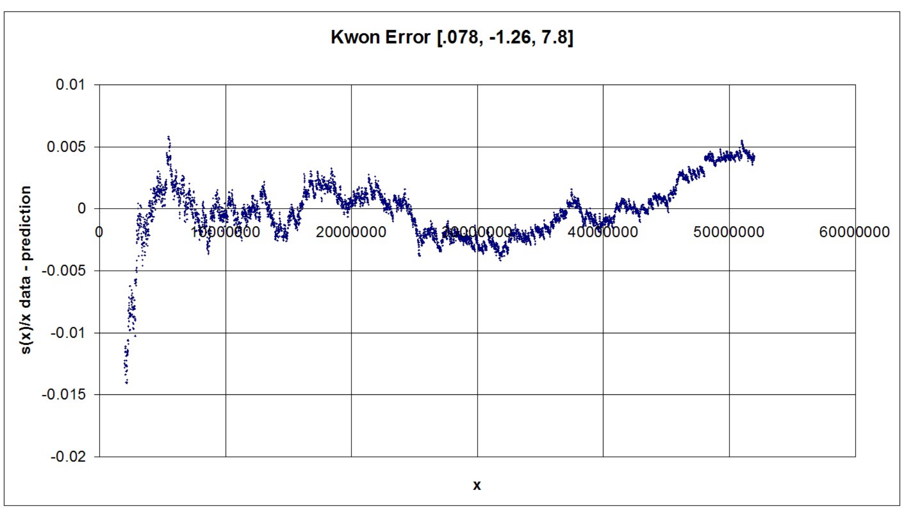
$$s(x) = \sum_{d \le x} h(d) = (0.0783 \pm .0017) x \log^2(x) - (1.26 \pm .06) x \log(x) + (7.8 \pm .6) x + o(x) \tag{2.3}$$

Conjecture 2.2.2. *For the class numbers $h(D)$ of classical positive discriminants D we have*

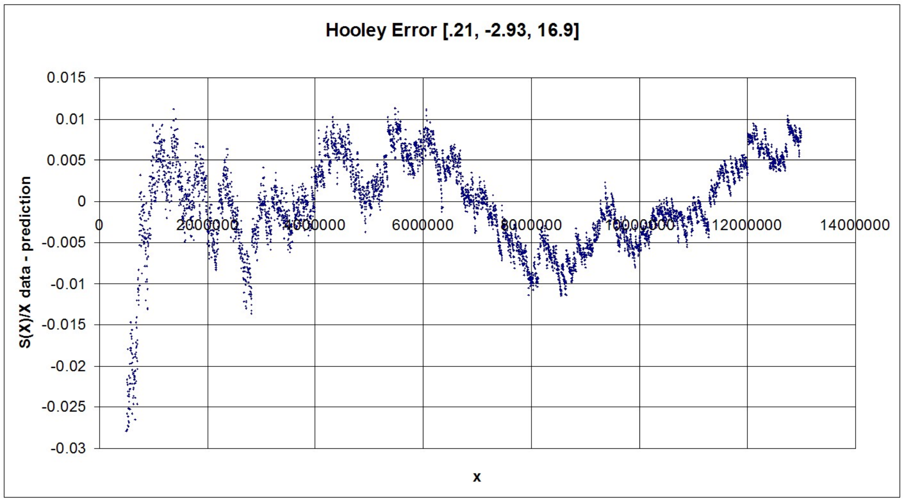
$$S(X) = \sum_{D \le X} h(D) = (0.2088 \pm .0045) X \log^2(X) - (2.93 \pm .14) X \log(X) + (16.9 \pm 1.1) X + o(X) \tag{2.4}$$

Clearly, these two results are highly tentative, however we are encouraged by the resemblance of Conjecture 2.2.2 to Hooley's conjecture 1.1.2. As mentioned before, this is the first time numerics and theory have predicted the same constant for this asymptotic.

