Computing in Quadratic Orders

John Robertson

November 27, 2009

Please direct comments, corrections, or questions to jpr2718@gmail.com

1 Introduction

When studying binary quadratic forms and quadratic number fields and their orders, it is often helpful to be able to compute examples. This essay gives reasonably simple methods for various standard computations with binary quadratic forms and ideals in quadratic orders. It is not intended as an introduction to the theory of binary quadratic forms (for which see [22, 16, 27, 5, 6, 10, 8]), or quadratic number fields (see [42, 40, 8, 6, 21, 27, 26, 31, 36, 10]). Also, [37] has a "popular" exposition of many topics discussed here, and [25] covers the equivalence of class numbers based on quadratic forms and based on ideals. Steve Finch has collected many conjectures related to topics herein in [13].

Algorithms given here will tend to be simple, and may or may not generalize to number fields of higher degree.

2 Background and notation

The *binary quadratic forms* (BQFs) we are interested in are

$$g(x, y) = ax^2 + bxy + cy^2 \tag{1}$$

for integers x and y with *discriminant* $\Delta = b^2 - 4ac \neq 0$, not a square. This last condition implies that neither a nor c is zero. Forms will be written as any of $g = g(x, y) = ax^2 + bxy + cy^2 = (a, b, c)$ (we will reserve "f" for another use).

Note that if Δ is a square, then $g = \frac{1}{4a}[(2ax + by)^2 - \Delta y^2]$ factors into linear terms, resulting in a substantially different theory. When Δ is not square, then g does not factor into linear terms.

If $\Delta < 0$ then $ac > 0$ so a and c have the same sign. We will require that $a > 0$ and $c > 0$ when $\Delta < 0$. That is, in the case of negative discriminants we will only consider *positive definite* forms. The situation when $a < 0$ and $c < 0$ is completely analogous.

We will also consider quadratic number fields, i.e., $\mathbf{Q}(\sqrt{d})$ for squarefree integers d, $d \neq 0, 1$. An *algebraic integer* is a root of a monic polynomial with integral coefficients. Swinnerton-Dyer [43, pp. 1-2] delineates reasons that it is natural to call these algebraic integers. The *ring of algebraic integers* (or *integers*) $\mathcal{O}$ of $\mathbf{Q}(\sqrt{d})$ is the $\mathbf{Z}$–module with basis $[1, \omega]$ where

$$\omega = \begin{cases} \sqrt{d} & \text{if } d \equiv 2 \text{ or } 3 \pmod 4, \\ \dfrac{1 + \sqrt{d}}{2} & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

(that is, $\mathcal{O} = \{a \cdot 1 + b \cdot \omega\}$ where $a, b \in \mathbf{Z}$). The ring $\mathcal{O}$ is the intersection of the set of all algebraic integers with $\mathbf{Q}(\sqrt{d})$.

The *conjugate* $\overline{\omega}$ of ω is

$$\overline{\omega} = \begin{cases} -\sqrt{d} & \text{if } d \equiv 2 \text{ or } 3 \pmod 4, \\ \dfrac{1 - \sqrt{d}}{2} & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

General definitions of an *order* in number fields are given in [43, p. 3] and [10, p. 133]. For quadratic number fields, the *order* $\mathcal{O}_f$ of *conductor* f within the ring of integers $\mathcal{O}$ is the $\mathbf{Z}$–module with basis $[1,\ f\omega]$, where ω is as above (and $\mathcal{O} = \mathcal{O}_1$) [8, pp. 83-84].

The *discriminant* Δ of $\mathcal{O}_f$ is

$$\Delta = \begin{cases} f^2 d & \text{if } d \equiv 1 \pmod 4, \\ 4f^2 d & \text{if } d \equiv 2 \text{ or } 3 \pmod 4. \end{cases}$$

Clearly $\Delta \equiv 0$ or $1 \pmod 4$, and Δ is not a square. It is not hard to see that any integer $\Delta \equiv 0$ or $1 \pmod 4$, not a square, is the discriminant of a unique order. To find the d and f that generate a particular Δ find the largest $j \geq 1$ so that j^2 divides Δ. Set $d = \Delta/j^2$. If $d \equiv 1 \pmod 4$ then take $f = j$. Otherwise $d \equiv 2$ or $3 \pmod 4$, and take $f = j/2$. **We will refer to these as the d and f that correspond to Δ.**

When $f = 1$, Δ is a *fundamental discriminant*, sometimes denoted Δ_0.

When $d \equiv 1 \pmod 4$, $\omega^2 = \omega + (d-1)/4$, $\omega + \overline{\omega} = 1$, and $\omega\overline{\omega} = (1-d)/4$.

When $d \not\equiv 1 \pmod 4$, $\omega^2 = d$, $\omega + \overline{\omega} = 0$, and $\omega\overline{\omega} = -d$.

The *norm* $N(x + yf\omega)$ of an element $x + yf\omega \in \mathcal{O}_f$ is the element times its conjugate, $N(x + yf\omega) = (x + yf\omega)(x + yf\overline{\omega})$. When $d \equiv 1 \pmod 4$,

$$N(x + yf\omega) = x^2 + fxy + f^2 \left(\frac{1-d}{4}\right) y^2,$$

and when $d \not\equiv 1 \pmod 4$,

$$N(x + yf\omega) = x^2 - f^2 d y^2.$$

An *ideal* I in an order $\mathcal{O}_f$ (or any ring) is a set closed under addition and subtraction so that if $\alpha \in I$ and $r \in \mathcal{O}_f$, then $r\alpha \in I$. One way of writing an ideal I in an order $\mathcal{O}_f$ ($I \subset \mathcal{O}_f$) is to list two basis elements, $I = [t + uf\omega, v + wf\omega]$ (so $I = \{r(t + uf\omega) + s(v + wf\omega) | r, s \in \mathbf{Z}\}$). Alternatively, this can be written as the matrix

$$\begin{pmatrix} t & v \\ u & w \end{pmatrix}.$$

The *norm* of the ideal is $|tw - uv|$.

Every ideal $I \subset \mathcal{O}_f$ can be written uniquely as

$$I = \begin{pmatrix} t & v \\ 0 & w \end{pmatrix}$$

where $t > 0$, $0 \leq v < t$, $w > 0$, $w|t$, and $w|v$ (see [5, p. 94] or the section "Hermite Normal Form," below). Any rational integer in I is a multiple of t.

As will be expanded upon below, there are connections between forms and orders with the same discriminant.

We will make very little mention of *fractional ideals*, because any fractional ideal can be written as rI where r is a rational number and I is an ideal. Methods for working with ideals are easily extended to fractional ideals.

3 Standard number theory routines

Without further discussion, we assume the availability of the following standard routines: compute the least nonnegative residue of a modulo m; compute greatest common divisor $g > 0$ of two integers x, y and find r, s so that $rx + sy = g$; determine whether an integer is a prime; factor integers into primes; determine whether an integer is squarefree; determine whether an integer is a square; find the largest square that divides a integer; list all divisors of an integer; compute the Legendre, Jacobi, and Kronecker symbols; list all primes up to a given limit; sort arrays; solve linear equations $ax + by = c$ in integers; apply the Chinese Remainder Theorem (CRT); solve equations $x^2 \equiv a \pmod{m}$; efficient computation of $a^k \pmod{m}$ by the power algorithm; efficient computation of $ab \pmod{m}$. Cohen [6] discusses most of these. (The CRT can be used as part of a method to solve $x^2 \equiv a \pmod{m}$.)

We give algorithms for ideal multiplication and addition that require the MLLL basis reduction algorithm. But we also give algorithms that do not require the MLLL algorithm. A good MLLL routine is given in [32, pp. 209-210]. Actually, the only routine needed here is one that takes as input three

vectors of length 2 and returns two vectors of length 2 that have the span of the original three (in the notation of [32], $n = k = 2$). The routine should also check whether two of the input vectors are dependent, and return the right output when this occurs.

4 Hermite Normal Form

Computations with ideals presented here will often require putting the bases for ideals into Hermite Normal Form (HNF). Let the basis for an ideal be $[t + uf\omega, v + wf\omega]$. We desire a basis of the form $[t', v' + w'f\omega]$ with $t' > 0$, $w' > 0$, and $0 \leq v' < t'$. If $u = 0$ then just make sure t and w are positive (change the sign of t, or change the signs of w and v if needed), and select v' as the least nonnegative residue of v modulo t. (If $w = 0$, switch the roles of t, u and v, w, and adjust as in the previous sentence.) If $u \neq 0$ and $w \neq 0$, let $g > 0$ be the GCD of u and w, and let r, s be such that $ru + sw = g$. Take $t' = |(tw - uv)|/g$, $v' = tr + vs$, and $w' = g$. Replace v' with the least nonnegative residue of v' modulo t'.

We write the ideal $[t' + 0 \cdot f \cdot \omega, v' + w'f\omega]$ in HNF as an array

$$\begin{pmatrix} t' & v' \\ 0 & w' \end{pmatrix}.$$

Thus, our HNF arrays are in upper triangular form (some authors take HNF to be lower triangular).

Cohen [6, pp. 66ff] discusses reduction to HNF for more general arrays than the 2×2 and 2×4 arrays that arise in quadratic orders.

5 Transformations between forms

A common way to write a transformation between forms is to write a matrix T that gives the relationship between the *variables*, with the coefficients of the transformed form following as a consequence. Specifically, suppose we

have a form (1) in variables x, y, and a transformation between these variables and the variables x', y' given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = T \begin{pmatrix} x' \\ y' \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Then

$$\begin{aligned} & ax^2 + bxy + cy^2 \\ &= a(\alpha x' + \beta y')^2 + b(\alpha x' + \beta y')(\gamma x' + \delta y') + c(\gamma x' + \delta y')^2 \\ &= a'x'^2 + b'x'y' + c'y'^2 \end{aligned}$$

where

$$\begin{aligned} a' &= a\alpha^2 + b\alpha\gamma + c\gamma^2 \\ b' &= 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta \\ c' &= a\beta^2 + b\beta\delta + c\delta^2 \end{aligned}$$

This defines a transform T from the form $g = (a, b, c)$ to the form $gT = g' = (a', b', c')$.

Note that this is "backwards" from what you might expect, in that the transformation takes (x', y') to (x, y) and not the other way around.

If T_1 takes g to g' and T_2 takes g' to g'', then T_1T_2 (regular matrix product) takes g to g'', i.e., $(gT_1)T_2 = g(T_1T_2)$.

If $\det(T) = \pm 1$ then there is an inverse transformation for T, given by the inverse of the matrix representing T.

6 Equivalence and reduction

The natural notion of equivalence for forms is that two forms g, h are *equivalent* if there is a transformation

$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \tag{2}$$

taking $g = (a,\, b,\, c) = g(x, y) = ax^2 + bxy + cy^2$ to $h = (A,\, B,\, C) = h(X, Y)$ $= AX^2 + BXY + CY^2$ where

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix},$$

$$A = a\alpha^2 + b\alpha\gamma + c\gamma^2,$$

$$B = 2a\alpha\beta + b\alpha\delta + b\beta\gamma + 2c\gamma\delta, \text{ and}$$

$$C = a\beta^2 + b\beta\delta + c\delta^2,$$

and $\det(T) = \pm 1$. Any two forms g and h related by such a transformation represent the same numbers (the form g *represents* m if there are x, y so that $g(x, y) = m$). If $\det(T) = +1$, then g and h are said to be *properly equivalent*, $g \approx h$, while if $\det(T) = -1$, then g and h are said to be *improperly equivalent*. The forms g and h are equivalent, $g \sim h$, if g and h are either properly or improperly equivalent. Proper equivalence is an equivalence relation, as is equivalence. Equivalent forms have the same discriminant. We will say that a matrix is *unimodular* if it has integer entries and $\det(T) = +1$ [16, p. 168] (note that some authors allow unimodular transforms T to have $\det(T) = \pm 1$).

Many computations involving forms are aimed at determining whether two forms are equivalent or properly equivalent. A concept that is often useful here is that of a *reduced form*. Given a form, one way to find a properly equivalent reduced form involves use of a *reduction step*, which is a particular transformation of one form to another. We will define reduced forms, give the reduction step, and show how to use the reduction step to find a reduced form equivalent to a given form. We will use all of this to give a test to determine whether two forms are equivalent, and find a suitable transformation between them when they are equivalent.

7 Reduction step for $\Delta < 0$

A form (a, b, c) with $\Delta < 0$ is said to be *reduced* if

$$|b| \leq a \leq c$$

and if either $|b| = a$ or $a = c$ then $b \geq 0$.

The *reduction step* for positive definite forms is as follows. Given a form (a, b, c) we will produce a (not necessarily distinct) form (a', b', c') and the transformation T between these forms.

If (a, b, c) is reduced, take $a' = a$, $b' = b$, $c' = c$ and

$$T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i.e., do nothing.

Otherwise, if $a > c$ then take $a' = c$, $b' = -b$, $c' = a$ and

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

If the above conditions did not apply, look at

$$k = \left\lfloor \frac{a - b}{2a} \right\rfloor.$$

If $k \neq 0$ then take $a' = a$, $b' = 2ak + b$, $c' = ak^2 + bk + c$, and

$$T = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$$

Finally, if none of the above applied, we must have $b < 0$ and $a = c$ in which case take $a' = a$, $b' = -b$, $c' = c = a$ and

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Given a form, to find an equivalent reduced form, apply the reduction step repeatedly until a reduced form results. The composition of the transforms T gives the transform from the original form to the reduced form.

8 Reduction step for $\Delta > 0$

A form (a, b, c) with $\Delta > 0$ is said to be reduced if

$$|\sqrt{\Delta} - 2|a|| < b < \sqrt{\Delta}.$$

An equivalent set of conditions is

$$0 < b < \sqrt{\Delta} \text{ and } \sqrt{\Delta} - b < 2|a| < \sqrt{\Delta} + b.$$

In either of the above sets of conditions, replacing a with c gives an equivalent set of conditions.

Cohen [6, p. 258] gives the following reduction step. Recall that neither a nor c can be zero for any forms we consider.

First define $r(-b, c)$ to be the unique integer r so that $r \equiv -b \pmod{2|c|}$ and $-|c| < r \leq |c|$ if $|c| > \sqrt{\Delta}$, or $\sqrt{\Delta} - 2|c| < r < \sqrt{\Delta}$ if $|c| < \sqrt{\Delta}$.

For the reduction step, take $a' = c$, $b' = r(-b, c)$, and $c' = ((b')^2 - \Delta)/4c$. The transform T is

$$T = \begin{pmatrix} 0 & 1 \\ -1 & -(b + b')/2c \end{pmatrix}.$$

Given a form, to find an equivalent reduced form, apply the reduction step repeatedly until a reduced form is produced. The composition of the transforms T gives the transform from the original form to the reduced form.

In the case of $\Delta < 0$, when a reduced form is input into the reduction step, the same form is output. When $\Delta > 0$, the output form is always different from the input form, even when the input form is reduced. If the input form is reduced, the output form will also be reduced.

There are always at least two reduced forms in any proper equivalence class of forms when $\Delta > 0$. To find all reduced forms in an equivalence class, begin with any form in that class. If that form is not reduced, apply the reduction step until a reduced form is produced. From here, repeated application of the reduction step will produce all reduced forms in the class,

and will eventually produce the original reduced form again. When the original reduced form is produced a second time, all equivalent reduced forms have been produced. The set of reduced forms in an equivalence class is called the *period* of reduced forms for the class.

9 All reduced forms of discriminant Δ

If $\Delta < 0$ and (a, b, c) is reduced, then $b^2 - 4ac = \Delta < 0$ and $|b| \leq a \leq c$. Then $a^2 \geq b^2$ and $4ac \geq 4a^2$, so $3a^2 = 4a^2 - a^2 \leq 4ac - b^2 = |\Delta|$, and $|a| \leq \sqrt{|\Delta|/3}$. Also, $a > 0$. So, it suffices to search on $0 < a \leq \sqrt{|\Delta|/3}$, $-a \leq b \leq a$, $b \equiv \Delta \pmod 2$. Cohen [6, p. 228] gives a more efficient algorithm based on this principle.

If $\Delta > 0$ and (a, b, c) is reduced, then $b^2 - 4ac = \Delta > 0$ and $|\sqrt{\Delta} - 2|a|| < b < \sqrt{\Delta}$. It follows that $-\sqrt{\Delta} < a < \sqrt{\Delta}$. So it suffices to search on $-\sqrt{\Delta} < a < \sqrt{\Delta}$, $|\sqrt{\Delta} - 2|a|| < b < \sqrt{\Delta}$, $b \equiv \Delta \pmod 2$.

When $\Delta > 0$, the forms can be arranged into periods through use of the reduction step.

10 Determine whether forms are properly or improperly equivalent

Given two forms g_1, g_2, with the same discriminant Δ, to determine whether they are properly equivalent, first find reduced forms $h_1 \approx g_1$, and $h_2 \approx g_2$.

If $\Delta < 0$, the original forms are properly equivalent if and only if $h_1 = h_2$.

In the case $\Delta > 0$, g_1 and g_2 are properly equivalent if and only if h_1 is in the period of reduced forms of h_2 (in which case, both periods are the same).

To see whether two forms are improperly equivalent, apply any transformation of determinant -1 to one form, and then test whether the transformed form is properly equivalent to the other form.

When two forms are properly equivalent, a unimodular transformation

between them can be found by composing the transforms (or inverses of the transforms) generated during the reduction steps. As an example, we find a unimodular transform between the properly equivalent forms $g_1 = (1, 3, 1)$ and $g_2 = (5, 5, 1)$, neither of which is reduced. Applying the reduction step to g_1 results in the reduced form $h_1 = (1, 1, -1)$, and the transform $T_1 = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$, for which $g_1 T_1 = h_1$. Also, applying the reduction step to g_2 results in the reduced form $h_2 = (1, 1, -1)$, and the transform $T_2 = \begin{pmatrix} 0 & 1 \\ -1 & -3 \end{pmatrix}$, for which $g_2 T_2 = h_2$. Since $h_1 = h_2$, setting $T = T_1 T_2^{-1}$, so $T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, we have $g_1 T = g_2$.

Note that two forms can be both properly and improperly equivalent. As an example, $(26, 55, -1)$ and $(26, 49, -7)$ are properly and improperly equivalent under the transforms

$$\begin{pmatrix} -15 & 2 \\ 7 & -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}.$$

If two forms are properly and improperly equivalent, any two forms in the same proper class of forms are properly and improperly equivalent. In particular, each form in that proper equivalence class is improperly equivalent to itself. For example $(26, 55, -1)$ is improperly equivalent to itself under the transform

$$\begin{pmatrix} -15 & 32 \\ 7 & -15 \end{pmatrix}$$

If a proper class of forms includes a form improperly equivalent to itself, the class is an *ambiguous class*.

A form (a, b, c) is *ambiguous* if $a|b$. An ambiguous form is improperly equivalent to itself, so any class with an ambiguous form is ambiguous. If $\Delta > 0$, every ambiguous class has exactly four ambiguous forms of the form $(a, 0, c)$ or (a, a, c) and exactly two reduced ambiguous forms. If $\Delta < 0$,

every ambiguous class has exactly two ambiguous forms of the form $(a, 0, c)$ or (a, a, c); at most one of these is reduced.

11 Solving $ax^2 + bxy + cy^2 = m$—the roadmap.

When the equation

$$ax^2 + bxy + cy^2 = m \tag{3}$$

has solutions, those solutions can be divided into equivalence classes. These equivalence classes, in turn, are related to automorphs (transformations that take a form to itself). So, we first have a section on equivalence classes of solutions, and then one on automorphs. Finally, we give a method of solving these equations, which actually just makes sure we find at least one solution in each equivalence class.

12 Equivalent solutions

There is a natural way that solutions to a BQF equation can be divided into equivalence classes so that any solution falls into one of a finite number of classes.