Also of note is that we have relaxed the error terms to $o(x)$ as opposed to Kwon's $O(x^\gamma)$. With such wide ranges for c in the above analysis, any attempt at investigating error terms is highly speculative. Nonetheless, if one chooses the average values in the above ranges the error terms can be plotted. For the first conjecture, we have



and for the second we have



From these graphs, it is impossible to predict what the decay of these terms might look like. Indeed, they do not appear to decrease at all within the range of our data, but appear to merely be bounded by $|\text{error}| < .005$ or $|\text{error}| < .01$, respectively. However it is notable that even though we do not have confidence of more than $\pm.6$ or ± 1.1 in c, the error terms are bounded below .005 for standard discriminants and below .01 for classical discriminants through the range of our data if we pick the average values.

Chapter 3

Hooley's Heuristic

Hooley's Heuristic method for computing the asymptotic on $S(X)$ relies entirely on the class number formula (1.1). For positive discriminants D the formula has three main components, $h(D)$, $\log(\epsilon_D)$, and $L_D(1)$. As $L_D(1) = O(\log(D))$, the formula links $\log(\epsilon_D)$ and $h(D)$ together in such a way that knowledge of either one restricts the other into a fairly narrow range. Therefore, Hooley's method works entirely with the units ϵ_D [Hoo].

3.1 A First Method for Counting Units

For any classical discriminant D, the fundamental unit is given by the fundamental solution to the Pellian equation

$$T^2 - DU^2 = 1, \tag{3.1}$$

Where

$$\epsilon_D = T + \sqrt{D}U.$$

All other units ϵ'_D and all other solutions to (3.1) are powers of this ϵ_D. In order to develop a sum over all $h(D)$, as is Hooley's aim, we must develop a method for counting fundamental solutions to the Pellian equation.

(3.1) is a inhomogeneous equation in three variables, T, D, and U, the fundamental solutions of which we would like to sum subject to some conditions. Clearly one would like the condition $D \leq X$, however, there are still an infinite number of solutions to (3.1) in this range (due to derived solutions), so we must further restrict ourselves by insisting that $\epsilon_D \leq D^{\frac{1}{2}+\alpha}$. The $\frac{1}{2}$ in this condition reflects the lower bound on ϵ_D from the introduction. Further, to ensure that we do not obtain any trivial solutions from the case that D is a square, we also insist that $U \geq 1$. Hooley therefore defines

$$G(X,\alpha) = \sum_{\substack{D \leq X \\ \epsilon'_D \leq D^{1/2+\alpha}}} 1 \tag{3.2}$$

Where the sum is over all solutions to (3.1), fundamental or derived. Our first impulse may be to develop a method for finding ϵ_D for each D and then sum over the D with respect to the above conditions. Although ϵ_D can be found for arbitrary D by a method of continued fractions, this is a technique that does not lend itself to this problem.

Instead, we fix U and solve over possible solutions for T in the range prescribed by the conditions of the sum (3.2). We take $Y_1 = (A(U)U)^{\frac{1}{\alpha}}$, where $A(U)$ is a constant depending on U but obeying $2 < A(U) < 3$, $Y_2 = (Y_1 U^2 + 1)^{\frac{1}{2}}$ and $Y_3 = (xU^2+1)^{\frac{1}{2}}$. For $X_\alpha = \frac{1}{2}(X^\alpha - X^{-1-\alpha})$ we have,

$$G(X,\alpha) = \sum_{1 \leq U \leq X_\alpha} \sum_{\substack{DU^2 = T^2 - 1 \\ Y_2 \leq T \leq Y_3}} 1 \tag{3.3}$$

Taking the inside sum, we use the multiplicative structure of solutions once U has been fixed,

$$\sum_{\substack{DU^2=T^2-1\\ Y_2\le T\le Y_3}} 1 = \sum_{\substack{T^2-1\equiv 0(U^2)\\ Y_2\le T\le Y_3}} 1 = \sum_{\substack{\Omega^2-1\equiv 0(U^2)\\ 0<\Omega\le U^2}} \sum_{\substack{T\equiv\Omega(U^2)\\ Y_2\le T\le Y_3}} 1$$

$$= \left(\frac{Y_3 - Y_2}{U^2} + O(1)\right)\rho(U^2),$$

where $\rho(\lambda)$ is the number of distinct roots of $\Omega^2 \equiv 0 \mod \lambda$.