Before getting to that, note the formula

$$(rt \pm Dsu)^2 - D(ru \pm ts)^2 = (r^2 - Ds^2)(t^2 - Du^2).$$

This can be used to generate new solutions to the Pell equation $r^2 - Ds^2 = N$ from a given solution and a solution to the Pell equation $t^2 - Du^2 = 1$. If $r^2 - Ds^2 = N$ and $t^2 - Du^2 = 1$, then $(rt \pm Dsu)^2 - D(ru \pm ts)^2 = N \cdot 1 = N$. For (t, u) a solution to the "+1" Pell equation, the solutions (r, s) and $((rt \pm Dsu), (ru \pm ts))$ are considered to be equivalent.

This idea generalizes to BQF equations, essentially through the use of the formula

$$ax^2 + bxy + cy^2 = \frac{1}{4a}[(2ax + by)^2 - \Delta y^2].$$

In particular, if (x', y') is a solution to (1) and (t, u) is any solution to $t^2 - \Delta u^2 = 4$, then (x, y) is also a solution to (1) where x and y are defined by

$$x = \left(\frac{t - bu}{2}\right) x' - cuy', \quad y = aux' + \left(\frac{t + bu}{2}\right) y', \tag{4}$$

or

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{t-bu}{2} & -cu \\ au & \frac{t+bu}{2} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

In turn, this is equivalent to

$$2ax + by + y\sqrt{\Delta} = \left(\frac{t + u\sqrt{\Delta}}{2}\right)(2ax' + by' + y'\sqrt{\Delta}).$$

The two solutions (x, y), (x', y') are said to be *equivalent*. It is not difficult to verify that this defines an equivalence relation.

An equivalent, but often easier, test is that two solutions (x, y), (x', y') are equivalent if and only if

$$2axx' + b(xy' + x'y) + 2cyy' \equiv 0 \pmod{|N|}, \text{ and}$$

$$xy' - x'y \equiv 0 \pmod{|N|}.$$

Solutions to the Pell equations $t^2 - \Delta u^2 = \pm 4$ are related to units in quadratic orders with discriminant Δ. This relationship is discussed in more detail in the section "Norm equations and units" below. See [39] for methods to solve these equations.

For $\Delta > 0$, all solutions (t, u) to $t^2 - \Delta u^2 = 4$ are generated from the solution (t_1, u_1) with the minimum positive t_1 and u_1 by

$$\frac{t + u\sqrt{\Delta}}{2} = (\pm 1)\left(\frac{t_1 + u_1\sqrt{\Delta}}{2}\right)^n$$

for some integer n (positive, zero, or negative), so it is easy, using (4), to generate all solutions to (3) equivalent to a given solution.

There are a finite number of equivalence classes of solutions to (3) (possibly zero). See [23] for more on the structure of solutions.

13 Automorphs

An *automorph* of a form g is a unimodular transform T (recall that unimodular means that $\det(T) = 1$) that maps g to itself, $gT = g$. If T is an automorph of $g = ax^2 + bxy + cy^2$, then

$$T = \begin{pmatrix} \frac{t-bu}{2} & -cu \\ au & \frac{t+bu}{2} \end{pmatrix}$$

for some t, u that satisfy the Pell equation $t^2 - \Delta u^2 = 4$ [12, p. 112] [22, p. 95] [16, p. 194].

14 Solving $ax^2 + bxy + cy^2 = m$

This will give the "traditional" method of solving the BQF equation $g(x, y) = ax^2 + bxy + cy^2 = m$, given a, b, c, and m. When $\Delta > 0$, a superior method is given in [23] and [38].

The basic principle is as follows. If x, y is a primitive solution (i.e., $\gcd(x, y) = 1$) to $g(x, y) = m$ then for r, s so that $xr - ys = 1$, the transform

$$T = \begin{pmatrix} x & s \\ y & r \end{pmatrix},$$

maps g to a form $g' = mx^2 + b'xy + c'y^2$ with x^2-coefficient m, and the solution $g'(1, 0) = m$ is mapped to (x, y),

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & s \\ y & r \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

So, the method is as follows. First find all forms $g' = mx^2 + b'xy + c'y^2$ of discriminant Δ with $0 \le b' < 2|m|$ (see comment below). These are the forms with a b' for which $c' = (b'^2 - \Delta)/(4m)$ is an integer. For small $|m|$ these b' are quickly determined by brute-force search. For large $|m|$, there are more efficient methods to solve $b'^2 \equiv \Delta \pmod{4|m|}$ [6, pp. 31-36, 44].

Second, for each form g' found in the first step, test to see whether it is properly equivalent to g. For each g' that is properly equivalent, use any transform T for which $gT = g'$ to map the solution $(1, 0)$ to $g' = m$ to a solution (x, y) to $g = m$.

At this point, we have one primitive solution from each class of solutions. To get non-primitive solutions, apply the above to every equation $g = m/j^2$ for every $j > 1$ so that j^2 divides m. Multiply the solutions found by j so that they are solutions to the original equation.

To find all solutions, apply the automorphs of g to the solutions found so far.

A word on why it suffices to only look at b' so that $0 \leq b' < 2|m|$. With r, s so that $xr - ys = 1$, as above, consider the transforms

$$\begin{pmatrix} x & s + kx \\ y & r + ky \end{pmatrix},$$

all of which have determinant $+1$. For these transforms,

$$\begin{aligned} b' &= 2ax(s + kx) + b(x(r + ky) + y(s + kx)) + 2cy(r + ky) \\ &= 2axs + b(xr + ys) + 2cyr + 2k\left(ax^2 + bxy + cy^2\right) \\ &= 2axs + b(xr + ys) + 2cyr + 2km. \end{aligned}$$

By picking k suitably, we can force $0 \leq b' < 2|m|$.

References for this method include [16, Chapter 6], [22, Chapter III], and [5, p. 53].

15 Examples of solutions of binary quadratic form equations

As an example for $\Delta > 0$, we solve

$$55x^2 + 315xy + 451y^2 = -3971. \tag{5}$$

Primitive Solutions

b'	c'	T	x	y
807	-41	$\begin{pmatrix} 1617 & -164 \\ -562 & 57 \end{pmatrix}$	1617	-562
2251	-319	$\begin{pmatrix} -1359 & 385 \\ 473 & -134 \end{pmatrix}$	-1359	473
5691	-2039	$\begin{pmatrix} 1562 & -1119 \\ -543 & 389 \end{pmatrix}$	1562	-543
7135	-3205	$\begin{pmatrix} 1327 & -1192 \\ -462 & 415 \end{pmatrix}$	1327	-462

Table 1: Primitive solutions to $55x^2 + 315xy + 451y^2 = -3971$

As $\Delta = 5$, we start with solutions to $b'^2 \equiv 5 \pmod{4 \cdot 3971}$, $0 \le b' < 2 \cdot 3971$, which are $b' = 807,\ 2251,\ 5691,\ 7135$. For $b' = 807$, $c' = -41$, and the form $g_1 = (55,\ 315,\ 451)$ is equivalent to the form $g_2 = (-3971,\ 807,\ -41)$ under the transform

$$T = \begin{pmatrix} 1617 & -164 \\ -562 & 57 \end{pmatrix}$$

giving the solution $g(1617,\ -562) = -3971$. The table "Primitive Solutions" gives a primitive solution corresponding to each b', along with the corresponding c' and the transform T from g_1 to $(-3971,\ b',\ c')$.

As $-3971 = -11 \cdot 19^2$, to find imprimitive solutions it suffices to solve

$$55x^2 + 315xy + 451y^2 = -11. \tag{6}$$

Solutions to $b'^2 \equiv 5 \pmod{4 \cdot 11}$, $0 \le b' < 2 \cdot 11$, are $b' = 7,\ 15$. For $b' = 7$,

$c' = -1$, a solution to (6) is $(x,\ y) = (95,\ -33)$, and an imprimitve solution to (5) is $(x,\ y) = (1805,\ -627)$. For $b' = 15$, $c' = -5$, a solution to (6) is $(x,\ y) = (121,\ -42)$, and an imprimitve solution to (5) is $(x,\ y) = (2299,\ -798)$.

The minimal positive solution $(t,\ u)$ to

$$t^2 - 5u^2 = 4 \tag{7}$$

is $(t,\ u) = (3,\ 1)$, from which all solutions to (7) can be generated, and so all solutions to (5). For instance, $(t,\ u) = (3,\ 1)$ gives the automorph

$$T = \begin{pmatrix} -156 & -451 \\ 55 & 159 \end{pmatrix}$$

and $T \begin{pmatrix} 1617 \\ -562 \end{pmatrix} = \begin{pmatrix} 1210 \\ -423 \end{pmatrix}$. Other solutions equivalent to $(1617,\ -562)$ include $(3641,\ -1263)$, and $(2013,\ -707)$.

It happens in this example that none of the solutions found are those with minimal positive y among all solutions for the class. A method that does find the solution with minimal positive y in each class when $\Delta > 0$ is given in [23] and [38].

As an example for $\Delta < 0$, we solve

$$x^2 + 3xy + 3y^2 = 7. \tag{8}$$

As $\Delta = -3$, we start with solutions to $b'^2 \equiv -3 \pmod{4 \cdot 7}$, $0 \le b' < 2 \cdot 7$, which are $b' = 5,\ 9$. For $b' = 5$, $c' = 1$, and the form $g = (1,\ 3,\ 3)$ is equivalent to the form $g' = (7,\ 5,\ 1)$ under the transform

$$\begin{pmatrix} 4 & 1 \\ -1 & 0 \end{pmatrix}$$

giving the solution $g(4, -1) = 7$. There are six solutions to $t^2 + 3u^2 = \pm 4$ (see next section), namely $(t, u) = (\pm 2, 0), (\pm 1, \pm 1)$. These give the automorphs

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & -3 \\ 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 \\ -1 & -1 \end{pmatrix} \quad \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix},$$

so the complete list of solutions equivalent to $(4, -1)$ is $(4, -1)$, $(-4, 1)$, $(-1, 2)$, $(5, -3)$, $(-5, 3)$, $(1, -2)$.

For $b' = 9$, $c' = 3$, and the form $g = (1,\ 3,\ 3)$ is equivalent to the form $g' = (7,\ 9,\ 3)$ under the transform

$$\begin{pmatrix} 5 & 3 \\ -2 & -1 \end{pmatrix}$$

giving the solution $g(5, -2) = 7$. The other solutions equivalent to this one are $(-5, 2)$, $(1, 1)$, $(4, -3)$, $(-4, 3)$, $(-1, -1)$.

As 7 is squarefree, there are no imprimitive solutions.

This last example is a little unusual in that for $\Delta < -4$, there are only 2 solutions in each equivalence class of solutions, not 6.

16 Norm equations and units

Finding all elements of a quadratic order with a given norm is equivalent to solving certain binary quadratic equations. For a given Δ, with associated d and f, suppose we want to find all elements $x + yf\omega$ of norm m, i.e., $N(x + yf\omega) = m$. Consider first the case where $d \equiv 1 \pmod 4$. Then $N(x + yf\omega) = (x + yf\omega)(x + yf\overline{\omega}) = x^2 + fxy + f^2((1 - d)/4)y^2$. So, the equation we want to solve is

$$x^2 + fxy + f^2((1 - d)/4)y^2 = m. \tag{9}$$

Sometimes it is helpful to multiply through by 4 and complete the square to convert the equation to the generalized Pell equation

$$(2x + fy)^2 - f^2dy^2 = 4m,$$

or

$$X^2 - \Delta Y^2 = 4m, \tag{10}$$

where $X = 2x + fy$ and $Y = y$. Because $d \equiv 1 \pmod 4$, for any solution to (10), $X \equiv fY \pmod 2$, and every solution to (10) yields a solution $x = (X - fY)/2$, $y = Y$ to (9).

If f is even, so $4|\Delta$, it can be easier to solve the equation

$$X^2 - (\Delta/4)Y^2 = m$$

where $X = x + (f/2)y$ and $Y = y$.

If $d \not\equiv 1 \pmod 4$ then $N(x + yf\omega) = (x + yf\omega)(x + yf\overline{\omega}) = x^2 - f^2dy^2$, and the equation to solve is

$$x^2 - f^2dy^2 = m.$$

Units in an order of a quadratic number field are elements of norm ± 1.

First consider $d \equiv 1 \pmod 4$. Then $x + yf\omega$ is a unit if

$$x^2 + fxy + f^2\left(\frac{1-d}{4}\right)y^2 = \pm 1. \tag{11}$$

As above, multiplying through by 4, taking $X = 2x + fy$, $Y = y$, and noting that $\Delta = f^2d$, (11) is

$$X^2 - \Delta Y^2 = \pm 4. \tag{12}$$

For every solution of (12) $x = (X - fY)/2$ is an integer, so every solution of (12) gives a solution of (11).

For $\Delta < 0$, only the equation

$$X^2 - \Delta Y^2 = +4 \tag{13}$$

has solutions, and (13) has finitely many in this case. If $d = -3$ and $f = 1$, so $\Delta = -3$, the solutions (X, Y) to (13) are $(\pm 2, 0)$, and $(\pm 1, \pm 1)$, which correspond to solutions (x, y) of $(\pm 1, 0)$, $(0, \pm 1)$, and $(\pm 1, \mp 1)$ to (11). It is easy to see that the six solutions $z = x + y\omega$ to (11) are the six complex solutions to $z^6 = 1$. For any other $\Delta < 0$ (with $d \equiv 1 \pmod 4$), the only

solutions (X, Y) to (13) are $(\pm 2, 0)$, which correspond to solutions (x, y) of $(\pm 1, 0)$, to (11).

For $\Delta > 0$,

$$X^2 - \Delta Y^2 = -4 \tag{14}$$

might have solutions, and in any event, (12) always has infinitely many solutions. For all solutions $X + Y\sqrt{\Delta}$ to (12), $(X + Y\sqrt{\Delta})/2 = \pm 1((X_1 + Y_1\sqrt{\Delta})/2)^n$, where $X_1 + Y_1\sqrt{\Delta}$ is the minimal positive solution to (12) and n is an integer (possibly zero or negative). See [39] for further information.

Now consider $d \not\equiv 1 \pmod 4$. Then $x + yf\omega$ is a unit if

$$x^2 - f^2 d y^2 = \pm 1. \tag{15}$$

Again, by multiplying through by 4, taking $X = 2x$, $Y = y$, and noting that $\Delta = 4f^2 d$, (15) can be written

$$X^2 - \Delta Y^2 = \pm 4. \tag{16}$$

While these equations are equivalent, equation (15) gets at units more directly, so we suggest solving it when units are wanted and $d \not\equiv 1 \pmod 4$.

For $\Delta < 0$, only the equation

$$x^2 - f^2 d y^2 = +1 \tag{17}$$

has solutions, and (17) has finitely many in this case. If $d = -1$, and $f = 1$, so $\Delta = -4$, the solutions (x, y) to (17) are $(\pm 1, 0)$, and $(0, \pm 1)$. It is easy to see that the four solutions $z = x + y\omega$ to (17) are the four complex solutions to $z^4 = 1$. For any other $\Delta < 0$ (with $d \not\equiv 1 \pmod 4$), the only solutions (x, y) to (17) are $(\pm 1, 0)$.

For $\Delta > 0$,

$$x^2 - f^2 d y^2 = -1 \tag{18}$$

might have solutions, and in any event, (15) always has infinitely many solutions. For all solutions $x + yf\sqrt{d}$ to (15), $x + yf\sqrt{d} = \pm 1 \cdot (x_1 + y_1 f\sqrt{d})^n$, where $x_1 + y_1 f\sqrt{d}$ is the minimal positive solution to (15) and n is an integer (possibly zero or negative). See [39] for further information.

17 Genus equivalence of forms

Loosely speaking, two forms with the same discriminant are in the same *genus* if they represent "some of the same numbers." More precisely, if $\Delta = f^2 d$ for $d \equiv 1 \pmod 4$, d squarefree, or $\Delta = 4f^2 d$ for $d \equiv 2$ or $3 \pmod 4$, d squarefree (as above), then two forms g_1 and g_2 are in the same genus if and only if there are integers $t_1^2 m$ and $t_2^2 m$ so that g_1 primitively represents $t_1^2 m$, g_2 primitively represents $t_2^2 m$, m is squarefree, and $\gcd(t_1, f) = \gcd(t_2, f) = 1$ [8, pp. 144-145].

As an example of the need for the "gcd" condition, consider the forms $g_1(x, y) = x^2 + 45xy - 9y^2$ and $g_2(x, y) = -x^2 + 45xy + 9y^2$. These forms have discriminant $2061 = 3^2 \cdot 229$. These forms are in different genuses (apply the test given below), but both represent $405 = 3^2 \cdot 5$: $g_1(3, 4) = g_2(12, 1) = 405$.

Alternatively, two forms of the same discriminant Δ are in the same genus if and only if they represent the same values in $(\mathbf{Z}/(\Delta\mathbf{Z})^*$, the non-zero-divisors in $(\mathbf{Z}/(\Delta\mathbf{Z})$ [10, p. 53].

A way to determine whether two forms are in the same genus is through the use of characters. For our purposes, a *character* is a function $\chi : \mathbf{Z}/m\mathbf{Z} \to \{-1, 0, +1\}$ with $m > 0$ so that

$$\chi(r) = 0 \iff \gcd(r, m) > 1, \text{ and}$$

$$\gcd(r, s) = 1 \implies \chi(rs) = \chi(r)\chi(s).$$

More general characters are defined in [43, p. 125] and [26, p. 270].

For a given discriminant Δ, make a list of characters as follows [10, p. 55] [12, pp. 82-87] [8, p. 150]. Begin with the characters $\chi(m) = \left(\frac{m}{p}\right)$ for every odd (positive) prime p dividing Δ. Here $\left(\frac{m}{p}\right)$ is the Kronecker symbol. If Δ is odd or $\Delta/4 \equiv 1 \pmod 4$, this is the complete list of characters. Otherwise add characters according to the table "Additional characters".

To test whether two forms, g and g', are in the same genus, find m and m' so that g represents m, g' represents m', and $\gcd(m, 2\Delta) = \gcd(m', 2\Delta) = 1$ (that this can always be done is discussed in [38]). Then g and g' are

Additional characters

$\Delta/4 \pmod 8$	Additional Characters
2	$\left(\frac{2}{m}\right)$
3, 4, or 7	$\left(\frac{-1}{m}\right)$
6	$\left(\frac{-2}{m}\right)$
0	$\left(\frac{-1}{m}\right)$ and $\left(\frac{2}{m}\right)$

Table 2: Additional characters when Δ is even, $\Delta/4 \not\equiv 1 \pmod 4$

in the same genus if and only if the values for all characters in the list for Δ agree for m and m'.

For example, for the discriminant $\Delta = 10336 = 2^5 \cdot 17 \cdot 19$, there are four characters, $\chi_1(m) = \left(\frac{m}{17}\right)$, $\chi_2(m) = \left(\frac{m}{19}\right)$, $\chi_3(m) = \left(\frac{-1}{m}\right)$, and $\chi_4(m) = \left(\frac{2}{m}\right)$. The form $g_1 = (-84, 68, 17)$ represents 1 ($g_1(1,1) = -84 \cdot 1^2 + 68 \cdot 1 \cdot 1 + 17 \cdot 1^2 = 1$), and clearly 1 is relatively prime to $2 \cdot 10336$. So, for g_1 we have $\chi_1(1) = \chi_2(1) = \chi_3(1) = \chi_4(1) = 1$. The form $g_2 = (84, 68, -17)$ represents -1, ($g_2(1,-1) = 84 \cdot 1^2 + 68 \cdot 1 \cdot (-1) - 17 \cdot (-1)^2 = -1$). For g_2 we have $\chi_1(-1) = \chi_4(-1) = 1$, and $\chi_2(-1) = \chi_3(-1) = -1$. Thus g_1 and g_2 are in different genus classes. Also consider $g_3 = 9x^2 + 88xy - 72y^2$, which represents 9 with $x = 1$, $y = 0$. For g_3 we have $\chi_1(9) = \chi_2(9) = \chi_3(9) = \chi_4(9) = 1$, so g_1 and g_3 are in the same genus class. Note that g_3 does not represent 1, while g_1 does, and g_3 represents 353 ($g_3(43, -4) = 353$), while g_1 does not, so they do not represent exactly the same integers. But both do represent $273 = 3 \cdot 7 \cdot 13$ ($g_1(1,3) = g_3(3,1) = 273$), which alone would show they are in the same genus class. Note also that g_1 and g_3 are neither properly nor improperly equivalent. (Also, from material below, you can see that g_1 and g_2 correspond to weakly equivalent ideal classes, but are in different genus classes.)

Here's another use of characters. If p is an odd prime, $p \nmid \Delta$, m is represented by some form in a given genus, $\gcd(m, 2\Delta) = 1$, and $\chi_i(p) =$

$\chi_i(m)$ for every character associated with the discriminant Δ, then some form in the given genus represents p [37, p. 140]. This does not hold for $p = 2$ or for composites. For example, if $d = 577$, $f = 1$, so $\Delta = 577$, then the only character is $\chi(m) = \left(\frac{m}{577}\right)$. The form $F(x, y) = x^2 + xy - 144y^2$ is in the principal genus, and clearly represents 1. Here $\chi(1) = 1 = \chi(2)$. But no form in the principal genus represents 2. Similarly, for $d = 229$, $f = 1$, $\Delta = 229$, there is one character and $\chi(1) = 1 = \chi(91) = \left(\frac{91}{229}\right)$ but no form in the principal genus represents 91.

18 Composition of forms

This algorithm is given in [37, p. 135].

The *composition* of two forms, $g = (a, b, c)$ and $g' = (a', b', c')$, of the same discriminant Δ, is as follows. Let h be the gcd of a, a', and $(b + b')/2$. Let u, v, w be any integers so that $au + a'v + ((b + b')/2)w = h$ (note that u, v, and w are not unique). Set

$$a'' = \frac{aa'}{h^2}$$

$$b'' = \frac{1}{h}\left(aub' + a'vb + \frac{(bb' + \Delta)}{2}w\right)$$

$$c'' = \frac{b''^2 - \Delta}{4a''}$$

Look at as a map from pairs of proper equivalence classes to proper equivalence classes.

Composition of forms corresponds to multiplication of ideals, discussed below.

19 Intermission

So far, routines discussed have been for forms. We now abruptly begin to discuss routines for ideals in modules. An early goal is to give formulas for maps between forms and ideals. Before getting to that, it is necessary

to cover a few routines for ideals. Many of the later routines for ideals require mapping the ideals to forms, doing a computation with the forms, and mapping back to ideals.

20 Is a module an ideal?

The question is whether the **Z**–module $[t + uf\omega, v + wf\omega]$ is an ideal in the order $[1, f\omega]$.

First, let $d \equiv 1 \pmod 4$. The **Z**–module is an ideal of the order if and only if $tw - uv$ divides each of the following

$$uwf^2\left(\tfrac{d-1}{4}\right) - v(t + uf)$$

$$u^2f^2\left(\tfrac{d-1}{4}\right) - t(t + uf)$$

$$w^2f^2\left(\tfrac{d-1}{4}\right) - v(v + wf)$$

The proof of this is as follows. We need any element of the order $[1, f\omega]$ times any element of the **Z**–module $[t + uf\omega, v + wf\omega]$ to be an element of the module. It suffices to check that $f\omega$ times $t + uf\omega$ is an element of the module, and that $f\omega$ times $v + wf\omega$ is an element of the module. Writing

$$f\omega(t + uf\omega) = r(t + uf\omega) + s(v + wf\omega)$$

and solving for r and s (note that $\omega^2 = (d-1)/4 + \omega$) gives the first two conditions above. Doing the same for the second product gives the third condition, and a condition equivalent to the first condition.

In the important special case where $u = 0$, the conditions above reduce to w divides t, w divides v, and

$$\frac{t}{w} \text{ divides } f^2\left(\frac{d-1}{4}\right) - \frac{v}{w}\left(\frac{v}{w} + f\right).$$

Mollin [28, Theorem 1.2.1, p. 9] states this equivalently as if $I = [t, v + w\omega]$, then I is a nonzero ideal of $\mathcal{O}_f$ if and only if $w|t$, $w|v$, and $tw|N(v+w\omega)$.

If both $u = 0$ and $w = 1$, then the conditions reduce to t dividing $f^2(d-1)/4 - v(v+f)$.

When $d \equiv 2$ or $3 \pmod 4$ the formulas are similar, but a little simpler. The $\mathbf{Z}$–module is an ideal of the order if and only if $tw - uv$ divides each of the following

$$uwf^2d - tv$$

$$u^2f^2d - t^2$$

$$w^2f^2d - v^2$$

Here, when $u = 0$, the conditions above reduce to w divides t, w divides v, and

$$\frac{t}{w} \text{ divides } f^2d - \left(\frac{v}{w}\right)^2.$$

If both $u = 0$ and $w = 1$, then the conditions reduce to t dividing $f^2d - v^2$.

Mollin's Theorem 1.2.1 [28, p. 9] gives conditions for a module to be an ideal.

Per [8, p. 140], a module $\mathcal{M}$ is *primitive* if no rational integer $t > 1$ can divide every element of $\mathcal{M}$. A primitive module $\mathcal{M}$ is an ideal of $\mathcal{O}_f$ exactly when it has the canonical form $[t, v + f\omega]$, $0 < t$, $0 \le v < t$, $0 < f$, and where t, the norm $[\mathcal{O}_f : \mathcal{M}]$, divides $N(v + f\omega)$.

21 Multiplication and addition of ideals

The product IJ of the ideals $I = [\alpha_1, \alpha_2]$ and $J = [\beta_1, \beta_2]$ is the $\mathbf{Z}$-module $[\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \alpha_2\beta_2]$. To reduce this to a two-element basis, you can write it as a 2×4 array (each $\alpha_i\beta_j$ is a column), and reduce that to HNF. We discuss this in more detail below, with a shortcut for the case where at least one of I or J is invertible.

It's a little overkill, but you can use the MLLL algorithm twice to reduce IJ to a two-element basis, namely feed in three of the $\alpha_i\beta_j$ to the MLLL

algorithm, get two vectors as output; feed these and the fourth $\alpha_i\beta_j$ into the MLL algorithm, and the output is a two element basis for IJ.

A similar algorithm can be used to add ideals. The sum of the ideals $I = [\alpha_1, \alpha_2]$ and $J = [\beta_1, \beta_2]$ is the $\mathbf{Z}$-module $[\alpha_1, \alpha_2, \beta_1, \beta_2]$. As with multiplication of ideals, this just needs to be reduced to HNF to have a two-element basis.

When Δ is large, faster ways to multiply ideals are given in [46]. These methods give fast methods to multiply ideals, because the product of the ideals $I = [\alpha_1, \alpha_2]$ and $J = [\beta_1, \beta_2]$ is the sum of the ideals $\alpha_1 J$ and $\alpha_2 J$.

Here's a more detailed walk-through of multiplication of two ideals. Every integral or fractional ideal can be written as $r[a, b+f\omega]$ where r is an integer or rational number, so it suffices to consider products of primitive ideals. Suppose we want to multiply $I = [a_1, b_1 + f\omega]$ and $J = [a_2, b_2 + f\omega]$. Then IJ is spanned by the columns of

$$A = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 \\ 0 & v_2 & v_3 & v_4 \end{pmatrix}$$

where $t_1 = a_1a_2$, $t_2 = a_1b_2$, $v_2 = a_1$, $t_3 = a_2b_1$, $v_3 = a_2$, and $t_4 + v_4 f\omega = (b_1 + f\omega)(b_2 + f\omega)$. Our goal is an array in HNF

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$$

that has the same span as the preceding array.

To this end, first find the greatest common divisor v of v_2, v_3, and v_4, and integers r_2, r_3, and r_4 such that $r_2v_2 + r_3v_3 + r_4v_4 = v$. Write $t = r_2t_2 + r_3t_3 + r_4t_4$ so we have

$$t + vf\omega = r_2(t_2 + v_2f\omega) + r_3(t_3 + v_3f\omega) + r_4(t_4 + v_4f\omega).$$

Now we get to the shortcut we can use when at least one of I, J is invertible. When at least one is invertible, we have $N(IJ) = N(I)N(J)$ ([6, p. 181]) so we can simply take $a = a_1a_2/v$, $c = v$, and b as the least nonnegative

residue of t modulo a. Then $IJ = [a, b + cf\omega]$. Because $ac = av = a_1a_2$, the norm of IJ is $N(I)N(J)$.

Otherwise, for $2 \le i \le 4$, set $s_i = v_i/v$ and $t'_i = t_i - s_it$. Set $a = \gcd(t_1, t'_2, t'_3, t'_4)$, and (as just above) $c = v$, and take b as the least nonnegative residue of t modulo a. Then

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$$

is in HNF and $IJ = [a, b + cf\omega]$.

Here are a few examples. Let $d = 2$, $f = 1$, $I = [7, 3+\omega]$, $J = [17, 6+\omega]$. Of course, $\omega = \sqrt{2}$. Because $(3 + \sqrt{2})(6 + \sqrt{2}) = 20 + 9\sqrt{2}$, our array A is

$$\begin{pmatrix} 119 & 42 & 51 & 20 \\ 0 & 7 & 17 & 9 \end{pmatrix}$$

We have $\gcd(7, 17, 9) = 1$ and we can pick $r_2 = 5$, $r_3 = -2$, $r_4 = 0$ so $5 \cdot 7 + (-2) \cdot 17 + 0 \cdot 9 = 1$. So $t = 108$ and $v = 1$. Because both ideals are invertible, the norm of their product is $7 \cdot 17 = 119$, so the product is

$$\begin{pmatrix} 119 & 108 \\ 0 & 1 \end{pmatrix}$$

For a second example, take $d = 5$, $f = 21$, $I = J = [3, f\omega]$, where $f\omega$ is

$$21\left(\frac{1+\sqrt{5}}{2}\right).$$

Neither I nor J is invertible. A is

$$\begin{pmatrix} 9 & 0 & 0 & 441 \\ 0 & 3 & 3 & 21 \end{pmatrix}$$

We have $\gcd(3, 3, 21) = 3$ and we can take $r_2 = 1$, $r_3 = r_4 = 0$, so $t = 0$. We have $s_2 = s_3 = 1$, $s_4 = 7$ so $t'_2 = t'_3 = 0$ and $t'_4 = 441$. As $\gcd(9, 0, 0, 441) = 9$, $IJ = [9, 3f\omega]$. Note that $N(IJ) = 27 > 9 = N(I)N(J)$.

As a final example, let $d = 5$, $f = 21$, $I = [3, f\omega]$, $J = [7, f\omega]$. Neither I nor J is invertible. A is

$$\begin{pmatrix} 21 & 0 & 0 & 441 \\ 0 & 3 & 7 & 21 \end{pmatrix}$$

We have $\gcd(3, 7, 21) = 1$ and we can take $r_2 = -2$, $r_3 = 1$, $r_4 = 0$, so $t = 0$. We have $s_2 = 3$, $s_3 = 7$, $s_4 = 21$ so $t'_2 = t'_3 = 0$ and $t'_4 = 441$. As $\gcd(21, 0, 0, 441) = 21$, $IJ = [21, f\omega]$. Note that $N(IJ) = 21 = N(I)N(J)$.

Williams [17, Sec. 5.4] and Mollin [28, pp. 10-11] [30, Sect 2.1, p. 59] give formulas for multiplying ideals under the assumption that at least one ideal is invertible. Cohen [6] discusses reducing arrays to HNF while minimizing the size of the numbers in the intermediate calculations on pages 66ff and multiplication of modules (including ideals) on pages 187 and 188.

22 Principal ideal generated by an element

The ideal generated by the element $t + uf\omega$ is $[t + uf\omega, (t + uf\omega)f\omega]$.

23 Equivalence classes of ideals

For ideals, the natural notion of equivalence is that two ideals I and J of $\mathcal{O}_f$ are equivalent $I \sim J$ if there are principal ideals (a) and (b) with $a,\ b \in \mathcal{O}_f$ so that $I(a) = J(b)$. There is also a notion of *strict equivalence* $I \approx J$, where we add the condition that $N(a/b) > 0$. The ideals are *weakly equivalent* if $N(a/b) < 0$.

Two ideals I and J are said to be in the same *genus class*, $I \approxeq J$, if there is an ideal K so that I and JK^2 are strictly equivalent. In particular, there is a genus class, called the *principal class*, that consists of the squares of ideals.

We will discuss methods for testing whether ideals are equivalent, and finding (a) and (b) when they are, after covering maps between classes of forms and classes of ideals.

24 Map form to ideal

The maps in this section and the next should not be looked at as maps between individual forms and individual ideals, but rather as maps between classes of forms and classes of ideals. The classes can be taken to be those of proper equivalence of forms and strict equivalence of ideals. In fact, narrower classes can be used. Forms can be considered equivalent here if there is a transform

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

that takes one to the other. Ideals can be considered equivalent if one is a rational multiple of the other [6, p. 221].

Given a form (a, b, c) of discriminant $\Delta = f^2 d$ or $4f^2 d$, set $\alpha = a$, $\gamma = 0$, and $\delta = 1$. If $d \equiv 1 \pmod 4$ then set $\beta = (-1)(b + f)/2$, otherwise set $\beta = -b/2$.

If $a < 0$ then multiply the ideal $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ by the ideal $(f\omega)$.

Reduce the result to HNF. If $\delta > 1$, divide out by δ. We write $\phi_{FI}(a, b, c)$ for (the strict class of) this ideal.

This map ϕ_{FI} is an isomorphism between the group of proper classes of forms, under composition, and strict classes of ideals, under multiplication [8, pp. 140-141] [6, pp. 220-225].

There are two immediate consequences of this map from forms to ideals. The forms (a, b, c) and $(-a, b, -c)$ are called negatives (?) of each other, and map to weakly equivalent classes. So, the class group, or weak class group, can be found without using ideals by joining classes with forms that are the negatives of each other.

Also, the forms (a, b, c) and $(a, -b, c)$ map to conjugate ideals, so inverse ideals.

25 Map ideal to form

See the section just above for discussion of equivalence of ideals and forms under this map.

Given d, f, the basis $[t+uf\omega, v+wf\omega]$ of an ideal is said to be *positively oriented* if $((v+wf\omega)(t+uf\overline{\omega})-(t+uf\omega)(v+wf\overline{\omega}))/\sqrt{\Delta}>0$. Note that if $t>0$, $u=0$, $v\geq 0$, and $w>0$, then the basis is positively oriented.

For d, f, we will map the ideal $I=\begin{pmatrix} t & v \\ u & w \end{pmatrix}$ with a positively oriented basis and a sign $s=\pm 1$ to a form. Set $m=s/(tw-uv)$. If $d\equiv 1 \pmod 4$,

$$\begin{aligned} a &= m\left(t^2+tuf+u^2f^2\left(\tfrac{1-d}{4}\right)\right), \\ b &= (-ms)\left(2tv+twf+vuf+2uwf^2\left(\tfrac{1-d}{4}\right)\right), \\ c &= m\left(v^2+vwf+w^2f^2\left(\tfrac{1-d}{4}\right)\right). \end{aligned}$$

If $d\not\equiv 1 \pmod 4$, then set

$$\begin{aligned} a &= m\left(t^2-u^2f^2d\right), \\ b &= (-2ms)\left(tv-uwf^2d\right), \\ c &= m\left(v^2-w^2f^2d\right). \end{aligned}$$

We write $\phi_{IF}(I)$ for (the proper class of) the form $g=(a,b,c)$ [8, pp. 140-141] [6, pp. 220-225].

26 Maps between ideals and quadratic irrationals

Similar to the maps above, these maps should be looked at as maps between equivalence classes of ideals and equivalence classes of quadratic irrationals. As above, ideals are equivalent if they are rational multiples of each other. Quadratic irrationals are equivalent, here, if they differ by a rational integer [6, p. 221].

Any number ξ that can be written as $\xi=(P+\sqrt{D})/Q$, with D not a square, is the root of and irreducible quadratic equation $ax^2+bx+c=0$,

with $a, b, c \in \mathbf{Z}$ and $ac \neq 0$. For example, take $a = Q^2$, $b = -2PQ$, and $c = P^2 - D$. We can, after canceling common factors if necessary, assume $\gcd(a, b, c) = 1$. Under this assumption, the discriminant associated with ξ is $\Delta = b^2 - 4ac$. In particular, any ξ that can be written, $(P + \sqrt{D})/Q$ can also be written so $P^2 \equiv D \pmod{2Q}$, for instance as $(-b + \sqrt{\Delta})/2a$.

With the above notation, if $I = [t + uf\omega, v + wf\omega]$ has an oriented basis, then

$$\phi_{IQ}(I) = \frac{v + wf\omega}{t + uf\omega}$$

gives a map from invertible ideals in the order of discriminant Δ to quadratic irrationals with the same discriminant [6, p. 221].

For $\tau = (-b + \sqrt{\Delta})/2a$ where $4a|\Delta - b^2$, set

$$\phi_{QI}(\tau) = a(\mathbf{Z} + \tau\mathbf{Z}).$$

As an example, consider the quadratic irrational, $\tau = (-4 + \sqrt{5})/11$. As the denominator is odd, we multiply top and bottom by 2, to get $\tau = (-8 + \sqrt{20})/22$, so $a = 11$, $b = -8$, and $44|(20 - 64)$. Also $\Delta = 20$, so $d = 5$ and $f = 2$. You can check that $11(-8 + \sqrt{20})/22 = -5 + 1 \cdot 2 \cdot \omega$, so τ maps to the ideal $[11, -5 + f\omega]$ in the order $\mathcal{O}_2$ of the ring of integers of $\mathbf{Q}(\sqrt{5})$.

27 Testing ideals for equivalence

Two ideals, I and J, are strictly equivalent, if and only if the forms $g = \phi_{IF}(I)$ and $h = \phi_{IF}(J)$ are properly equivalent.

The two ideals are weakly equivalent if and only if $I(f\omega)$ and J are strictly equivalent.

When I and J are strictly equivalent, there are several ways to find a and b so that $I(a) = J(b)$.

One method starts by letting $m = N(I)/g$, and $n = N(J)/g$, where $g = \gcd(N(I), N(J))$. Then for $k = 1, 2, 3, \ldots$, in turn, generate all elements a of norm kn (actually, take one element from each equivalence class of

solutions to $N(x) = kn$) and all elements b of norm km. See if for any pair $I(a) = J(b)$. If not, try the next k until a suitable pair is found.

If $I = (1)$ or $I = (f\omega)$, and $I \approx J$ then J is principal so it suffices to find a generator for J. This amounts to finding all $a \in \mathcal{O}_f$ for which $N(J) = |N(a)|$, and then determining which of these satisfies $(a) = J$. In turn, this amounts to either solving

$$x^2 + fxy + f^2\left(\frac{1-d}{4}\right)y^2 = \pm N(J)$$

if $d \equiv 1 \pmod 4$, or solving

$$x^2 - f^2 dy^2 = \pm N(J)$$

if $d \not\equiv 1 \pmod 4$.

Another alternative is to invert one ideal, multiply this inverse by the other ideal, and look for generator of the resulting principal ideal. If $(a) = I^{-1}J$, then $I(a) = J(1)$.

For $\Delta > 0$, if there are forms g and g' so that $g' = \rho^n(g)$, where ρ is the reduction step above, and if $I = \phi_{FI}(g)$ and $I' = \phi_{FI}(g')$, then there is an efficient algorithm for computing γ so that $(\gamma)I = I'$ [6, p. 276]. It suffices to show this for $n = 1$, as the γ's for successive steps can be multiplied together.

Let $g = (a, b, c)$, $g' = (a', b', c')$, $g' = \rho(g)$, $I = \phi_{FI}(g)$, and $I' = \phi_{FI}(g')$.