$\rho(U^2)$ is multiplicative and hence we can write a Dirichlet series for it and evaluate it via euler product. Perron's formula then yields an estimation for partial sums of $\rho(U^2)$, and we obtain Hooley's first

Theorem 3.1.1.

$$G(X,\alpha) = \frac{4\alpha^2}{\pi^2} X^{\frac{1}{2}} \log^2(X) + O(X^{\frac{1}{2}} \log(2X) + O(X^{\alpha} \log(2X))$$

uniformly for $X \ge 1$ *and* $0 \le \alpha \le 1$.

Because of the lower bound on ϵ_D, we have that all units with $\epsilon'_D < D$ must be fundamental. Hence, for $\alpha \le \frac{1}{2}$ we also have

Corollary 3.1.2.

$$\Gamma(X,\alpha) \sim \frac{4\alpha^2}{\pi^2} X^{\frac{1}{2}} \log^2(X), \tag{3.4}$$

where $\Gamma(X,\alpha)$ *is defined as* $G(X,\alpha)$ *except that we restrict ourselves to fundamental units.*

Clearly, the third term in the above equation is troublesome for all but very small α. It arises from the $O(1)$ term in the above computation of $G(X,\alpha)$. Also, $G(X,\alpha)$ has the defect of being a sum over all units as opposed to merely the fundamental ones. We are instead more interested in the sum $\Gamma(X,\alpha)$. Thus to get results which lend themselves more meaningfully to the conjecture we are aiming at, we take a slightly different approach.

3.2 A Second Method for Counting Units

In the previous section we held U fixed and then counted over T which gave solutions to the congruence $T^2 - 1 \equiv 0 \mod U^2$, relying on the multiplicative structure of $\rho(U^2)$ and the bounds Y_2 and Y_3. We now different approach for counting solutions to $T^2 - 1 \equiv 0 \mod U^2$.

Factoring the equation, we see that each solution to the above equation for odd U corresponds to a choice of U_1, U_2 such that

$$T - 1 \equiv 0 \mod U_1^2 \quad \text{and} \quad T + 1 \equiv 0 \mod U_2^2 \tag{3.5}$$

where

$$U = U_1 U_2, \qquad (U_1, U_2) = 1. \tag{3.6}$$

Because we have $(U_1, U_2) = 1$, each such factorization yields a solution to (3.5) and hence a unique solution to (3.1). This fact reflects the multiplicativity of our $\rho(\lambda)$ from the previous method. Given the relaxation of the bounds $Y_2 \le T \le Y_3$ to the bounds $T \le x^{\frac{1}{2}} U$ and $T > 1$, we impose the restrictions of factorizations $U = U_1 U_2$ that $U_1^2 \le T - 1 \le x^{\frac{1}{2}} U - 1$ and $U_2^2 \le T + 1 \le x^{\frac{1}{2}} U + 1$. Dropping the ± 1 from the above allows us to impose the conditions

$$U_1 \le x^{\frac{1}{2}} U_2 \quad \text{and} \quad U_2 \le x^{\frac{1}{2}} U_1. \tag{3.7}$$

We have a slightly modified construction for even U. This reflects the fact that $T^2 - 1 \equiv 0 \mod U^2$ has 2 solutions if U is a power of an odd prime but 4 solutions if U is 2^n for $n \ge 2$. Correspondingly, we choose

$$T - 1 \equiv 0 \mod 2U_1^2 \quad \text{and} \quad T + 1 \equiv 0 \mod 2U_2^2 \tag{3.8}$$

where

$$U = 2U_1U_2, \qquad (U_1, U_2) = 1. \tag{3.9}$$

And one solution to the above set of congruences yields one solution to $T^2 - 1 \equiv 0 \mod \frac{1}{2}U^2$. The have likewise that $2U_1^2 \le T - 1 \le 2x^{\frac{1}{2}}U - 1$ and $2U_2^2 \le T + 1 \le 2x^{\frac{1}{2}}U + 1$. This again yields the restrictions (3.7), however with a slightly different meaning for U_1, U_2.