Let $\gamma_0 = (b+\sqrt{\Delta})/(2a)$. When $d \equiv 1 \pmod 4$, $\gamma_0 = \frac{1}{a}\left(\frac{b-f}{2} + f\omega\right)$. When $d \not\equiv 1 \pmod 4$, $\gamma_0 = \frac{1}{a}\left(\frac{b}{2} + f\omega\right)$.

If $aa' > 0$ then $\gamma = \gamma_0$.

If $d \equiv 1 \pmod 4$, $a < 0$, and $a' > 0$ then there is a factor of $(f\omega)$ in I that we need to remove. Here $(f\omega) = \left[f^2\left(\frac{d-1}{4}\right), f\omega\right]$, and $(f\omega)^{-1} = (-f + f\omega) = \left[f^2\left(\frac{d-1}{4}\right), -f + f\omega\right]$. Now,

$$\gamma = \gamma_0(-f + f\omega) = \frac{1}{a}\left(\frac{b-f}{2} + f\omega\right)(-f + f\omega)$$

$$= \frac{1}{a}\left(\left(\frac{f-b}{2}\right)f + f^2\left(\frac{d-1}{4}\right) + \left(\frac{b-f}{2}\right)(f\omega)\right)$$

If $d \equiv 1 \pmod 4$, $a > 0$, and $a' < 0$ then we need to add a factor of $(f\omega)$. So,

$$\gamma = \gamma_0(f\omega) = \frac{1}{a}\left(\frac{b-f}{2} + f\omega\right)(f\omega)$$
$$= \frac{1}{a}\left(f^2\left(\frac{d-1}{4}\right) + \left(\frac{b+f}{2}\right)(f\omega)\right)$$

When $d \not\equiv 1 \pmod 4$, note that $(f\omega)^{-1} = (f\omega) = [f^2 d, f\omega]$. If $aa' < 0$, we need to add a factor of $(f\omega)$, so

$$\gamma = \gamma_0(f\omega) = \frac{1}{a}\left(\frac{b}{2} + f\omega\right)(f\omega) = \frac{1}{a}\left(f^2 d + \frac{b}{2}(f\omega)\right).$$

28 Invertibility of Ideals

For d, f consider I an ideal in the order $\mathcal{O}_f$ (so we have $K = \mathbf{Q}(\sqrt{d}) \supset \mathcal{O} = \mathcal{O}_1 \supset \mathcal{O}_f \supset I$). The question is, is I invertible? That is, is there a fractional ideal I' of $\mathcal{O}_f$ so that $II' = \mathcal{O}_f = (1)$?

There are some relatively simple tests, and others that are more complex. Perhaps the simplest test, but one that does not work in all situations, is that if $\gcd(N(I), f) = 1$ then I is invertible. Said another way, every ideal relatively prime to (f) is invertible. (I is relatively prime to (f) if and only if $(f) + I = (1)$, which occurs if and only if $N(I)$ is relatively prime to f.) In particular, in the maximal order, i.e., when $f = 1$, every ideal is invertible. While there are invertible ideals that are not relatively prime to (f), every invertible ideal is strictly equivalent to an ideal that is prime to (f). (And there is unique factorization into primes for the set of ideals relatively prime to (f).)

Another simple test is that if

$$I = \begin{pmatrix} t & v \\ 0 & 1 \end{pmatrix}$$

is invertible, then the inverse I' of I is

$$I' = \begin{pmatrix} t & -v-f \\ 0 & 1 \end{pmatrix}.$$

when $d \equiv 1 \pmod 4$, and

$$I' = \begin{pmatrix} t & -v \\ 0 & 1 \end{pmatrix}.$$

when $d \not\equiv 1 \pmod 4$. So, multiply I by I' and see if the product is

$$(t) = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}.$$

If it is, I is invertible, otherwise I is not invertible. Any fractional ideal can be written as $r \begin{pmatrix} t & v \\ 0 & 1 \end{pmatrix}$ for some rational r, so it is easy to extend this test to arbitrary fractional ideals.

Another simple test is that I is invertible if and only if $N(I^2) = N(I)^2$.

Our last simple test, and perhaps the easiest to use, is that I is invertible if and only if the form $\phi_{IF}(I)$ (above) is primitive [6, p. 222]. Mollin [30, p. 260] notes that any primitive ideal can be written as $[a, (b+\sqrt{\Delta})/2]$ and this test is equivalent to determining whether $\gcd(a, b, (b^2-\Delta)/(4a))$ is 1. Equivalently, if the primitive ideal is $[a, b+\omega]$, then it is invertible if for $d \not\equiv 1 \pmod 4$,

$$1 = \gcd(a, 2b, (b^2 - f^2 d)/a)$$

or when $d \equiv 1 \pmod 4$

$$1 = \gcd(a, 2b+f, ((2b+f)^2 - f^2 d)/(4a)).$$

Here's a more complicated test, but one that gets more to the heart of the matter. We say the ideal I is *proper* if when $\beta \in K$ and $\beta I \subset I$, then $\beta \in \mathcal{O}_f$. The ideal I is invertible if and only if it is proper.

As an example, for $d = 229$, $f = 5$, $\omega = (1 + \sqrt{229})/2$, we will show directly that $I = [11, 1 + 5\omega]$ is proper. First suppose that $\beta I \subset I$ and note that this holds if and only if $11\beta \in I$, and $(1 + 5\omega)\beta \in I$. Let $11\beta = 11T + U(1 + 5\omega)$, where this last is some element of I. Then $\beta = (11T + U(1 + 5\omega))/11$, and we must have

$$\left(\frac{11T + U(1 + 5\omega)}{11}\right)(1 + 5\omega) = 11V + W(1 + 5\omega).$$

Multiplying this out, and noting that $\omega^2 = 57 + \omega$, we have

$$11T + U + 5U\omega + 55T\omega + 5U\omega + 25 \cdot 57U + 25U\omega = 121V + 11W + 55W\omega.$$

Collecting terms with and without ω we have the two equations

$$11T + 1426U = 121V + 11W,$$

and

$$11T + 7U = 11W.$$

Taking the difference, we readily deduce that $129U = 11V$, and, in particular, that U is divisible by 11. Letting $U/11 = X$, we have that $\beta = T + X(1 + 5\omega) = (T + X) + 5X\omega \in \mathcal{O}_5$. Hence I is proper.

For $d = 229$, $f = 5$, we will show that $I = [5, 5\omega]$ is not proper, where ω is as just above. Simply take $\beta = \omega$. Clearly $\beta \notin \mathcal{O}_5$, and yet $5\beta \in I$, and $5\omega\beta \in I$, as is easily seen. In fact, for any $\beta \in \mathcal{O}$, $\beta I \subset I$.

Now consider the ideals $[25, 5k + 5\omega]$ in $\mathcal{O}_5$. We will see which of these are proper. For any given k, if $\beta I \subset I$, then $25\beta \in I$, so $\beta = (25T + U(5k + 5\omega))/25$ for some T, U. Also

$$\left(\frac{25T + U(5k + 5\omega)}{25}\right)(5k + 5\omega) = 25V + W(5k + 5\omega)$$

for some V, W. This gives

$$(5T + U(k + \omega))(k + \omega) = 25V + 5W(k + \omega).$$

Multiplying out,

$$5Tk + 5T\omega + Uk^2 + 2Uk\omega + 57U + U\omega = 25V + 5Wk + 5W\omega.$$

Separating into two equations,

$$5Tk + Uk^2 + 57U = 25V + 5Wk$$

and

$$5T + 2Uk + U = 5W.$$

Multiply the second times k and subtract from the first to get

$$U(57 - (k^2 + k)) = 25V.$$

If $5 \nmid (57 - (k^2 + k))$ then $25|U$ and it is easy to see that $\beta \in \mathcal{O}_5$. On the other hand, if $5|(57-(k^2+k))$, then it is easy to find U that is not a multiple of 25, and for which $\beta \notin \mathcal{O}_5$. We conclude that I is proper for $k = 0, 2, 4$ and I is not proper for $k = 1, 3$. In fact, for $k = 1, 3$ we can use $\beta = \omega$ to show that I is not proper.

29 Inverse of $(f\omega)$

For $d \equiv 1 \pmod 4$, the inverse of $(f\omega) = [f^2\left(\frac{d-1}{4}\right), f\omega]$ in $\mathcal{O}_f$, is

$$(f\omega)^{-1} = (-f + f\omega) = \left[f^2\left(\frac{d-1}{4}\right), -f + f\omega\right].$$

For $d \not\equiv 1 \pmod 4$, the inverse of $(f\omega) = [f^2 d, f\omega]$ in $\mathcal{O}_f$ is itself

$$(f\omega)^{-1} = (f\omega) = [f^2 d, f\omega].$$

30 Relation between ideals in the maximal order and smaller orders

If I is an ideal of $\mathcal{O}$, then $I' = I \cap \mathcal{O}_f$ is an ideal of $\mathcal{O}_f$. If $I = [t, v + \omega]$ then I' is invertible in $\mathcal{O}_f$ if and only if $\gcd(t, f) = 1$.

If $I = [t, v + f\omega]$ is an ideal of $\mathcal{O}_f$ and $\gcd(t, f) = 1$, then there is a v' so that $I = [t, fv' + f\omega]$, $I' = [t, v' + \omega]$ is an ideal of $\mathcal{O}$, and $I = I' \cap \mathcal{O}_f$. Also, I is invertible.

If $I = [t, v + f\omega]$ is an ideal of $\mathcal{O}_f$ and $\gcd(t, f) \neq 1$, then either I is not invertible, or there is no $I' \subset \mathcal{O}$ so that $I = I' \cap \mathcal{O}_f$, or both. In the following examples "I'" is $I' = [t, v' + \omega] \subset \mathcal{O}$ so that $I = I' \cap \mathcal{O}_f$, if there is such. Each I is an ideal of the order $\mathcal{O}_6 \subset \mathbf{Q}(\sqrt{229})$.

For $I = [4, 6\omega]$, I is invertible and there is no I'.

For $I = [3, 6\omega]$, I is not invertible and $I' = [3, \omega] \subset \mathcal{O}$.

For $I = [2, 6\omega]$, I is not invertible and there is no I'.

Every invertible ideal in $\mathcal{O}_f$ is strictly equivalent to an ideal relatively prime to f.

See [8, p. 144] for a discussion of ideal equivalence modulo f. For I, $J \subset \mathcal{O}$, $I' = I \cap \mathcal{O}_f$, and $J' = J \cap \mathcal{O}_f$, we have that I' and J' are equivalent in $\mathcal{O}_f$ if there are α, $\beta \in \mathcal{O}$ so that $\alpha I = \beta J$, $\gcd(\alpha\beta, f) = 1$, and $\alpha \equiv z\beta \pmod{f}$ for some $z \in \mathbf{Z}$. For strict equivalence of I' and J', we also need $N(\alpha/\beta) > 0$.

31 Computing the order of an ideal in the class group

The *order of an ideal* I in the strict class group is the smallest n so that

$$I^n \approx (1).$$

We start with a very simple strategy for finding the order of an element in the strict class group. First make a table with an ideal of smallest norm in each strict equivalence class. To do this, first generate all reduced forms, and if $\Delta > 0$, group the forms into periods. This gives the classes. Then generate

all ideals of norm $1, 2, 3, 4, \ldots$. For quadratic number fields, if $\Delta > 0$ then every class has an ideal I with $N(I) < \sqrt{\Delta/5}$, and if $\Delta < 0$ then every class has an ideal I with $N(I) < \sqrt{|\Delta|/3}$ [8, Cor. 13.18]. Test each such ideal against each class that does not yet have an ideal assigned to it. To test an ideal I apply the transformation $g = \phi_{IF}(I)$, find a reduced form g' equivalent to g, and then see if g' is equal to any of the forms in the period for the class being tested. Continue until every class has an ideal assigned.

To find the order of an ideal I, start with $n = 1$ and $J = I$, and see if $J \approx (1)$. If it is, then the order of I is 1. If it is not, do $n = n + 1$, $J = JI$. If now $J \approx (1)$ then the order of I is n. Otherwise, continue the iteration until $J \approx (1)$. To keep the precision needed for the computations low, at each step replace J with the ideal of smallest norm in the same class as J.

Cohen [6, pp. 235-264] gives more sophisticated methods for computing the class group, and so the orders of elements in the group, when $h_+(\Delta)$ is large. A newer method due to Terr [44, 4] is as follows. Write the group operation multiplicatively. Set $A_0 = B_0 = 1$, and $A_1 = B_1 = g$, where 1 is the unit element for the group and g is the group element. Then take $A_k = gA_{k-1}$, $B_k = A_k B_{k-1}$. At each step, see whether $B_k = A_j$ for some $0 \le j \le k$. If $B_k = A_j$, then the order of the element is $k(k+1)/2 - j$. If the order of g is i, then this method finds that order in at most k steps where $k(k-1)/2 < i \le k(k+1)/2$.

32 Computing the class group

We give a simple method for computing the strict class group that makes use of the fact that this group is abelian, and so is a product of subgroups $\mathbf{Z}/p^n\mathbf{Z}$ for primes p. We assume that a table has been created with an element of each class, and its order in the class group.

At each step, we will test a certain element of the class group, "checking off" elements as we go along. First, check off the identity element (the element of order 1). For each prime p that divides the order of the class

group we do some tests, in turn. Given a p, make a list of all elements of order p^i for some i, sorted in decreasing order of i. At each step, pick the next element a in the list that has not been checked off. Say a is of order p^i. For each of a, a^2, a^3, ..., a^{p^i-1}, multiply by all elements that have been checked off, and tentatively check off the product. If a product was previously checked off, or tentatively checked off, undo all the tentative check marks, and move on to the next a. If no product was checked off, add a to the list of generators, and check off all the elements tentatively checked off.

When there are no more elements of order p^i to test, move on to the next prime. When there are no more primes to test, you have a complete list of generators. The structure of the group is now easily determined.

Also, prime ideals I of norm

$$N(I) \leq \frac{1}{2}\left(\frac{4}{\pi}\right)^{r_2}\sqrt{|\Delta|},$$

where $r_2 = 0$ if $\Delta > 0$ and $r_2 = 1$ if $\Delta < 0$, generate the strict class group [8, Thm. 13.20, p. 129] [42, Cor. 10.3, p. 175]. An example of this approach is given in

http://www.imsc.ernet.in/~kapil/crypto/notes/node36.html.

Note that not every ideal class contains one of these prime ideals (for example, for $d = 1009$, $f = 1$), but every ideal class contains a product of these ideals.

More sophisticated methods for computing the class group when $h_+(\Delta)$ is large are given in [6, pp. 235-264] and in [4].

33 Computing the class number

The *strict class number* $h_+(\Delta)$ is the number of strict ideal equivalence classes, which is the same as the number of proper equivalence classes of forms. The *class number* $h(\Delta)$ is the number of weak equivalence classes of ideals.

Improper equivalence classes of forms result in a different class number. For example, $h_+(-303) = h(-303) = 10$, but there are only 6 equivalence classes (allowing proper and improper equivalence) of forms [37, pp. 124-126].

As an aside, we note that there is an alternate definition of equivalence of forms, whereby we define two forms g_1, g_2 to be equivalent if there is a transform T with $\det(T) = \pm 1$ so that $g_2 = \det(T)Tg_1$. Under this alternate equivalence, the equivalence classes of ideals and the equivalence classes of forms are in one-to-one correspondence under the maps ϕ_{FI} and ϕ_{IF}. For example, under this alternate definition of equivalence, the form (a, b, c) is equivalent to the form $(-a, b, -c)$ under the transform

$$T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The proper classes of the forms (a, b, c) and $(-a, b, -c)$ correspond to weakly equivalent ideal classes under the maps ϕ_{FI} and ϕ_{IF}. Note that in the case of negative discriminants, in this paragraph we have relaxed our normal restriction to consideration of only positive definite forms. We can put this alternative equivalence another way: in terms of forms, $h(\Delta)$ is the number of equivalence classes of forms where the proper class of (a, b, c) is considered equivalent to the proper class of $(-a, b, -c)$.

The simplest method (and often most efficient when $f = 1$) to obtain $h_+(\Delta)$ is to count reduced forms for $\Delta < 0$, and periods of reduced forms for $\Delta > 0$. Methods for counting reduced forms are given above in the section, "All reduced forms of discriminant Δ." If $\Delta < 0$, then $h(\Delta) = h_+(\Delta)$. For $\Delta > 0$, if any (a, b, c) is properly equivalent to $(-a, b, -c)$, then $h(\Delta) = h_+(\Delta)$. Otherwise, no (a, b, c) is properly equivalent to $(-a, b, -c)$, and $h(\Delta) = h_+(\Delta)/2$. Alternatively, if $N(\epsilon) = -1$, where ϵ is a fundamental unit in $\mathcal{O}_f$, then $h(\Delta) = h_+(\Delta)$, otherwise $h(\Delta) = h_+(\Delta)/2$.

When $f > 1$, the following formula (and related explanation taken almost verbatim) from Cohn [8, pp. 181-182] may shorten the computation. As usual, let Δ be the discriminant associated with d, f, and let Δ_0 be the fundamental discriminant associated with d. Then

$$(19) \qquad \left\{ \begin{array}{c} h_+(\Delta)/h_+(\Delta_0) \\ \text{or} \\ h(\Delta)/h(\Delta_0) \end{array} \right\} = f \prod_{p|f} \left(1 - \frac{(\Delta_0/p)}{p} \right) \cdot \left\{ \begin{array}{c} 1/E_+ \\ \text{or} \\ 1/E \end{array} \right\}$$

where, for $f \geq 2$, if $\Delta < 0$ then $h_+ = h$ and $E_+ = E = w/2$ where $w = 6$ when $\Delta_0 = -3$, $w = 4$ when $\Delta_0 = -4$, and $w = 2$ otherwise. For $f \geq 2$ and $\Delta > 0$, E is the earliest power of ϵ_1 (the fundamental unit in $\mathcal{O}$) for which $\epsilon_1^E \in \mathcal{O}_f$, and E_+ is the earliest power of ϵ_+ (the generator of totally positive units in $\mathcal{O}$, so $N(\epsilon_+) = +1$) for which $\epsilon_+^{E_+} \in \mathcal{O}_f$. Note that (Δ_0/p) is the Kronecker symbol.

When $\Delta > 0$:

Either $E_+ = E$ or $E_+ = E/2$.

Podsypanin [33] proves that $E \leq 2f$ (the erroneous parts of Podsypanin's proof are easily done using classical results due to Siebeck, summarized in [11, p. 394]. See also [22, p. 94] or [19, Theorem 11]).

If $E_+ = E$, then either $E = 3f/2$ or $E \leq f$.

For a negative fundamental discriminant, $\Delta_0 < 0$, an alternative formula is

$$h_+(\Delta_0) = h(\Delta_0) = \frac{w}{2\Delta_0} \sum_{k=1}^{-\Delta_0 - 1} k \left(\frac{\Delta_0}{k} \right)$$

where w is as in the paragraph containing (19) and $\left(\frac{\Delta_0}{k}\right)$ is the Kronecker symbol.

For a positive fundamental discriminant, $\Delta_0 > 0$, an alternative formula is

$$(20) \qquad h(\Delta_0) = \frac{-1}{2\ln(\epsilon_1)} \sum_{k=1}^{\Delta_0 - 1} \ln(\sin(k\pi/\Delta_0)) \left(\frac{\Delta_0}{k} \right)$$

where ϵ_1 is the fundamental unit in $\mathcal{O}$ and $\left(\frac{\Delta_0}{k}\right)$ is the Kronecker symbol. If $N(\epsilon_1) = -1$ then $h_+(\Delta_0) = h(\Delta_0)$, while if $N(\epsilon_1) = +1$ then $h_+(\Delta_0) =$

$2h(\Delta_0)$. Alternatively, the right hand side of (20) gives $h_+(\Delta_0)$ if $2\ln(\epsilon_1)$ is replaced by $\ln(\epsilon_+)$ where $\epsilon_+ = \epsilon_1$ when $N(\epsilon_1) = +1$, and $\epsilon_+ = \epsilon_1^2$ when $N(\epsilon_1) = -1$. Similar formulas are given in [8, p. 135].

Here's a method to compute $\ln(\epsilon_1)$ when ϵ_1 is too large to compute conveniently. This will use the PQa algorithm [39], to which refer for notation used here. If $\Delta_0 \equiv 1 \pmod 4$ then take $P_0 = 1$, $Q_0 = 2$ and $D = \Delta_0$. If $\Delta_0 \equiv 0 \pmod 4$ then take $P_0 = 0$, $Q_0 = 1$ and $D = \Delta_0/4$. It is not necessary to compute A_i, B_i, or G_i. Then

$$\ln(\epsilon_1) = \ln(\sqrt{D} + Q_0 a_0 - P_0) - P_0 \ln(2) + \sum_{i=1}^{\ell-1} \ln\left(\frac{P_i + \sqrt{D}}{Q_i}\right)$$

where ℓ is the length of the period of the continued fraction expansion of $\frac{P_0+\sqrt{D}}{Q_0}$.

The following are given in [8, p. 135] for Δ_0 a fundamental discriminant:

$$\epsilon_1^h = \frac{\prod_u \sin \pi u/\Delta_0}{\prod_t \sin \pi t/\Delta_0}$$

for $\Delta_0 > 0$, and

$$h = \frac{w}{2|\Delta_0|}\left(\sum u - \sum t\right)$$

for $\Delta_0 < 0$, where t (and u) are the positive residues (and non-residues) less than Δ_0, so $(\Delta_0/t) = +1$, while $(\Delta_0/u) = -1$ [and $\ln(\epsilon_1) > 0$].

To find E and E_+, it suffices to do much of the work modulo f. First apply the PQa algorithm to solve $X^2 - \Delta Y^2 = \pm 4$ or $x^2 - (\Delta/4)y^2 = \pm 1$, as appropriate, but for B_i and G_i, only compute their least positive residues modulo f (and there is no need to compute A_i at all). Then apply the appropriate recursion $y_n = Ky_{n-1} \pm y_{n-2}$, again, only keeping the values modulo f. Here K is X_1, $2X_1$, x_1, or $2x_1$ as appropriate. See ([39]) for the recursions that go with each case.

As an example, we will compute E and E_+ for $d = 229$, $f = 6$ ($\Delta = 8244$). We start by finding the fundamental unit in $\mathcal{O}_1$, modulo 6. In this case we solve $X^2 - 229Y^2 = \pm 4$, and retain the minimal positive solution modulo 6. Applying the PQa algorithm (see ([39])), with $P_0 = 1$, $Q_0 = 2$, and $D = 229$, and retaining only the least positive residues of B_i and G_i modulo 6, we find that the minimal positive solution is a solution to the -4 equation, and that $X_1 \equiv 3 \pmod 6$ and $Y_1 \equiv 1 \pmod 6$. The recursion, for this case, taking $y_0 = 0$, is then $y_n = 3y_{n-1} + y_{n-2}$, giving $y_0, y_1, y_2, \ldots, y_6$, modulo 6, as 0, 1, 3, 4, 3, 1, 0. So, the minimal n for which $y_n \equiv 0 \pmod 6$ is $n = 6$. It then follows that $E = 6$ and $E_+ = 3$. This last is due to the fact that the minimal positive solution to $X^2 - 229Y^2 = +4$ (the minimal totally positive solution) is X_2, Y_2 and $6/2 = 3$. (Not needed, but included for reference: the minimal positive solution to $X^2 - 229Y^2 = \pm 4$ is $X = 15$, $Y = 1$, and this is a solution to the -4 equation; the fundamental unit in $\mathcal{O}_1$ is $7 + 1 \cdot \omega$; for $\tau = 15 + 1 \cdot \sqrt{229}$, $\tau^2/2 = 227 + 15\sqrt{229}$; $\tau^6/2^5 = 11646902 + 128820 \cdot 6 \cdot \sqrt{229}$.)

34 Genus classes for ideals

Two ideals, I, J are said to be in the same genus if there is an ideal K so that I is strictly equivalent to JK^2.

Under the maps ϕ_{FI} and ϕ_{IF}, above, two ideals are in the same genus class of ideals if and only if the corresponding forms are in the same genus class of forms. More precisely, two strict classes of ideals are in the same genus class if and only if the corresponding proper classes of forms are in the same genus class.

Note that for forms, two classes that are improperly equivalent are always in the same genus class (they represent the same numbers), while for ideals, two ideals that are weakly equivalent might be in different genus classes. For example, for $d = 2233$, $f = 1$, the ideals (1) and $(-24 + \omega)$ are weakly equivalent (both are principal), but they lie in different genus classes. Here $\phi_{FI}(36, 35, -7) = (1)$, and $\phi_{FI}(-36, 35, 7) = (-24 + \omega)$.

To test whether two ideals are in the same genus class, apply ϕ_{IF} to both, and apply the character-based test given in the section "Genus equivalence of forms".

From a slightly more advanced standpoint: If H_+ is the strict class group, then the genus group G is $H_+/2H_+$, which is the same as H_+ tensored with $\mathbf{Z}/2\mathbf{Z}$. A tensor map F from $H_+ \otimes \mathbf{Z}/2\mathbf{Z}$ to $H_+/2H_+$ is $F(h, 0) = 0$, and $F(h, 1) = h + 2H_+$ [34].

35 Overall structure

It can be useful to make an exhibit displaying some elements of the overall structure of the strict class group, the weak class group, the genus group, and relations among them. Here's an outline of a way to do that for a given discriminant Δ and its corresponding d and f (d is squarefree, $d \neq 1$, $\Delta = f^2 d$ or $\Delta = 4f^2 d$ depending on whether $d \equiv 1 \pmod 4$ or not).

Start by listing all primitive reduced forms. If $\Delta > 0$, separate them into periods, and select one member of each period. In any given period, if there is a form with x^2 coefficient $+1$, you might select that form; otherwise if there is a form with x^2 coefficient -1, you might select that form; otherwise if there is an ambiguous form (in which case there are exactly two), you might select one of the ambiguous forms. This gives one element of each proper class of forms.

These proper classes of forms are in one-to-one correspondence with strict classes of invertible ideals. For each proper/strict class, find an ideal of minimal norm that corresponds to the form for the class under the maps ϕ_{IF} and ϕ_{FI}. To do this, find all ideals of norm i for $i = 1,\ 2,\ 3, \cdots$, map the ideal to a form using ϕ_{IF}, and see which of the forms in the list of reduced forms this is equivalent to. For each class, keep the first ideal that maps to the class. Continue until you have an ideal for each class.

If $\Delta < 0$, then strict and weak classes of invertible ideals are identical. If $\Delta > 0$, to find which strict classes are weakly equivalent, take the form

(a, b, c) representing the strict class, see which class has the form $(-a, b, -c)$ as a member. If some strict ideal class is weakly equivalent to itself, all strict ideal classes are weakly equivalent to themselves.

Again, if $\Delta < 0$, then proper and improper classes of primitive forms are identical. If $\Delta > 0$, to find which proper classes are improperly equivalent, take the form (a, b, c) representing the proper class, see which class has the form $(a, -b, c)$ as a member. There will always be at least one proper class that is improperly equivalent to itself, namely the principal class (1). Generally, some proper classes are improperly equivalent to themselves, while others are not. Classes that are improperly equivalent to themselves are ambiguous classes and have order 1 or 2 in the strict class group. For $\Delta > 0$, each ambiguous class has exactly two reduced ambiguous forms and exactly two primitive ambiguous ideals. An ambiguous class might have no reduced ambiguous ideals. As an example, for $d = 34$, $f = 1$, the strict class of the ideal $[9, 4 + \sqrt{34}]$ has no reduced ambiguous ideals, although this class includes the ambiguous ideals $[17, \sqrt{34}]$ and $[34, \sqrt{34}]$.

Note that there can be noninvertible ambiguous ideals in the order. They are not elements of classes in the class group, because the class group is by definition the equivalence classes of invertible ideals. Noninvertible ideals map to imprimitive forms, and are not relatively prime to f. An invertible ideal might or might not be relatively prime to f, and every invertible ideal is strictly equivalent to an ideal that is relatively prime to f.

When $\Delta < 0$, $h_+(\Delta) = h(\Delta)$ and $H_+(\Delta) = H(\Delta)$.

When $\Delta > 0$, if any strict ideal class is weakly equivalent to itself, then every strict ideal class is weakly equivalent to itself, $N(\epsilon) = -1$ where ϵ is a fundamental unit for the order, and $h_+(\Delta) = h(\Delta)$ and $H_+(\Delta) = H(\Delta)$. Otherwise, no strict ideal is weakly equivalent to itself, $N(\epsilon) = +1$, and $h_+(\Delta) = 2h(\Delta)$. In this case, $H_+(\Delta)$ might or might not be $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. See the section "Norm equations and units" for discussion of how to find a fundamental unit for an order. See the section "The relation between $H_+(\Delta)$ and $H(\Delta)$" (and the section preceding that) for further discussion of

the possible relations between the class group and the strict class group.

To separate the strict classes into genus classes, first determine the characters for the discriminant (see section above, "Genus equivalence of forms"). For each form representing a strict class, find an integer m represented by the form and relatively prime to 2Δ. Compute each character at this m. Two strict classes are in the same genus if the values of all the characters are the same.

The principal genus consists of the squares of the elements of the strict class group.

Classes that are properly equivalent will be in the same genus class. Classes that are weakly equivalent might or might not be in the same genus class. But, either each strict class and its weak equivalent are all in the same genus class, or each such pair is in two distinct genus classes.

The order of the genus group is equal to the number of ambiguous classes in the strict class group. This is discussed in more detail below.

Each order has two singular ideals, which are ambiguous principal ideals generated by elements of positive norm. One of these is (1). If $h_+(\Delta) = h(\Delta)$, then the other is $(f\sqrt{d})$. When $h_+(\Delta) \neq h(\Delta)$, the process of computing the other singular ideal is a bit more complicated, and is given below in the section "Singular ideals."

To find the order of an element of the strict class group, take the representative ideal I, successively raise it to the powers 1, 2, 3, ... Stop when you find a j so that $I^j \approx (1)$. The order of the element is then j. This can be done by successively setting $J = I \cdot J$, and then substituting for J the ideal of smallest norm for its strict class group. To get the group structure, see the section, "Computing the class group".

If $h_+(\Delta) = h(\Delta)$ then there is exactly one principal class. If $h_+(\Delta) = 2h(\Delta)$, there are exactly two principal classes, one generated by elements of positive norm, and one generated by elements of negative norm.

Using the overall structure of the strict class group $H_+(\Delta)$ and the facts that

Elements of the principal genus are the squares of strict classes, and

Ambiguous classes are those of order 1 or 2 in the strict class group.

it is easy to see that

The order of the genus group $|G|$ is equal to the number of ambiguous classes, and

There is one ambiguous class in each coset of the principal genus class if and only if $h_+(\Delta)/|G|$ is odd,

as we will show.

For r a prime power, let $S(r)$ denote the cyclic group of order r, so (as an additive group) $S(r) = \mathbf{Z}/r\mathbf{Z}$. Write

$$H_+(\Delta) = S(2) \times \cdots \times S(2) \times S(2^{k_1}) \times \cdots \times S(2^{k_m}) \times S(q_1) \times \cdots \times S(q_n)$$

where there are ℓ factors $S(2)$, $k_i > 1$, q_i is an odd prime power, and any or all of ℓ, m, n might be zero.

Given the above, it is easy to see that

$$h_+(\Delta) = 2^{\ell+k_1+\cdots+k_m} q_1 \dots q_n.$$

Elements of the principal genus class are those with even "coefficient" in the groups $S(2^i)$ (above, for $i \geq 1$) and any coefficient in the groups $S(q_i)$, so $|G| = 2^{\ell+m}$. A class is ambiguous if its coefficient in $S(2^i)$ (for $i \geq 1$) is 0 or 2^{i-1}, and is 0 in the groups $S(q_i)$, so the number of ambiguous classes is also $2^{\ell+m}$. This shows that the number of ambiguous classes is equal to the order of the genus group.

In fact, each of 2^ℓ genus classes contain 2^m ambiguous ideals. In particular, $h_+/|G|$ is odd if and only if each genus class contains exactly one ambiguous class.

For fundamental discriminants of Type IV (see "The relation between $H_+(\Delta)$ and $H(\Delta)$," below), $h_+(\Delta)$ is even ($= 2h(\Delta)$), but $H_+(\Delta)$ is not $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$, so $H_+(\Delta)$ must have a cyclic subgroup $\mathbf{Z}/2^k\mathbf{Z}$ where $k > 1$. It follows that for Type IV, $h_+(\Delta)/|G|$ is always even. For Types II and III, this ratio can be either even or odd. Also, for Type IV, $h_+(\Delta) \equiv 0 \pmod 4$.

36 A little group theoretic result

We will use this in the following section.

Let G_1, G_2 be groups, with the group operation written additively. Let $=$ denote isomorphism.

Lemma 1 *Let $a \in G_1$, $a \neq 0$, $2a = 0$, $G_1/\{0, a\} = G_2$, and $G_1 = G_2 \oplus \mathbf{Z}/2\mathbf{Z}$. Then there is no $b \in G_1$ so that $2b = a$.*

Write $[r, s]$ to represent an element of $G_2 \oplus \mathbf{Z}/2\mathbf{Z}$, with $r \in G_2$ and $s \in \mathbf{Z}/2\mathbf{Z}$. We can write the isomorphism $G_1 = G_2 \oplus \mathbf{Z}/2\mathbf{Z}$ so that a corresponds to $[0, 1]$. If $b \in G_1$ corresponds to $[r, s]$, then $2b$ corresponds to $[2r, 2s] = [2r, 0] \neq [0, 1]$.

37 The relation between $H_+(\Delta)$ and $H(\Delta)$

If $\Delta < 0$ then $H_+(\Delta) = H(\Delta)$, and so $h_+(\Delta) = h(\Delta)$. End of story for $\Delta < 0$.

We will discuss each of the following points below.

> If $\Delta > 0$ then we can have $h_+(\Delta) = h(\Delta)$ or $h_+(\Delta) = 2h(\Delta)$. When $h_+(\Delta) = h(\Delta)$, we always have $H_+(\Delta) = H(\Delta)$. When $h_+(\Delta) = 2h(\Delta)$, in some cases $H_+(\Delta) = H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$, and in some cases this does not hold.
>
> Take ϵ to be the fundamental unit for the order of discriminant Δ. If $N(\epsilon) = -1$ then $h_+(\Delta) = h(\Delta)$, while if$N(\epsilon) = +1$, then $h_+(\Delta) = 2h(\Delta)$.
>
> When $N(\epsilon) = +1$, the ideals (1) and $(f\omega)$ are not strictly equivalent. But they might or might not be in the same genus class. If they are in the same genus class, then $H_+(\Delta) \neq H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$, while if they are not in the same genus class then $H_+(\Delta) = H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$.

The relation between $H_+(\Delta)$ and $H(\Delta)$ follows from the relation between the ideals (1) and $(f\omega)$. These two ideals are always weakly equivalent, $(1)(f\omega) = (f\omega)(1)$ and $N(f\omega/1) < 0$.

If $(1) \approx (f\omega)$ then every ideal is strictly equivalent to any ideal that it is weakly equivalent to, so strict classes and weak classes are identical (if $(1)(\alpha) = (f\omega)(\beta)$ with $N(\alpha/\beta) > 0$ and $I(\gamma) = J(\delta)$ with $N(\gamma/\delta) < 0$ then $I(\gamma\alpha) = J(\delta\beta f\omega)$ and $N((\gamma\alpha)/(\delta\beta f\omega)) > 0$). So $H_+(\Delta) = H(\Delta)$ and $h_+(\Delta) = h(\Delta)$.

We will show that $(1) \approx (f\omega)$ if and only if $N(\epsilon) = -1$. This is most easily seen using the corresponding forms.

Consider first $d \equiv 1 \pmod 4$. Here the strict class of the ideal (1) corresponds to the proper class of the form $g_1 = (1, -f, f^2(1-d)/4)$ and the strict class of the ideal $(f\omega)$ corresponds to the proper class of the form $g_2 = (-1, -f, -f^2(1-d)/4)$. The ideals (1) and $(f\omega)$ are strictly equivalent if and only if the forms g_1 and g_2 are properly equivalent. If $g_1 \approx g_2$, then there is a unimodular transform

$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

so that $g_1 T = g_2$. In particular $N(\alpha + \gamma f\omega) = g_1(\alpha, \gamma) = -1$. This means $\alpha + \gamma f\omega$ is a unit with norm -1, so for the fundamental unit, $N(\epsilon) = -1$.

Conversely, assume $N(\epsilon) = -1$ and set $\epsilon = r + sf\omega$. Then the transform

$$T = \begin{pmatrix} r & sf^2\left(\dfrac{1-d}{4}\right) \\ s & -r - fs \end{pmatrix}$$

is unimodular and takes g_1 to g_2, so these forms are properly equivalent.

When $d \not\equiv 1 \pmod 4$, similar arguments apply using the forms $(1, 0, -f^2 d)$ and $(-1, 0, f^2 d)$, and the transform

$$T = \begin{pmatrix} r & -sf^2 d \\ s & -r \end{pmatrix}.$$

Now assume that $(1) \not\approx (f\omega)$, so $H_+(\Delta) \neq H(\Delta)$. From the above, this happens if and only if $N(\epsilon) = +1$. Note that $H(\Delta) = H_+(\Delta)/\{(1), (f\omega)\}$.

In this case, either (1) and $(f\omega)$ are in the same genus class, or they are not. This is easily tested using characters for forms. The ideal (1) corresponds to the form, $g = (1, -f, f^2(1-d)/4)$ or $(1, 0, -f^2 d)$ depending on whether $d \equiv 1 \pmod 4$ or not. The form g represents $+1$, and for every character $\chi_i(1) = 1$. The ideal $(f\omega)$ corresponds to negative of the form g, $(-1, -f, -f^2(1-d)/4)$ or $(-1, 0, f^2 d)$ depending on whether $d \equiv 1 \pmod 4$ or not, and this form represents -1. It therefore suffices to compute $\chi_i(-1)$ for each character. If these values are all $+1$, then (1) and $(f\omega)$ are in the same genus class, while if any $\chi_i(-1)$ is -1, then the two ideals are in different genus classes. For there to be a character with $\chi_i(-1) = -1$, either there is an odd prime p dividing fd so that $p \equiv 3 \pmod 4$ or Δ is even and $\Delta/4 \equiv 4 \pmod 8$.

Suppose now that there is a character so $\chi_i(-1) = -1$. When this happens, $H_+(\Delta) = H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. In fact, we can easily pick ideal classes to represent the elements of $H(\Delta)$. We can pick ideals $I_2, I_3, \ldots, I_{h(\Delta)}$ so that $(1), (f\omega), I_2, I_2(f\omega), I_3, I_3(f\omega), \ldots, I_{h(\Delta)}, I_{h(\Delta)}(f\omega)$ represent the classes of $H_+(\Delta)$. For each pair $\{I_j, I_j(f\omega)\}$, exactly one of $\chi_i(I_j)$, $\chi_i(I_j(f\omega))$ is $+1$ and the other is -1. Pick the ideal with $\chi_i(\ ) = +1$ as the representative in H. There could be more than one character with $\chi_i(-1) = -1$, and different such characters result in different representatives for H. An example using $\Delta = 2233 = 7 \cdot 11 \cdot 29$ is given below.