In essence, what we have done here is developed a method of counting solutions to $T^2 - 1 \equiv 0 \mod U^2$ in terms of factorizations of U while at the same time imposing bounds on T. We have thus created an analogue for $\rho(\lambda)$ from the previous section. We define likewise $r(U^2)$ to be the number of solutions to (3.6) and (3.7) if U is odd, and twice the number of solutions to (3.9) and (3.7) if U is even. Therefore, we would like to investigate the sum

$$K(X, \alpha) = x^{\frac{1}{2}} \sum_{1 \le U \le X_\alpha} \frac{r(U^2)}{U} \tag{3.10}$$

to verify its resemblance to $G(x, \alpha)$. By splitting (3.10) into even and odd cases, and using the möbius function to remove the relatively prime condition we can evaluate (3.10). For the odd part we get

$$X^{\frac{1}{2}} \sum_{\substack{m \le X_\alpha^{\frac{1}{2}} \\ m\text{odd}}} \frac{\mu(m)}{m^2} \sum_{\substack{1 \le U_1U_2 \le \frac{X_\alpha}{m^2} \\ U_1 \le x^{\frac{1}{2}}U_2,\, U_2 \le x^{\frac{1}{2}}U_1 \\ U\text{odd}}} \frac{1}{U_1U_2} \sim \frac{1}{\pi^2}\left(\alpha - \frac{1}{4}\right) X^{\frac{1}{2}} \log^2(X), \tag{3.11}$$

where the asymptotic is obtained by integrating over the appropriate region in the U_1-U_2 plane. In similar fashion we also evaluate the even case and combining the two obtain

$$K(X,\alpha) \sim \frac{4}{\pi^2}\left(\alpha - \frac{1}{4}\right) X^{\frac{1}{2}} \log^2(X). \tag{3.12}$$

This does not seem to resemble Hooley's theorem 1. The discrepancy is due to the fact that this method of counting units misses the derived solutions. Indeed if we take $t_r + \sqrt{D}u_r = (t_s + \sqrt{D}u_s)^2$ where $r = 2s$, we have

$$t_r = t_s^2 + Du_s^2 \qquad \text{and} \qquad u_r = 2t_s u_s$$

giving

$$t_r - 1 = Du_s^2 \qquad \text{and} \qquad t_r + 1 = t_s^2.$$

This bares a striking resemblance to (3.8) and (3.9). Hence, we see that for derived solutions of even order we take $U_1 = u_s$ and $U_2 = t_s$. As a result, we have that $\sqrt{D}U_1$ is almost equal to U_2. Given this fact, we observe that restricting our D to range between $\frac{x}{\log^2(x)} \le D \le x$ will not affect the asymptotic estimate, and hence we have roughly

$$\frac{X^{\frac{1}{2}}U_1}{\log(X)} \le U_2 \le X^{\frac{1}{2}}U_1. \tag{3.13}$$

Hence, derived solutions of even order lie just within the region given by (3.7). If we sum over the region (3.13) instead of that prescribed by (3.7), one finds that the corresponding series $G'(X,\alpha)$ is $O(X^{\frac{1}{2}}\log(X)\log(\log(X)))$. From theorem 1, it is possible to show that for $\alpha > \frac{7}{2}$, $G'(X,\alpha)$ is bounded below by a function that is $O(X^{\frac{1}{2}}\log^2(X))$, and hence $K(X,\alpha)$ cannot possibly take account of derived solutions of even order.

A similar but much more complicated process shows that $K(X,\alpha)$ cannot take account of derived solutions of odd order either. It as well relies on the fact that removal of the region in which derived solutions are found will not affect the asymptotic estimate of $K(X,\alpha)$. Therefore, we have constructed a method of counting fundamental solutions alone.