If $\chi_i(-1) = +1$ for every i, then $(f\omega)$ is in the principal genus, so there is an ideal I so that $I^2 = (f\omega)$ [8, Cor. 14.44a, p. 151]. By the lemma of the previous section, $H_+(\Delta) \neq H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. The above method of selecting a representative for H will not work because $I^2 \approx (I(f\omega))^2 \approx (f\omega) \not\approx (1)$.

Note that $f^2 d$ or $f^2 d/4$ is a sum of relatively prime squares if and only if for all characters $\chi_i(-1) = +1$.

For fundamental discriminants Cohn [8, p. 142] summarizes these relationships in the table below. Singular ideals are discussed in a later section.

Type	Fundamental Discriminant	$N(\epsilon_1)$	$h_+(\Delta_0)$	$H_+(\Delta_0)$	$\chi_i(-1)$	Singular Ideals
I	$0 > \Delta_0$	$\cdots$	$h(\Delta_0)$	$= H(\Delta_0)$	$\cdots$	(1), $(\sqrt{d})$
II	$0 < \Delta_0 = A^2 + B^2$	-1	$h(\Delta_0)$	$= H(\Delta_0)$	all $+1$	(1), $(\sqrt{d})$
III	$0 < \Delta_0 \neq A^2 + B^2$	$+1$	$2h(\Delta_0)$	$= H(\Delta_0) \times \mathbf{Z}/2\mathbf{Z}$	some -1	(1), (λ)
IV	$0 < \Delta_0 = A^2 + B^2$	$+1$	$2h(\Delta_0)$	$\neq H(\Delta_0) \times \mathbf{Z}/2\mathbf{Z}$	all $+1$	(1), (λ)

Table 3: Relations for fundamental discriminants

Here are some examples illustrating possible relations between $H_+(\Delta)$ and $H(\Delta)$ for different orders in given quadratic fields.

Consider $d > 0$ so that for $f = 1$, Δ_0 is Type II, i.e., $N(\epsilon_1) = -1$ (for example, $d = 2$, $\Delta_0 = 8$, $\epsilon_1 = 1 + \sqrt{2}$). Then if f is such that for Δ, $N(\epsilon_1) = -1$, then $h_+(\Delta) = h(\Delta)$ and $H_+(\Delta) = H(\Delta)$. But if f is such that for Δ, $N(\epsilon_1) = +1$, then $h_+(\Delta) = 2h(\Delta)$. In this case, $H_+(\Delta)$ might or might not be $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. For example, for $d = 2$, $f = 17$, $\epsilon_1 = 577 + 408\sqrt{2}$, $H(\Delta) = \mathbf{Z}/2\mathbf{Z}$, and $H_+(\Delta) = \mathbf{Z}/4\mathbf{Z}$ is not $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. But for $d = 2$, $f = 6$, $\epsilon_1 = 17 + 12\sqrt{2}$, $H(\Delta) = \mathbf{Z}/2\mathbf{Z}$, and $H_+(\Delta) = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ is $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$.

If $d > 0$ is such that for $f = 1$, Δ_0 is Type III, i.e., $N(\epsilon_1) = +1$ and d is not a sum of two squares (for example, $d = 3$, $\Delta_0 = 12$, $\epsilon_1 = 2+\sqrt{3}$), then for any f and Δ, we always have $h_+(\Delta) = 2h(\Delta)$ and $H_+(\Delta) = H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$.

If $d > 0$ is such that for $f = 1$, Δ_0 is Type IV, i.e., $N(\epsilon_1) = +1$ and d is a sum of two squares (for example, $d = 34$, $\Delta_0 = 136$, $\epsilon_1 = 35 + 6\sqrt{34}$), then for any f and Δ, we always have $h_+(\Delta) = 2h(\Delta)$, but we might or might not have $H_+(\Delta) = H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. For example, for $d = 34$, $f = 5$, $\epsilon_1 = 2449 + 420\sqrt{34}$, $H(\Delta) = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, and so $H_+(\Delta) = \mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ is not $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$. But for $d = 34$, $f = 2$, $\epsilon_1 = 35 + 6\sqrt{34}$, $H(\Delta) = \mathbf{Z}/4\mathbf{Z}$,

and so $H_+(\Delta) = \mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ is $H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$.

The general rule is that if $f^2 d = a^2 + b^2$ for some a, b with $\gcd(a, b) = 1$ or 2, $8 \nmid f^2 d$, then $H_+(\Delta) \neq H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$; otherwise $H_+(\Delta) = H(\Delta) \times \mathbf{Z}/2\mathbf{Z}$.

(Cannot allow $d = f = 10$, with $f^2 d = 1000$, as $1000 = 18^2 + 26^2$, but this does not work. If d is even, then f has to be odd, while if d is odd, f can have up to one factor of 2. Quadratic residues of 16 are 0, 1, 4, 9, so if $4|f$, then $4|a$ and $4|b$.)

Further to some of the arguments above, it is easy to see that if d is a sum of squares, then there is a principal ideal in $\mathcal{O}_1$ that is generated by an element of negative norm and is the square of another ideal. For $t^2 + v^2 = d$ with v odd, consider the ideals $J = [v, t + \sqrt{d}]$, $K = J^2$, and the element $\alpha = t + \sqrt{d}$, with norm $N(\alpha) = t^2 - d = -v^2$. Then

$$J^2 = [v, t+\sqrt{d}]^2 = [v^2, v(t+\sqrt{d}), t^2+2t\sqrt{d}+d] = [v^2, v(t+\sqrt{d}), t^2+2t\sqrt{d}+v^2+t^2] =$$

$$[v^2, v(t + \sqrt{d}), 2t(t + \sqrt{d}) + v^2] = [v^2, \gcd(v, 2t)(t + \sqrt{d})].$$

Since $\gcd(v, 2t) = 1$, then $J^2 = [v^2, t + \sqrt{d}]$. Also,

$$(\alpha) = (t + \sqrt{d}) = [t + \sqrt{d}, t\sqrt{d} + d] =$$

$$[t + \sqrt{d}, t\sqrt{d} + t^2 + v^2] = [t + \sqrt{d}, t(t + \sqrt{d}) + v^2] = [t + \sqrt{d}, v^2] = J^2.$$

So, $(\alpha) = J^2$. Also $(\alpha) \sim (\sqrt{d})$ because both are principal ideals generated by elements of negative norm ($N(\sqrt{d}) = -d$). So the class of $(\sqrt{d})$ is the square of the class of J.

If $t^2 + v^2 = f^2 d$, the same sort of argument carries through in $\mathcal{O}_f$.

More generally, for $f \geq 1$ we can define Types I, II, III, IV as follows:

Type I—$d < 0$

Type II—$d > 0$ and $N(\epsilon_1) = -1$

Type III—$d > 0$, $N(\epsilon_1) = +1$, and for some i, $\chi_i(-1) = -1$

Type IV—$d > 0$, $N(\epsilon_1) = +1$, and for all i, $\chi_i(-1) = +1$

Then we have the following as theorems:

For Type I or II, $H_+ = H$ (so $h_+ = h$).

For Type III, $H_+ = H \times \mathbf{Z}/2\mathbf{Z}$ (so $h_+ = 2h$).

For Type IV, $h_+ = 2h$ and $H_+ \neq H \times \mathbf{Z}/2\mathbf{Z}$.

Δ is Type III if and only if $16|\Delta$ or there is a prime $p \equiv 3 \pmod 4$ so that $p|\Delta$

Δ is Type III if and only if neither Δ nor $\Delta/4$ is as sum of two relatively prime squares.

An alternative statement of some of the above for $\Delta > 0$ is that $H_+ = H \times \mathbf{Z}/2\mathbf{Z}$ if and only if $16|\Delta$ or there is a prime $p \equiv 3 \pmod 4$ so that $p|\Delta$. Otherwise, i.e., if $\Delta = 2^\alpha \prod p_i^{\alpha_i}$ with $\alpha \in \{0, 2, 3\}$ and $p_i \equiv 1 \pmod 4$ for all i, then

if $N(\epsilon_1) = -1$ then $H_+ = H$ (so $h_+ = h$), and

if $N(\epsilon_1) = +1$ then $h_+ = 2h$ and $H_+ \neq H \times \mathbf{Z}/2\mathbf{Z}$.

38 Example of picking representatives for H

As an example, here are two ways to pick representative ideals for H for $\Delta = d = 2233$, $f = 1$ (Type III). The characters for $\Delta = 2233 = 7 \cdot 11 \cdot 29$ are $\chi_1(m) = \left(\frac{m}{7}\right)$, $\chi_2(m) = \left(\frac{m}{11}\right)$, and $\chi_3(m) = \left(\frac{m}{29}\right)$. As $\chi_1(-1) = (-1/7) = -1$, $\chi_2(-1) = (-1/11) = -1$, and $\chi_3(-1) = (-1/29) = +1$, either of the first two can be used to select elements of H.

The table, "Characters of $H_+(2233)$" has a row for each class, and gives a representative form (a, b, c), a value m represented by the form so that $\gcd(2\Delta, m) = 1$, and the values of the three characters for the form. (The numbering of the classes, "Index," is arbitrary.) The classes selected using

Characters in $H_+(2233)$

Index	Form	m	$\left(\frac{m}{7}\right)$	$\left(\frac{m}{11}\right)$	$\left(\frac{m}{29}\right)$
1	$(-1, 47, 6)$	-1	-1	-1	1
2	$(1, 47, -6)$	1	1	1	1
3	$(-32, 21, 14)$	3	-1	1	-1
4	$(32, 21, -14)$	-3	1	-1	-1
5	$(-32, 43, 3)$	3	-1	1	-1
6	$(32, 43, -3)$	-3	1	-1	-1
7	$(-11, 33, 26)$	27	-1	1	-1
8	$(29, 29, -12)$	-27	1	-1	-1
9	$(-28, 21, 16)$	9	1	1	1
10	$(28, 21, -16)$	-9	-1	-1	1
11	$(-28, 35, 9)$	9	1	1	1
12	$(28, 35, -9)$	-9	-1	-1	1

Table 4: Characters in $H_+(2233)$

Structure of $H_+(2233)$

Index	Ideal	Form	Weakly Equiv to	Improperly Equiv to	Genus Class	Group Element
1	$[6, \omega]$	$(-1, 47, 6)$	2	1	A	100
2	$[1, \omega]$	$(1, 47, -6)$	1	2	B	000
3	$[3, 2+\omega]$	$(-32, 21, 14)$	4	5	C	011
4	$[2, \omega]$	$(32, 21, -14)$	3	6	D	111
5	$[3, \omega]$	$(-32, 43, 3)$	6	3	C	012
6	$[2, 1+\omega]$	$(32, 43, -3)$	5	4	D	112
7	$[12, 2+\omega]$	$(-11, 33, 26)$	8	7	C	010
8	$[8, 2+\omega]$	$(29, 29, -12)$	7	8	D	110
9	$[4, 1+\omega]$	$(-28, 21, 16)$	10	11	B	001
10	$[6, 3+\omega]$	$(28, 21, -16)$	9	12	A	101
11	$[4, 2+\omega]$	$(-28, 35, 9)$	12	9	B	002
12	$[6, 2+\omega]$	$(28, 35, -9)$	11	10	A	102

Table 5: Various relations in $H_+(2233)$

χ_1 are 2, 4, 6, 8, 9, 11, and the classes selected using χ_2 are 2, 3, 5, 7, 9, 11. Using the table, "Structure of $H_+(2233)$," you can verify that either of these choices works. Each row in this table represents a class and gives a representative ideal, the same representative form as above, the index of the weakly equivalent class, the index of the properly equivalent class, the genus class, and the group element. The group $H_+(2233)$ is $\mathbf{Z}/2\mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z} \oplus \mathbf{Z}/3\mathbf{Z}$, and in the table we write the element (i_1, i_2, i_3) as "i_1 i_2 i_3." For example, the identity element is written "0 0 0." The group operation is additive.

This method cannot work for $\Delta = d = 1345$, $f = 1$ (Type IV) because the class of $I = (\sqrt{1345})$, which is the class of principal ideals generated by elements of negative norm, is the square of an ideal class, containing, say, J. Then $J^2 \approx (JI)^2 \approx I$, so neither the class of J nor the class of JI can be in H, as H is closed under multiplication, and the class of I is not in H. Here's an explicit example:

$$(35{+}2\omega) = \begin{pmatrix} 7 & 0 \\ 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 49 & 42 \\ 0 & 1 \end{pmatrix} \approx \begin{pmatrix} 336 & 0 \\ 0 & 1 \end{pmatrix} = (-1{+}2\omega) = (\sqrt{1345}).$$

For more on the material above, see chapter 14 of Cohn [8], especially pages 142 and 143; 150 to 153, including remark 14.47; and exercise 14.15 on page 161.

39 Singular ideals

An ideal in the maximal order is *singular* if it is strictly principal and ambiguous. It is strictly principal if is strictly equivalent to (1), or, equivalently, if it is generated by an element of positive norm. It is ambiguous if it is equal to its conjugate, i.e., $[t + u\omega, v + w\omega]$ is equal to $[t + u\overline{\omega}, v + w\overline{\omega}]$. An ideal class is called ambiguous if it is improperly equivalent to itself (an ambiguous ideal class contains two ambiguous ideals).

Note that (1) is always a singular ideal. Also, $(\sqrt{d})$ is always equal to its conjugate $(-\sqrt{d})$. So if $(\sqrt{d}) \approx (1)$, as for Type I and Type II, $(\sqrt{d})$ is the other singular ideal.

If $(\sqrt{d}) \not\approx (1)$, then some other ideal has to be the other singular ideal. Let $t + u\omega$ be a fundamental unit in $\mathcal{O}$. Then $N(t + u\omega) = +1$, as otherwise we'd be in Type II. Set $g = \gcd(1 + t, u), v = (1 + t)/g, w = u/g$. Then $(v + w\omega)$ is a singular ideal. This ideal might or might not be reduced.

40 Prime and irreducible elements of an order

The factorization of elements in an order is driven by the fact that the maximal order ($f = 1$) is always a Dedekind domain, while the smaller orders ($f > 1$) are Noetherian, but not Dedekind domains. Because orders with $f > 1$ are not Dedekind domains, they can never be UFDs.

A ring is *Noetherian* if it satisfies the following three equivalent conditions:

> Every ideal is finitely generated.
>
> The ascending chain condition (for any infinite sequence of ideals $I_1 \subset I_2 \subset \cdots \subset I_k \subset \cdots$, there is an i so that if $j > i$ then $I_j = I_i$).
>
> Maximal ideal condition (for any set of ideals, there is at least one not contained in any of the others).

A ring is a *Dedekind domain* if it is Noetherian, it is integrally closed, and every prime ideal is maximal. A ring R is *integrally closed* if whenever a/b with $a \in R$, $b \in R - \{0\}$ satisfies a monic polynomial with coefficients in R, then $a/b \in R$ [8, p. 9] [40, pp. 27-30]. Alternatively, a ring is a Dedekind domain if every ideal is invertible.

Every order $\mathcal{O}_f$ with $f \geq 1$ is Noetherian, and in every order every prime ideal is maximal. So what distinguishes an order $\mathcal{O}_f$ with $f > 1$ from $\mathcal{O}_1$ is that $\mathcal{O}_1$ is integrally closed while $\mathcal{O}_f$ with $f > 1$ is never integrally closed, as we now show.

Every element $a + b\omega$ of the maximal order is in the integral closure of $\mathcal{O}_f$ for any $f \geq 1$. To see this, note that $\mathbf{Z} \subset \mathcal{O}_f$, and every element of the

maximal order is an algebraic integer and so is the root of a monic polynomial with coefficients in $\mathbf{Z}$ hence in $\mathcal{O}_f$. We can take $r = af + bf\omega$, $s = f$ and we have $r \in \mathcal{O}_f$ and $s \in \mathcal{O}_f - \{0\}$. So $r/s = a + b\omega$ is in the integral closure of $\mathcal{O}_f$. It might be simpler to just note that ω is in the integral closure of any order (any f) because $\omega = \frac{f\omega}{f}$ is a root of either $x^2 + x + (1-d)/4$ or $x^2 - d$ depending on whether $d \equiv 1 \pmod 4$ or not, but ω is an element of the order only if the order is the maximal order.

A Dedekind domain is a UFD if and only if its class number is 1. The only UFDs $\mathbf{Q}(\sqrt{d})$ with negative discriminant are those for $d = -1,\ -2,\ -3,\ -7,\ -11,\ -19,\ -43,\ -67,\ -163$ [41, 1, 15]. It is not known whether there are arbitrarily large positive d so that $\mathbf{Q}(\sqrt{d})$ has class number 1. Those d up to 500 so that $\mathbf{Q}(\sqrt{d})$ has class number 1 are

$d =$ 2, 3, 5, 6, 7, 11, 13, 14, 17, 19, 21, 22, 23, 29, 31, 33, 37, 38, 41, 43, 46, 47, 53, 57, 59, 61, 62, 67, 69, 71, 73, 77, 83, 86, 89, 93, 94, 97, 101, 103, 107, 109, 113, 118, 127, 129, 131, 133, 134, 137, 139, 141, 149, 151, 157, 158, 161, 163, 166, 167, 173, 177, 179, 181, 191, 193, 197, 199, 201, 206, 209, 211, 213, 214, 217, 227, 233, 237, 239, 241, 249, 251, 253, 262, 263, 269, 271, 277, 278, 281, 283, 293, 301, 302, 307, 309, 311, 313, 317, 329, 331, 334, 337, 341, 347, 349, 353, 358, 367, 373, 379, 381, 382, 383, 389, 393, 397, 398, 409, 413, 417, 419, 421, 422, 431, 433, 437, 446, 449, 453, 454, 457, 461, 463, 467, 478, 479, 487, 489, 491, 497.

When $f > 1$ an order might have class number 1, but, as noted above, the order will not be a UFD. For example, for $d = 5$, $f = 2$, $\Delta = 20$, we have that $h_+(20) = h(20) = 1$, and $4 = 2 \cdot 2 = (\sqrt{5} - 1)(\sqrt{5} + 1)$ gives two factorizations of 4 into irreducibles in $\mathcal{O}_2$.

For $d = 13$, $f = 7$, $\Delta = 637$, we have that $h_+(637) = 2$, $h(637) = 1$, and $49 = 7^2 = (91 + 10\omega)(\text{-}161 + 10\omega) = (\text{-}14 + 1\omega)(7 + 1\omega) = (\text{-}532 + 33\omega)(\text{-}532 + 33\omega) = (28 + 3\omega)(\text{-}49 + 3\omega)$, where each term shown is irreducible.

For $d = -3$, $f = 3$, $\Delta = -27$, we have that $h_+(-27) = h(-27) = 1$, and $27 = 3^3 = (-\omega)^3 = (3 - 2\omega)^2 = (6 - \omega)(3 + \omega) = (-3 + \omega)^3 = (3)(\omega)(3 - \omega)$.

Note that some of these factorizations involve 2 irreducibles, and some involve 3.

For $d = 79$, $f = 1$, $\Delta = 316$, we have that $h_+(316) = 6$, $h(316) = 3$, and $27 = 3^3 = (-17 + 2\omega)(17 + 2\omega)$.

To see whether an element α with $N(\alpha) = n$ is irreducible, for every $1 < m \leq \sqrt{n}$, $m|n$, test divide by one element β from each class of elements with $N(\beta) = m$.

If $\alpha \in \mathcal{O}_f$ is prime, then $\alpha|p$ for some rational prime p. In particular, $N(\alpha)$ is p or p^2. If $N(\alpha) = p$ for a rational prime p then α is prime. If $N(\alpha) = p^2$, and α is irreducible, then α is prime if and only if there are no solutions to

$$x^2 \equiv f^2 d \pmod 8$$

if $d \equiv 1 \pmod 4$ and $p = 2$, or to

$$x^2 \equiv f^2 d \pmod p$$

otherwise.