3.3 Phase Transitions

One might expect that $\Gamma(X,\alpha)$ would be of the from of (3.12) for all $\alpha > 1$. However, a more detailed analysis shows that the form of this sum changes as α increases. $K(X,\alpha)$ actually excludes more than the derived solutions than we would like. It also serves to exclude fundamental solutions which are derived from cubes of solutions to the equation $t^2 - Du^2 = 4$. If we let $\eta_D = t + \sqrt{D}u$, then in the case that $D \equiv 5 \mod 8$ we have that

$$\epsilon_D = \left(\frac{1}{2}\eta_D\right)^3,$$

and the ϵ_D is fundamental. The counting process used to construct $K(X,\alpha)$ excludes these solutions for the same reason it excludes cubes of any ϵ_D. Let $\Gamma^*(X,\alpha)$ be the contribution to $\Gamma(X,\alpha)$ from D not of the above form. Let $\Delta(X,\alpha)$ be the number of non-square determinants D not exceeding X for which $\eta_D \leq 2D^{\frac{1}{2}+\alpha}$. We then have

$$\Gamma(X,\alpha) = \Gamma^*(X,\alpha) + \Delta\left(X, \frac{1}{3}(\alpha-1)\right).$$

It appears that no other types of derived solution affects the counting process. Derived solutions corresponding to solutions to the equations $t^2 - Du^2 = -1$ or ± 2 occur for so few discriminants as to not affect the asymptotic. Therefore, we expect that

$$\Gamma^*(X,\alpha) \sim K(X,\alpha) \sim \frac{4}{\pi^2}\left(\alpha - \frac{1}{4}\right) X^{\frac{1}{2}} \log^2(X)$$

By minor changes to theorem 1, we have

$$\Delta(X,\alpha) \sim \frac{1}{2}X^{\frac{1}{2}} \sum_{\substack{1 \le U \le X_\alpha \\ U \text{odd}}} \frac{\rho(U^2)}{U} \sim \frac{\alpha^2}{2\pi^2} X^{\frac{1}{2}} \log^2(X)$$

for $0 < \alpha \le \frac{1}{2}$, but in analogy to $K(X,\alpha)$ hypothesize that

$$\Delta(X,\alpha) \sim \frac{1}{2}X^{\frac{1}{2}} \sum_{\substack{1 \le U \le X_\alpha \\ U \text{odd}}} \frac{r(U^2)}{U} \sim \frac{1}{2\pi^2}\left(\alpha - \frac{1}{4}\right) X^{\frac{1}{2}} \log^2(X)$$

for $\alpha \ge \frac{1}{2}$ in analogy with $K(X,\alpha)$.

Seeing no other reasons for the form of $\Gamma(X,\alpha)$ to change, Hooley establishes the first of his main conjectures,

Conjecture 3.3.1. *For any given $\alpha > \frac{1}{2}$, we have*

$$\Gamma(X,\alpha) \sim B(\alpha) X^{\frac{1}{2}} \log^2 X,$$

as $X \longrightarrow \infty$, where

$$B(\alpha) = \begin{cases} \frac{4}{\pi^2}\left(\alpha - \frac{1}{4}\right), & \text{if } \frac{1}{2} < \alpha \le 1, \\ \frac{4}{\pi^2}\left(\alpha - \frac{1}{4}\right) + \frac{1}{18\pi^2}(\alpha - 1)^2, & \text{if } 1 \le \alpha \le \frac{5}{2}, \\ \frac{4}{\pi^2}\left(\alpha - \frac{1}{4}\right) + \frac{1}{6\pi^2}\left(\alpha - \frac{7}{4}\right) & \text{if } \alpha > \frac{5}{2}. \end{cases} \tag{3.14}$$

3.4 Leap of Faith

In what precedes, Hooley's arguement is not rigorous largely because he assumes but cannot prove that there are more phase transitions for $\alpha > \frac{5}{2}$. However, in formulating his conjecture for the asymptotic on class numbers, this assumption is minor in comparison to his next.

At the end of this analysis, once we have introduced the sum over $h(D)$, we would like to be able to remove the condition on the size of units $\epsilon_D \le D^{\frac{1}{2}+\alpha}$. However, we cannot do this by merely allowing $\alpha \to \infty$ in the above conjecture because for any α fixed, there are infinitely many units with $\epsilon_D > D^\alpha$. However,

$$\limsup_{D\to\infty} \frac{\log(\log(\epsilon_D))}{\log(D)} = \frac{1}{2}$$

so that if we replace $D^{\frac{1}{2}+\alpha}$ in the definition of $\Gamma(X,\alpha)$ with a function of the type $f(D) = e^{D^\beta}$ we can hope to encompass all units by letting $\beta \to \frac{1}{2}$ at a later time.

To assume that his previous methods hold with such rapidly increasing functions of D is Hooley's major leap of faith in his heuristic. If $f(D)$ is some smooth increasing function we define $\Gamma_f(X)$ to be the number of fundamental solutions to $T^2 - DU^2 = 1$ for which $D < x$ and $\epsilon_D \le f(D)$. We define $g(D)$ with an eye towards $\Delta(X,\alpha)$ to be $g(D) = (2D^{-1}f(D))^{\frac{1}{3}}$, and f^{-1} and g^{-1} to be the inverse functions of f and g respectively. Hence, we are led to the formula

$$\begin{aligned}
\Gamma_f(X) &\sim X^{\frac{1}{2}} \sum_{U \le f(X)} \frac{r(U^2)}{U} + \frac{1}{2} \sum_{\substack{U \le g(X) \\ U odd}} \frac{r(U^2)}{U} \\
&- \sum_{U \le f(X)} \frac{r(U^2)(f^{-1}(U))^{\frac{1}{2}}}{U} - \frac{1}{2} \sum_{\substack{U \le g(X) \\ U odd}} \frac{r(U^2)(g^{-1}(U))^{\frac{1}{2}}}{U} \\
&\sim \frac{4}{\pi^2} X^{\frac{1}{2}} \log(X) \log(f(X)X^{-\frac{1}{4}}) + \frac{1}{2\pi^2} X^{\frac{1}{2}} \log(X) \log(g(X)X^{-\frac{1}{4}}) - \Sigma
\end{aligned} \tag{3.15}$$

where the two terms with f^{-1} and g^{-1} derive from the next lower order term. For functions f which increase more rapidly than any power of D, these terms affect the asymptotic, but if f is bounded by some power law, they may be ignored. These terms derive from the $\frac{-Y_3}{U^2}\rho(U^2)$ term in the formula preceding theorem 3.1.1. For $f(D) = e^{D^\beta}$, we can actually compute Σ and find that

$$\Gamma_f(X) \sim \frac{25}{6\pi^2(1+2\beta)} X^{\frac{1}{2}+\beta} \log(X). \tag{3.16}$$

3.5 From Units to $h(D)$

With a machinery for summing over fundamental units in hand, we are able to proceed directly to the sum $S(X)$. We do this by attaching a weight to each fundamental unit. Define $P(X,\alpha)$ the analogous sum to $\Gamma(X,\alpha)$, with weight $L_D(1)$. Define $P_f(X)$ analogous to $\Gamma_f(X)$ and if $f(D) = e^{D^\beta}$, define $P^*_\beta(X) = P_f(X)$. The calculation of $P(X,\alpha)$ is long and technical, so we omit most of the details, which can be found in Hooley's paper. As the important ideas in Hooley's heuristic have occurred previously, we instead go for the general ideas.

When $\alpha < \frac{1}{2}$, we begin

$$P(X,\alpha) = \sum_{1\le U\le X_\alpha} \sum_{\substack{DU^2=T^2-1 \\ Y_2\le T\le Y_3}} L_D(1).$$

Next, taking $Y_4 = X^{\frac{1}{4}+\delta}$, with δ small, we have the estimate

$$L_D(1) = \sum_{\substack{m\le Y_4 \\ m \text{ odd}}} \left(\frac{D}{m}\right)\frac{1}{m} + O(X^{-\frac{\delta^2}{2}}).$$

We have then

$$P(X,\alpha) = \sum_{1\le U\le X_\alpha} \sum_{\substack{DU^2=T^2-1 \\ Y_2\le T\le Y_3}} \sum_{\substack{m\le Y_4 \\ m \text{ odd}}} \left(\frac{D}{m}\right)\frac{1}{m} + O(X^{-\frac{\delta^2}{2}}\Gamma(X,\alpha)).$$