In particular, if $p|f$ then p is irreducible but not prime (when $p|f$ there are no solutions to $N(x) = p$, but there are always solutions to the appropriate equation just above). This makes it easy to generate examples of non-unique factorization in orders with $f > 1$. For instance, for $d \not\equiv 1 \pmod 4$ and p odd, $(f\sqrt{d})^2 = (p)^2(f/p)^2(d)$. The factor $f\sqrt{d}$ is irreducible in $\mathcal{O}_f$. While (f/p) and d might factor further into rational primes, none of those primes will divide $f\sqrt{d}$ in $\mathcal{O}_f$. When $d \equiv 1 \pmod 4$, examples look more complicated but follow from the same considerations.

Note that the norm of an irreducible element need not be prime. But if

$$N(\alpha) = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k} q_1^{b_1} q_2^{b_2} \cdots q_\ell^{b_\ell}$$

where α is irreducible, $\gcd(p_i, f) = 1$, and $q_i|f$, then $a_1 + a_2 + \cdots + a_k \leq h(\Delta)$.

41 Prime ideals

An ideal $I \in \mathcal{O}_f$ is said to *divide* an ideal $J \in \mathcal{O}_f$ (written $I|J$) if $I \supset J$. To see whether I divides J it suffices to determine whether each element of a basis for J is in I. Alternatively $I|J \iff I + J = I$. Along these lines, note that $IJ \subset I \cap J \subset I \subset I + J$ and similarly for J [6, p. 180]. The notion of divisibility of ideals is a little different from that for elements. If the element a divides the element b, then there is an element c so that $ac = b$. If an ideal I divides an ideal J, there is not necessarily an ideal K so that $IK = J$, although in many cases there is such a K. For instance, in $\mathcal{O}_2$ of $\mathbf{Q}(\sqrt{5})$, for $I = [2, \omega]$ and $J = [4, \omega]$, $I|J$, but there is no K so that $IK = J$. Note that here, $I^2 = [4, 2\omega]$. (But be aware that authors differ on the definition of divisibility of ideals. For instance, Narkiewicz defines divisibility as $I|J$ if there is an ideal K so that $IK = J$ [31, p. 8].)

The ideal I of $\mathcal{O}_f$ is a *prime ideal* if $I \neq \mathcal{O}_f$ and if the quotient ring $\mathcal{O}_f/I$ is an integral domain, i.e., $xy \in I$ implies $x \in I$ or $y \in I$ [8, p. 43] [6, pp. 182-183]. Equivalently, an ideal I is prime if whenever I divides $I_1 I_2$, then I divides at least one of I_1, I_2 [8, p. 43]. Additional equivalent definitions are given in [8, pp. 42-43].

In general, a maximal ideal is always a prime ideal (in any ring), but a prime ideal is not always maximal. But, in an order of a quadratic number field, a prime ideal (other than $\{0\}$) is always maximal.

In general, in orders, we do not have factorization of ideals into primes. But every ideal contains a product of prime ideals. In fact, for every ideal I there are prime ideals P_1, P_2, ..., P_k so that

$$P_1 P_2 \cdots P_k \subset I \subset P_1 \cap P_2 \cap \cdots \cap P_k.$$

For example, for $d = -3$, $f = 2$, $I = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$ and $P = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, we have that $P^2 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$, and

$$P^2 \subset I \subset P = P \cap P.$$

Here, P is prime, while the other ideals discussed are not.

In general, $N(I)N(J) = N(IJ)$ if at least one of I, J is proper, but this might not hold if neither I nor J is proper. In the example just given, $N(P) = 2$, while $N(P^2) = 8$. In any event, $N(I)N(J) \leq N(IJ)$.

In the maximal order ($f = 1$), any ideal can be factored uniquely into a product of prime ideals. In smaller orders ($f > 1$), ideals relatively prime to f can be factored uniquely into a product of prime ideals. Note that an order with $f > 1$ will have ideals that are invertible but not relatively prime to f. These ideals will not be a product of prime ideals. An order with $f > 1$ will also have ideals that are not invertible (which will not be relatively prime to f). These ideals might or might not be a product of prime ideals.

Every prime ideal divides (p) for some rational prime $p > 0$. So, to find prime ideals, it suffices to consider divisors of these (p). This means that prime ideals are either $(p) = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$ or $\begin{pmatrix} p & v \\ 0 & 1 \end{pmatrix}$ for some v. If $d \equiv 1 \pmod 4$ and there is a solution v to

$$f^2(d-1)/4 - v(v+f) \equiv 0 \pmod p \tag{21}$$

then there is a prime ideal $\begin{pmatrix} p & v \\ 0 & 1 \end{pmatrix}$ that divides (p). If there is no such v, then (p) is prime. Another way to write (21) that can be useful is

$$f^2 d - (2v+f)^2 \equiv 0 \pmod{4p}. \tag{22}$$

If $d \equiv 2$ or $3 \pmod 4$ then the equivalent of equation (21) is

$$f^2 d - v^2 \equiv 0 \pmod p. \tag{23}$$

Some specific cases are as follows (see also [8, p. 90]). If $p|f$ (including $p = 2$) then $I = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$ is the unique irreducible ideal that divides (p). This ideal is not invertible. Here, (p) divides I^2 but is not equal to $I^2 = \begin{pmatrix} p^2 & 0 \\ 0 & p \end{pmatrix}$ (the norm of I^2 is p^3, while the norm of (p) is p^2.)

If $p \nmid f$, p is odd, and $p|d$ then p ramifies, i.e., there is a unique v so that for $I = \begin{pmatrix} p & v \\ 0 & 1 \end{pmatrix}$, I divides (p). In fact, $I^2 = (p)$. If $d \equiv 1 \pmod 4$, then v can be taken to be either $-f/2$ of $(p-f)/2$, depending on the parity of f. If $d \not\equiv 1 \pmod 4$ then v can be taken to be 0.

If $p \nmid f$, p is odd, $p \nmid d$, and $x^2 \equiv d \pmod p$ has a solution then p splits, i.e. there are $v_1 \neq v_2$ so that for $I_i = \begin{pmatrix} p & v_i \\ 0 & 1 \end{pmatrix}$, with $i = 1, 2$, $I_1 I_2 = (p)$.

If $p \nmid f$, p is odd, $p \nmid d$, and $x^2 \equiv d \pmod p$ has no solutions then p remains inert, i.e., (p) is prime.

If $2 \nmid f$, $d \equiv 2$ or $3 \pmod 4$ (so 2 divides the discriminant $4f^2 d$), then 2 ramifies, i.e., there is a v so that for $I = \begin{pmatrix} 2 & v \\ 0 & 1 \end{pmatrix}$, $I^2 = (2)$. If $d \equiv 2 \pmod 4$ then $v = 0$, and if $d \equiv 3 \pmod 4$ then $v = 1$.

If $f^2 d \equiv 1 \pmod 8$ then 2 splits, i.e.,

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

If $f^2 d \equiv 5 \pmod 8$ then 2 remains inert, i.e., (2) is prime.

Alternatively, we can test whether a general ideal $I = [t, v + wf\omega]$ is prime relatively easily. We just generate $(tw)^2$ pairs of elements from the cosets of I in $\mathcal{O}_f$ and see whether some product of two nonzero elements is in I. Representatives of the cosets are given by $i + jf\omega$ for $0 \leq i \leq t-1$, $0 \leq j \leq w-1$. If any $i_1 + j_1 f\omega \neq 0$ times $i_2 + j_2 f\omega \neq 0$ is in I, then I is not prime. If no such product is in I then I is prime.

42 Ambiguous classes

Gauss discovered that proper classes of forms have a natural group structure, induced by the operation of composition of forms. Proper classes of forms correspond to strict classes of ideals, and the structure Gauss discovered

applies to either set of classes. For strict classes of ideals, this structure is induced by multiplication of ideals.

Forms and ideals have broader definitions of equivalence than proper and strict equivalence. For forms, there is improper equivalence and for ideals there is weak equivalence. These two broader definitions of equivalence do not correspond in any natural way. Weak equivalence of ideals results in either exactly the same classes as strict equivalence, or exactly half as many classes. That is, under weak equivalence, either every class is weakly equivalent to itself and no other classes, or every class is weakly equivalent to some one other class and not to itself (you cannot have some classes weakly equivalent to themselves and others weakly equivalent to classes other than themselves). Classes under weak equivalence also have a natural group structure, again induced by multiplication of ideals.

For forms under improper equivalence, generally some proper classes will be improperly equivalent to themselves, while other classes will be improperly equivalent to proper classes other than themselves. This wreaks complete havoc with the group structure on proper classes, and it is (generally) impossible to put a natural group structure on proper-and-improper classes of forms. This is part of what makes ambiguous classes so interesting.

A form (a, b, c) is *ambiguous* if $a|b$. An ideal I is *ambiguous* if $I = I'$ where I' is the conjugate of I.

A class of forms in $H_+(\Delta)$ is *ambiguous* if it contains an ambiguous form. A class of ideals in $H_+(\Delta)$ is *ambiguous* if it contains an ideal I for which $I \approx I'$.

Under any of the usual isomorphisms between proper classes of forms and strict classes of ideals, ambiguous classes of forms and ambiguous classes of ideals correspond. That is, the class of forms is ambiguous if and only if the corresponding class of ideals is ambiguous.

For the remainder of this section, "class" will mean proper class of forms or strict class of ideals, i.e., a class of $H_+(\Delta)$.

Here are some properties of ambiguous classes of forms. A class of forms

is ambiguous if and only if

1. Some form in the class is improperly equivalent to itself, in which case every form in the class is improperly equivalent to itself.

2. The class is improperly equivalent to itself.

3. The class is its own inverse.

4. The class has order 1 or 2 in $H_+(\Delta)$.

5. If $d > 0$, the class has exactly two reduced ambiguous forms. If $d < 0$, the class has two ambiguous forms of the form $(a, 0, c)$ or (a, a, c) (at most one of which is reduced).

A form (a, b, c) is ambiguous if it is properly equivalent to its inverse $(a, -b, c)$, or, equivalently, to (c, b, a). Note that (a, b, c) is always improperly equivalent to both of these last two. When $d > 0$, all ambiguous forms are properly equivalent to a reduced ambiguous form.

A class of ideals is ambiguous if and only if

1. Some ideal in the class is strictly equivalent to its conjugate, in which case every ideal in the class is strictly equivalent to its own conjugate.

2. The class is its own conjugate.

3. The class is its own inverse.

4. The class has order 1 or 2 in $H_+(\Delta)$.

5. The class has exactly two primitive ambiguous ideals.

When $d > 0$ the number of ambiguous ideals in an order is equal to the number of reduced ambiguous forms. Each of these is twice the number of genera. The number of reduced ambiguous ideals in an order is equal to the number of genera.

Classes for $\mathbf{Q}(34)$

Class	Ideal	Order
A	$[1, \sqrt{34}]$	1
B	$[9, 4 + \sqrt{34}]$	2
C	$[3, 1 + \sqrt{34}]$	4
D	$[3, 2 + \sqrt{34}]$	4

Table 6: Classes, representative ideals, and orders for maximal order of $\mathbf{Q}(34)$

When $d < 0$ the number of ambiguous ideals in an order is equal to the number of ambiguous forms of the form $(a, 0, c)$ or (a, a, c). Each of these is twice the number of genera. The number of reduced ambiguous ideals in an order is less than or equal to twice the number of genera.

If $d < 0$, an ambiguous class of ideals can have exactly 0 or 1 reduced ambiguous ideals as members, while if $d > 0$ an ambiguous class of ideals can have exactly 0, 1, or 2 reduced ambiguous ideals as members. If $d > 0$ and an ambiguous class does not have a reduced ambiguous ideal, or has two reduced ambiguous ideals, then the order has no unit of norm -1. The converse is not true.

Here are two examples. We consider the maximal orders of $\mathbf{Q}(34)$ and $\mathbf{Q}(146)$. Both have strict ideal class groups that are cyclic of order 4. For $\mathbf{Q}(34)$, we can label the classes A, B, C, D, with representative ideals and orders in $H_+(\Delta)$ as in the table, "Classes for $\mathbf{Q}(34)$."

Class A is the principal class generated by elements of positive norm and so has order 1 in $H_+(4 \cdot 34)$. Class B is the principal class generated by elements of negative norm and so has order 2 in $H_+(4 \cdot 34)$.

The table "Ambiguous Ideals in $\mathbf{Q}(34)$" lists the ambiguous ideals in $\mathbf{Q}(34)$, tells what class each is in, and whether the ideal is reduced. The maximal order of $\mathbf{Q}(34)$ has four ambiguous ideals, two of which are reduced, and both reduced ambiguous ideals are in class A. The ideal $[2, \sqrt{34}]$ is the

Ambiguous Ideals in $\mathbf{Q}(34)$

Ideal	Class	Reduced?
$[1, \sqrt{34}]$	A	Yes
$[2, \sqrt{34}]$	A	Yes
$[17, \sqrt{34}]$	B	No
$[34, \sqrt{34}]$	B	No

Table 7: Ambiguous ideals, classes, and whether reduced for maximal order of $\mathbf{Q}(34)$

Classes for $\mathbf{Q}(146)$

Class	Ideal	Order
A	$[1, \sqrt{146}]$	1
B	$[2, \sqrt{146}]$	2
C	$[5, 1 + \sqrt{146}]$	4
D	$[5, 4 + \sqrt{146}]$	4

Table 8: Classes, representative ideals, and orders for maximal order of $\mathbf{Q}(146)$

non-trivial singular ideal, generated by $6 + \sqrt{34}$.

Similarly, for $\mathbf{Q}(146)$, we can label the classes A, B, C, D, with representative ideals and orders in $H_+(\Delta)$ as in the table, "Classes for $\mathbf{Q}(146)$." Again, the ambiguous classes are A and B. The table "Ambiguous Ideals in $\mathbf{Q}(146)$" lists the ambiguous ideals in $\mathbf{Q}(146)$, tells what class each is in, and whether the ideal is reduced. The maximal order of $\mathbf{Q}(146)$ has four ambiguous ideals, two of which are reduced, one reduced ambiguous ideal is in class A, and one reduced ambiguous ideal is in class B. The non-trivial singular ideal is $[73, \sqrt{146}]$, which is generated by $73 + 6\sqrt{146}$.

Ambiguous Ideals in $\mathbf{Q}(146)$

Ideal	Class	Reduced?
$[1, \sqrt{146}]$	A	Yes
$[2, \sqrt{146}]$	B	Yes
$[73, \sqrt{146}]$	A	No
$[146, \sqrt{146}]$	B	No

Table 9: Ambiguous ideals, classes, and whether reduced for maximal order of $\mathbf{Q}(146)$

43 Numeri idonei

Euler used his *numeri idonei* in 1778 to show that

$$18518809 = 197^2 + 1848 \cdot 100^2$$

is prime, an impressive feat at the time (here 1848 is a numerus idoneus). As there are better methods for testing primality now, this method is not currently in vogue, but numeri idonei (aka *idoneal numbers, suitable numbers, convenient numbers*) are of certain inherent interest. The main fact about numeri idonei is that if n is a numerus idoneus, $n = ab$, and $m > 1$ is an odd integer for which there is exactly one representation $m = ax^2 + by^2$ with $x, y \geq 0$, and if for this representation $\gcd(ax, by) = 1$, then m is prime. Conversely, numeri idonei are the only numbers with this property. Euler had a method for determining whether an n was a numerus idoneus, which apparently was not quite correct. He did however identify the 65 numeri idonei known today. Ribenboim [37, p. 357] gives the following criterion due to Grube for determining whether a number is idoneal.

> Thus, n is a convenient number if and only if for every $x \geq 1$ such that $q = n + x^2 \leq \frac{4n}{3}$, if $q = rs$ and $2x \leq r \leq s$, then $r = s$ or $r = 2x$.

But there is more than the main fact, above, that can be said about numeri idonei and the form of numbers uniquely, or nearly uniquely, represented by idoneal forms ($ax^2 + by^2$, where $ab = n$ is a numerus idoneus). First, a criterion for n in the context of the structure of the quadratic orders is that n is one of Euler's numeri idonei if $n > 0$ and the order with discriminant $\Delta = -4n$ has exactly one class (proper class of forms or strict class of ideals) in each genus class. For n a numerus idoneus, and $ab = n$, consider the form $g(x, y) = ax^2 + by^2$. With two minor exceptions noted below, we have:

> If $m > 1$ is represented by g in a unique way (see comment below) by $x, y > 0$ with $\gcd(ax, by) = 1$, and **is not** represented by any other $x, y \geq 0$, then m is a prime or 2 times an odd prime (*Type 1* representation).
>
> If $m > 1$ is represented by g in a unique way by $x, y > 0$ with $\gcd(ax, by) = 1$, and **is** represented by some other $x, y \geq 0$, then m is a non-trivial prime power or 2 times a non-trivial odd prime power (*Type 2* representation).

We consider a representation of m as $g(x_1, y_1)$ with $x_1, y_1 \geq 0$, $\gcd(ax_1, by_1) = 1$ to be unique if for any other x_2, y_2 so that $x_2, y_2 \geq 0$, $\gcd(ax_2, by_2) = 1$, and $m = g(x_2, y_2)$ either $x_1 = x_2$ and $y_1 = y_2$, or $n = a = b = 1$ and $x_1 = y_2$ and $y_1 = x_2$.

The two small exceptions are that 8 is represented uniquely by nonnegative x, y by the forms $x^2 + 7y^2$ and $3x^2 + 5y^2$.

There are 65 known numeri idonei, and it is known that there is at most one more square-free numerus idoneus. If there is such a 66th numerus idoneus, it is larger than $10^{60}/4$ [37, pp. 142, 161, 357]. The 65 known numeri idonei are:

> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 18, 21, 22, 24, 25, 28, 30, 33, 37, 40, 42, 45, 48, 57, 58, 60, 70, 72, 78, 85, 88, 93, 102, 105, 112, 120, 130, 133, 165, 168, 177, 190, 210, 232, 240, 253,

273, 280, 312, 330, 345, 357, 385, 408, 462, 520, 760, 840, 1320, 1365, 1848.

Here's an example of a non-idoneal number. Consider that the form x^2+11y^2 represents 15 uniquely with x, y positive, $15 = 2^2+11\cdot 1^2$, but 15 is composite, so 11 is not idoneal.

The known numeri idonei produce 204 forms $ax^2 + by^2$ with $0 < a \leq b$, $\gcd(a, b) = 1$. Table 1 gives a little more information about the properties of numbers represented by these idoneal forms. In this table, there are five forms that represent slightly different classes of numbers than the remaining 199 forms, and we refer to these five as "special." "Type 1" and "Type 2" are defined above, while p represents an odd prime. Note that for "Type 2" representations, the powers of p are all at least 2.

For more extensive discussion, see [45, pp. 188, 219-226], [10, pp. 59-63], [14, §303, pp. 361-363, §334 pp. 404-406], [37, pp. 356-358, 142, 161], [12, p. 89], [5, p. 193], [3], [11, p. 361]. Brown [3] gives further sources, and some cautions about statements in the literature.

Types of Representation

Form	Type 1	Type 2
Special Forms		
x^2+y^2	2, p, $2p$	p^{k+1}, $2p^{k+1}$
x^2+3y^2	p	2^{2k}, p^{k+1}
x^2+7y^2	8, p	2^{k+3}, p^{k+1}
x^2+15y^2	p	2^{2k}, p^{k+1}
$3x^2+5y^2$	8, p	2^{2k+3}, p^{2k+1}
Other than Special Forms		
$a=1$, b even	p	p^{k+1}
$a=1$, b odd	p, $2p$	p^{k+1}, $2p^{2k+1}$
$1<a<b$, ab odd	p, $2p$	p^{2k+1}, $2p^{2k+1}$
$1<a<b$, ab even	p	p^{2k+1}

Table 10: Forms of numbers represented by idoneal forms ($k \geq 1$)

Here are a few examples of representations of various m by forms ax^2+by^2 where ab is idoneal. In each case, we give all representations of m by the form with $x, y \geq 0$.