Taking

$$I^*(Y_2,Y_3,U) = \sum_{\substack{DU^2=T^2-1 \\ Y_2\le T\le Y_3}} \sum_{\substack{m\le Y_4 \\ m \text{ odd}}} \left(\frac{D}{m}\right)\frac{1}{m},$$

we have

$$P(X,\alpha) = \sum_{1\le U\le X_\alpha} I^*(Y_2,Y_3,U) + O(X^{\frac{1}{2}-\frac{\delta^2}{2}}\log^2(2x))$$

$$= P^*(X,\alpha) + O(X^{\frac{1}{2}}\log(2X))$$

by the first section of chapter 3 and theorem 1. Next compute $I^*(Y_2, Y_3, U)$.

$$I^*(Y_2,Y_3,U) = \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{1}{m} \sum_{\substack{DU^2=T^2-1\\ Y_2\le T\le Y_3}} \left(\frac{D}{m}\right)$$

$$= \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{1}{m} \sum_{\substack{\Omega^2-1\equiv 0(U^2)\\ 0<\Omega\le U^2}} \sum_{Y_2\le \Omega+nU^2\le Y_3} \left(\frac{U^2n^2+2\Omega n+\frac{\Omega^2-1}{U^2}}{m}\right)$$

$$= \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{1}{m} \sum_{\substack{\Omega^2-1\equiv 0(U^2)\\ 0<\Omega\le U^2}} \sum_{0<l\le m} \left(\frac{U^2l^2+2\Omega l+\frac{\Omega^2-1}{U^2}}{m}\right) \sum_{\substack{Y_2U^{-2}-\Omega U^{-2}\le n\le Y_3U^{-2}-\Omega U^{-2}\\ n\equiv l \mod m}} 1$$

$$= \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{1}{m} \sum_{\substack{\Omega^2-1\equiv 0(U^2)\\ 0<\Omega\le U^2}} \left(\frac{Y_3-Y_2}{U^2m} \sum_{0<l\le m} \left(\frac{U^2l^2+2\Omega l+\frac{\Omega^2-1}{U^2}}{m}\right) + O(m)\right)$$

$$= \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{1}{m} \sum_{\substack{\Omega^2-1\equiv 0(U^2)\\ 0<\Omega\le U^2}} \left(\frac{Y_3-Y_2}{U^2m}\Psi(m;U,\Omega) + O(m).\right) \tag{3.17}$$

$\Psi(m;U,\Omega)$ is multiplicative in m, therefore to compute the above sum it suffices to compute it on prime powers. Checking cases reveals that $\Psi(m;U,\Omega)$ actually has no Ω dependence, and hence we re-designate it $\Psi(m;U)$. We hence reduce the above to

$$I^*(Y_2,Y_3,U) = \frac{(Y_3-Y_2)\rho(U^2)}{U^2} \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{\Psi(m;U)}{m^2} + O(\rho(U^2)\sum_{m\le Y_4} 1)$$

$$= \frac{(Y_3-Y_2)\rho(U^2)}{U^2} \sum_{\substack{m\le Y_4\\ m \text{ odd}}} \frac{\Psi(m;U)}{m^2} + O(X^{\frac{1}{4}+2\delta}) \tag{3.18}$$

Using the euler product to compute the inner sum here yields a constant and multiplicative function of U which we designate B and $\Phi(U)$ respectively. We thus have

$$P(X,\alpha) = B \sum_{1 \le U \le X_\alpha} \frac{(Y_3 - Y_2)\rho(U^2)}{U^2} \Phi(U) + O(X^{\frac{1}{4}+2\delta}). \tag{3.19}$$

Computing the Dirichlet Series of $\rho(U^2)\Phi(U)$ and utilizing Perron's formula to obtain a partial sum, we finally have

$$P(X,\alpha) \sim BX^{\frac{1}{2}} \sum_{1 \le U \le X_\alpha} \frac{\rho(U^2)\Phi(U)}{U} \sim \frac{4\alpha^2}{\pi^2} X^{\frac{1}{2}} \log^2(X) \tag{3.20}$$

We now make the same leap of faith as in section 3.4 and hypothesize that

$$P^*_\beta(X) \sim \frac{25}{6\pi^2(1+2\beta)} X^{\frac{1}{2}+\beta} \log(X) = P^\dagger_\beta(X), \text{say.} \tag{3.21}$$

Using integration by parts, the Stieltjes integral, and the fact

$$\limsup_{D\to\infty} \frac{\log(\log(\epsilon_D))}{\log(D)} = \frac{1}{2},$$

we obtain

$$S(X) = \sum_{D \le X} \frac{2\sqrt{D} L_D(1)}{\log(\epsilon_D)} \sim 2 \int_{\frac{3}{2}}^{X} \int_0^{\frac{1}{2}} y^{\frac{1}{2}-\beta} \, dP^*_\beta(y).$$

Substituting $P^\dagger_\beta(X)$ for $P^*_\beta(X)$, we have

$$\frac{25}{3\pi^2} \int_{\frac{3}{2}}^{X} \int_0^{\frac{1}{2}} y^{\frac{1}{2}-\beta} \frac{\partial^2 P^\dagger_\beta(y)}{\partial\beta\partial y} \, d\beta \, dy \sim \frac{25}{6\pi^2} \int_{\frac{3}{2}}^{X} \int_0^{\frac{1}{2}} \log^2(y) \, d\beta \, dy \sim \frac{25}{12\pi^2} X \log^2(X). \tag{3.22}$$

We therefore have

Conjecture 3.5.1. *As $X \to \infty$, we have*

$$\sum_{D \le X} h(D) \sim \frac{25}{12\pi^2} X \log^2(X).$$

Chapter 4

Conclusion

The main idea in Hooley's heuristic is to construct a sum that counts fundamental units ϵ_D. We obtain this sum with some bound on ϵ_D, by summing solutions of $T^2 - DU^2 = 1$ in U and T as opposed to summing over D. Hooley proves his results for only very restrictive bounds on ϵ_D. However, by assuming that the form of his sum holds for very large bounds, he collects all fundamental units in his sum. Weighting each one by the class number associated with that discriminant D provides a heuristic for $S(X)$.

Because of the delicate nature of his techniques, and his unjustified assumptions, one might at first be skeptical of this result. There could easily be an extra set of fundamental solutions that our counting process excludes by accident, or our second method for counting fundamental units could change significantly when ϵ_D is allowed to grow exponentially. However, the numerical results contained in this paper lend quite a bit of confidence to his formerly unchecked conjecture.

There are multiple ways in which I would like to further develop the work presented in this senior thesis. First, because our numerics necessitated a study of lower order terms in the asymptotic expansion of $S(X)/X$, one would like to be able to derive these lower order terms from the theory contained in Hooley's paper. If one were able to predict these terms based on theory alone, it would lend significant confidence to his conjecture. Moreover, If the exact asymptotic to 3 terms was known, we would be able to more meaningful conclusions about the error terms $o(x)$ from section 2.2.3.

Secondly, Hooley's work is defined for classical discriminants, which only represent half of the possible standard discriminants. One would also like to use the same technique as Hooley to predict an asymptotic for the slightly more general case of $s(x)$, and see it agree with the numerics presented for that case herein. Doing so would formulate the important conjecture in the modern terminology, establishing it for the work of future generations of mathematicians.

Bibliography

[Apo] T. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1976.

[Coh] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[Dav] H. Davenport, *Multiplicative Number Theory*, Springer, Berlin, Heidelberg, New York, 2000.

[Dic] L.E. Dickson, *Introduction to the Theory of Numbers*, University of Chicago Press, Chicago, 1929.

[Gau] C.F. Gauss, *Disquisitiones Arithmeticae*, 1801; English Translation, Yale University Press, New Haven, Conn, 1966.

[Hoo] C. Hooley, *On the Pellian equation and the class number of indefinite binary quadratic forms*, Jnl. Reine Angew Math. **353** (1984), 98-131.

[Kwo] S. Kwon, *On the Average of Class Numbers*, Undergraduate Thesis, Princeton University, 2006.

[PARI] PARI/GP, version `2.3.1`, Bordeaux, 2005, `http://pari.math.u-bordeaux.fr/`.

[Sie] C. L. Siegel, *The Average Measure of Quadratic Forms with Given Determinant and Signature*, Annals of Mathematics **45** (1944), 667-685.