$191 = 3 \cdot 7^2 + 11 \cdot 2^2$

$2 \cdot 79 = 3 \cdot 7^2 + 11 \cdot 1^2$

$19^3 = 82^2 + 15 \cdot 3^2 = 38^2 + 15 \cdot 19^2$

$3^4 = 1^2 + 5 \cdot 4^2 = 6^2 + 5 \cdot 3^2 = 9^2 + 5 \cdot 0^2$

$2^7 = 3 \cdot 1^2 + 5 \cdot 5^2 = 3 \cdot 6^2 + 5 \cdot 2^2 = 3 \cdot 4^2 + 5 \cdot 4^2$

$2^6 = 7^2 + 15 \cdot 1^2 = 2^2 + 15 \cdot 2^2 = 8^2 + 15 \cdot 0^2$

$2 \cdot 3^5 = 19^2 + 5 \cdot 5^2 = 21^2 + 5 \cdot 3^2 = 9^2 + 5 \cdot 9^2$

$2 \cdot 17^2 = 23^2 + 7^2 = 17^2 + 17^2$

44 So how are idoneal numbers a special case of something?

You might ask are there other forms with analogous properties, why was it mentioned that $n > 0$ is an idoneal number if the discriminant $-4n$ has one class per genus class, and what about forms $ax^2 + bxy + cy^2$ with $b \neq 0$ and discriminant Δ with one class per genus class? And how about forms with positive discriminant? Well, forms of discriminants with one class per genus have properties that are just the same as the idoneal forms above.

First, let's give some background on how many ways a prime can be represented by forms of a given discriminant. The short answer will turn out to be that any prime p that is represented by a form is represented by that form in an "essentially unique" way, and that at most two proper classes of forms with the same discriminant can represent p. We begin by elaborating on this.

Consider first the case where, given a discriminant Δ, we let p be an odd prime so that $\gcd(\Delta, p) = 1$. Recall from the section "Solving $ax^2 + bxy + cy^2 = m$," that any form g that represents p is properly equivalent to a form $g' = px^2 + bxy + cy^2$, with $0 \leq b < 2p$, and the representation of p by g corresponds to the representation of p by $(1, 0)$ in g'. So, how many such forms are there? Well, let's count solutions to

$$b^2 \equiv \Delta \pmod{4p}. \tag{24}$$

This has no solutions unless

$$x^2 \equiv \Delta \pmod{p} \tag{25}$$

has solutions. When this last equation has solutions, which we assume henceforth, it has exactly 2 with $0 < x < p$. And $x^2 \equiv \Delta \pmod 4$ always has exactly 2 solutions with $0 \leq x < 4$. So, by the Chinese Remainder Theorem, (24) has exactly 4 solutions with $0 \leq b < 4p$. Clearly if b is a solution, then so are $4p - b$, $2p + b$, and $2p - b$. So, exactly 2 solutions satisfy $0 \leq b < 2p$

and we get two forms, $g_1 = px^2 + b_1xy + c_1y^2$, and $g_2 = px^2 + b_2xy + c_2y^2$ with $0 \leq b_i < 2p$. How are these forms related, and what does that tell us about representations of p by forms of discriminant Δ? Well, g_1 and g_2 are in the same genus class because they both represent p and p is relatively prime to Δ. In fact, g_1 and g_2 are improperly equivalent under the transform

$$T = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

If g_1 and g_2 are properly equivalent, then they are in an ambiguous class and each is improperly equivalent to itself. If g_1 and g_2 are not properly equivalent, then neither is in an ambiguous class and neither is improperly equivalent to itself.

As an example, consider representations of 241 by the form $g = -x^2 + 75xy + 25y^2$, a form of discriminant $5725 = 5^2 \cdot 229$. The two solutions to $x^2 \equiv 5725 \pmod{4 \cdot 241}$ with $0 < x < 2 \cdot 241$ are $x = 69,\ 413$. So our forms are $g_1 = 241x^2 + 69xy - y^2$ and $g_2 = 241x^2 + 413xy + 171y^2$. These are properly equivalent, as shown by the transform

$$T = \begin{pmatrix} 1 & 1 \\ 69 & 70 \end{pmatrix},$$

for which $g_1T = g_2$. From the representation $241 = g_1(1, 0)$ and the transform

$$g \begin{pmatrix} 72 & -1 \\ 1 & 0 \end{pmatrix} = g_1$$

we have that g_1 and g are properly equivalent, and we get the representation $241 = g(72, 1)$. Similarly from the representation $241 = g_2(1, 0)$ and the transform

$$g \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} = g_2$$

we get the representation $241 = g(3, 1)$. From these two representations we get all representations of 241 by g from the automorphs of g. But since either

of these two representations can be derived from the other by a transformation of determinant -1, (for example,

$$g \begin{pmatrix} -1 & 75 \\ 0 & 1 \end{pmatrix} = g$$

carries the representation by $(72, 1)$ to the representation by $(3, 1)$). So any representation of 241 by g can be derived from any other such representation by a transform of determinant ± 1 that takes g to itself. In this sense, the representation of 241 by g is "essentially unique".

Continuing with odd primes p, if $p^2|\Delta$, then no primitive form of discriminant Δ can represent p (for the form $px^2 + bxy + cy^2$, $p^2|b^2 - 4pc$, so $p|b$ and $p|c$). If $p|\Delta$, $p^2 \nmid \Delta$, then p can be represented by primitive forms of discriminant Δ. Here (25) has just one solution with $0 \leq x < p$, namely $x = 0$. The result is that if such a p is represented by a form g, that representation is essentially unique, in the sense above.

There are several cases that need to be considered when $p = 2$, but, as with odd primes, if 2 is represented by a form g, that representation is essentially unique.

Now, for comparison, let's consider representations of the composite number pq by forms of discriminant Δ, where p and q are distinct odd primes, and $\gcd(pq, \Delta) = 1$. As above, to find all such representations, we start by looking for forms $pqx^2 + bxy + cy^2$ with,

$$b^2 \equiv \Delta \pmod{4pq}, \tag{26}$$

and

$$0 \leq b < 2pq. \tag{27}$$

How many solutions will there be? If the equation (26) has solutions, there will be exactly four solutions that satisfy (27). To see this, observe that each of the equations $b^2 \equiv \Delta \pmod 4$, $b^2 \equiv \Delta \pmod p$, and $b^2 \equiv \Delta \pmod q$ must have solutions, and each has exactly 2 non-negative solutions less than

the modulus (4, p, or q). So we get 8 solutions b satisfying $0 \le b < 4pq$, and of these 4 satisfy $0 \le b < 2pq$ (because if b is a solution, so is $4pq - b$). This gives us 4 forms that represent pq. There might or might not be pairs of these forms that are properly equivalent. Each of these 4 forms will be improperly equivalent to one of the 4, possibly itself (if it is in an ambiguous class).

Let's look at some examples. Consider $\Delta = 401$, $p = 5$, and $q = 7$. The four forms from the above procedure are $(35, 11, -2)$, $(35, 59, 22)$, $(35, 31, 4)$, and $(35, 39, 8)$. No two of these are properly equivalent. The first and second are improperly equivalent, as are the third and fourth. The representations of 35 by any of these four forms, or any form properly equivalent to any of these four, are essentially unique. Note that $h_+(401) = h(401) = 5$.

Next, let's look at $\Delta = 29$, again with $p = 5$, and $q = 7$. Note that $h_+(29) = h(29) = 1$. The four forms from the above procedure are $(35, 13, 1)$, $(35, 27, 5)$, $(35, 43, 13)$, and $(35, 57, 23)$. Any pair of these are properly and improperly equivalent. For any of these forms, there are two essentially different representations of 35. For example, for the form $g = (35, 13, 1)$, there is no transformation of determinant ± 1 that takes either of the solutions $(x, y) = (1, 0)$ and $(2, -5)$ to the other. Note that a transformation of determinant -1 that takes g to itself is

$$\begin{pmatrix} -1 & 0 \\ 13 & 1 \end{pmatrix}.$$

So what's the difference between the cases $\Delta = 401$ and $\Delta = 29$? In the first case, $h_+(401) = 5$, and the four forms corresponding to the solutions to $x^2 \equiv 401 \pmod{140}$ are each in separate proper equivalence classes. So each can represent 35 essentially uniquely. But in the second case, $h_+(29) = 1$, so the four forms corresponding to the solutions to $x^2 \equiv 29 \pmod{140}$ are all in the same proper equivalence class. The transformations of determinant -1 can associate pairs of solutions, but this leaves two pairs of solution that cannot be related by transformations of determinant ± 1.

For the idoneal forms, the m with essentially unique representations are

those m with exactly 1 positive solution.

We hope the above has motivated the more precise description that follows.

Recall the following, which defines the d and f associated with Δ. For a discriminant Δ, define k to be the largest integer so that $k^2|\Delta$ and define $d = \Delta/k^2$. If $d \equiv 1 \pmod 4$ then define $f = k$, otherwise define $f = k/2$. (Note that Δ is then the discriminant of the quadratic order $\mathcal{O}_f$ in the quadratic number field $\mathbf{Q}(\sqrt{d})$.)

Call a representation $g(x, y) = ax^2 + bxy + cy^2 = m$ of m by the form g *primitively essentially unique* if every primitive representation of m by g can be derived from any other primitive representation of m by g by a transformation of g to itself of determinant ± 1.

We say a representation of m by g is *Type 1* if the representation is primitively essentially unique, and there are **no** other representations of m by g. For example $5 = 1^2 + 2^2$ is a Type 1 representation of 5 by $g = x^2 + y^2$.

We say a representation of m by g is *Type 2* if the representation is primitively essentially unique, and there **are** other representations of m by g. For example $25 = 3^2 + 4^2 = 0^2 + 5^2$ is a Type 2 representation of 25 by $g = x^2 + y^2$.

We say a form g is *good* if the only m with $\gcd(m, fd) = 1$ and Type 1 representations by g are those m that are prime or 2 times an odd prime, or, equivalently, the only m with $\gcd(m, fd) = 1$ and Type 2 representations are those m that are a prime power, with exponent at least 2, or 2 times an odd prime power, again with exponent at least 2. That the condition $\gcd(m, fd) = 1$ is needed is shown by the primitive essentially unique representation of the composite 15 by the form $6x^2 + 6xy - y^2$ of discriminant 60 (and more specifically that there are many composites primitively essentially uniquely represented by g, but all such composites have $\gcd(m, 15) > 1$).

As usual, let $h(\Delta)$ be the class number of the discriminant Δ, $h_+(\Delta)$ be the strict (or proper) class number of Δ, and $|G|$ the number of genus classes for Δ.

Then

Lemma 2 *A form g of discriminant Δ is good if and only if one of the following is true:*

$h_+(\Delta) = |G|$, *or*

$h_+(\Delta) = 2|G|$ *and g is not in an ambiguous class.*

For forms $g = ax^2 + bxy + cy^2$ with $a, c > 0$, $b = 0$ (so the form is $g = ax^2 + cy^2$), the above conditions have an equivalent, but easier statement. Namely, if the discriminant $-4ac$ has one class per genus then if there is a unique $x > 0$ and $y > 0$ so that $g(x, y) = m$, and if for this x, y, $\gcd(ax, cy) = 1$, then m is a prime or 2 times an odd prime (with two minor exceptions for $x^2 + 7y^2 = 8$ and $3x^2 + 5y^2 = 8$). For these forms, $n = ac$ is called an *idoneal number*. Note that the only transformations that take any of these forms of these forms to itself are $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$, and, if $a = c = 1$, $\begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}$. The $\pm$ signs can be taken independently.

To recap the main points. Consider a composite pq, with p, q distinct odd primes. If there are solutions to $x^2 \equiv \Delta \pmod{4pq}$ with $0 \le x < 2pq$ then there will be exactly 4 solutions. These give 4 inequivalent solutions for some forms of discriminant Δ. No two of these 4 can be equivalent. That means that if two (or more) are equiv to the same form, there is no transformation of determinant $+1$ that takes that form to itself, and maps one solution to another. But there might be a transform of determinant -1 that maps the form to itself and takes one of the solutions to another.

So, if $h_+/|G| = 1$, all 4 solutions are in the same class, and even after possible pairing, there are at least one pair for which there is no transform of determinant ± 1 that carries either solution to the other, so no form can primitively essentially uniquely represent a composite pq.

If $h_+/|G| \ge 3$, two classes can each have one solution, and a third can have 2. If this last is ambiguous, all 3 can essentially uniquely represent pq.

If $h_+/|G| = 2$, then the two classes can each have 2 solutions. If these two classes are ambiguous, each can essentially uniquely represent pq. But, if

one of these classes is not ambiguous, that class cannot essentially uniquely represent pq.

So, a form of discriminant Δ is good if either $\{h_+/|G| = 1\}$ or $\{h_+/|G| = 2$ and the form is not in an ambiguous class$\}$.

Note that the above does not apply to $2p$ as there is only one solution to $x^2 \equiv \Delta \pmod 2$, and so only 2 solutions to $x^2 \equiv \Delta \pmod{8p}$ with $0 \leq x < 4p$, and these can be equivalent under a transform of determinant -1.

45 Notation

Refer to the tables "List of Notations" for notations used.

List of Notations

Symbol	Definition
d	A squarefree integer $\neq 0, 1$
ω	With 1, a basis for the ring of integers in $\mathbf{Q}(\sqrt{d})$,
	$= \sqrt{d}$ if $d \not\equiv 1 \pmod 4$,
	$= (1+\sqrt{d})/2$ if $d \equiv 1 \pmod 4$.
f	Conductor of the quadratic order $\mathcal{O}_f = [1, f\omega]$.
Δ	Discriminant of an order or of a quadratic form,
	$= 4f^2 d$ if $d \not\equiv 1 \pmod 4$,
	$= f^2 d$if $d \equiv 1 \pmod 4$,
	$= b^2 - 4ac$ for the form $g(x,y) = ax^2 + bxy + cy^2$.
Δ_0	Fundamental discriminant ($f = 1$).
(a, b, c)	The form $g(x,y) = ax^2 + bxy + cy^2$.
$\overline{\alpha}$	The conjugate of $\alpha \in \mathcal{O}_f$.
$N(\alpha)$	The norm of the element α. $N(\alpha) = \alpha\overline{\alpha}$.
$N(I)$	The norm of the ideal I.
$\left(\frac{r}{s}\right)$	Kronecker symbol.
(α)	the ideal generated by α.
$I \vert J$	The ideal I divides the ideal J.
ϵ	Any unit. $N(\epsilon) = \pm 1$.
ϵ_1	The fundamental unit in $\mathcal{O}$. Every unit in $\mathcal{O}$ is $\pm\epsilon_1^n$, $n \in \mathbf{Z}$.
	If ϵ_1 is written as $(x + y\sqrt{d})/2$ then $x, y > 0$ and
	x, y are the smallest positive solutions to $x^2 - dy^2 = \pm 4$
ϵ_+	The generator of totally positive units in $\mathcal{O}$; $N(\epsilon_+) = 1$.
	$\epsilon_+ = \epsilon_1$ or ϵ_1^2.
$h(\Delta)$	The ideal class number of the discriminant Δ.
$h_+(\Delta)$	The strict ideal class number of the discriminant Δ.
$H(\Delta)$	The ideal class group for the discriminant Δ.
$H_+(\Delta)$	The strict ideal class group for the discriminant Δ.
$\sim$	Equivalent forms (properly or improperly) or ideals (strictly or weakly).
$\approx$	Properly equivalent forms or strictly equivalent ideals.
$\approxeq$	Denotes genus equivalence of forms or ideals.
χ	Character.

Table 11: Definitions of symbols used

List of Notations

Symbol	Definition
E	The smallest power of ϵ_1 that lies in $\mathcal{O}_f$.
E_+	The smallest power of ϵ_+ that lies in $\mathcal{O}_f$.
ϕ_{FI}	A map from proper classes of forms to strict classes of ideals.
ϕ_{IF}	A map from strict classes of ideals to proper classes of forms.
ϕ_{QI}	A map from classes of quadratic irrationals to strict classes of ideals.
ϕ_{IQ}	A map from strict classes of ideals to classes of quadratic irrationals.

Table 12: Definitions of symbols used

References

[1] A. Baker, Linear forms in the logarithms of algebraic numbers I, II, III, *Mathematika* **13** (1966), 204-216; ibid. **14** (1967), 102-107; ibid. **14** 1967 220-228.

[2] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Academic Press, New York, 1966.

[3] Kevin Brown, Numeri Idonei, http://www.mathpages.com/home/kmath058.htm.

[4] Johannes Buchmann and Arthur Schmidt, Computing the Structure of a Finite Abelian Group, *Mathematics of Computation* **74**, No. 252, 2005, pp. 2017-2026.

[5] D. A. Buell, *Binary Quadratic Forms, Classical Theory and Modern Computations*, Springer–Verlag, New York, 1989. As the title says, including solution of binary quadratic form equations. The standard method is summarized in Theorem 4.4 on page 53. The forms listed on the last line of this page do not necessarily give rise to inequivalent solu-

tions. A somewhat different approach is summarized in Theorems 4.26 and 4.27 on page 75.

[6] Henri Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993.

[7] Henri Cohen et al., PARI, freeware available at http://www.parigp-home.de/.

[8] Harvey Cohn, *A Classical Introduction to Algebraic Numbers and Class Fields*, Springer–Verlag, 1978.

[9] Harvey Cohn, *Advanced Number Theory*, Dover, New York, 1980.

[10] David A. Cox, *Primes of the Form* $x^2 + ny^2$, John Wiley & Sons, New York, 1989.

[11] L. E. Dickson, *History of the Theory of Numbers*, Vol. 1, AMS Chelsea Publishing, Providence, Rhode Island, 1999.

[12] L. E. Dickson, *Introduction to the Theory of Numbers*, Dover, New York, 1957. Reprint of 1927 University of Chicago edition.

[13] Steven Finch, *Class Number Theory*, 2005, available at http://pauillac.inria.fr/algo/csolve/clss.pdf.

[14] Carl Friedrich Gauss, *Disquisitiones Arithmeticae*, English edition, translated by Arthur A. Clarke, revised by William C. Waterhouse, Springer-Verlag, New York, 1986.

[15] Dorian Goldfeld, Gauss' class number problems for imaginary quadratic fields, *Bull. Amer. Math. Soc.* **13** (1985), 23–37.

[16] Adolf Hurwitz, *Lectures on Number Theory*, Springer-Verlag, New York, 1986. Chapter 6 is a wonderful exposition of methods to solve binary quadratic form equations, $Ax^2+Bxy+Cy^2 = N$. Much theory of simple continued fractions is also developed.

[17] M. J. Jacobson, Jr. and H. C. Williams, *Solving the Pell Equation*, Springer-Verlag, 2009.

[18] D. H. Lehmer, An Extended Theory of Lucas' Functions, *The Annals of Mathematics*, 2nd Ser., Vol. 32, No. 3. (Jul., 1930), pp. 419-448.

[19] D. H. Lehmer, On the Multiple Solutions of the Pell Equation, *The Annals of Mathematics*, 2nd Ser., Vol. 30, No. 1/4. (1928 - 1929), pp. 66-72.

[20] William Judson Leveque, *Topics in Number Theory*, Volumes I and II, Dover, New York , 2002. Chapter 1 of Volume II gives the standard method of solution of binary quadratic form equations for positive and negative discriminants.

[21] Daniel Marcus, *Number Fields*, Springer–Verlag, New York, 1977.

[22] G. B. Mathews, *Number Theory*, Chelsea, New York, undated (also as *Number Theory, Part I*, G. E. Stechert & Co, New York, 1927). Includes a classic exposition of binary quadratic forms.

[23] Keith Matthews, The diophantine equation $ax^2 + bxy + cy^2 = N$, $D = b^2 - 4ac > 0$, *J. Théor. Nombres Bordeaux*, **14** (2002) 257-270. For additions see http://www.numbertheory.org/papers.html#jntb (which is in "publications" at http://www.numbertheory.org/keith.html).

[24] Keith Matthews, The diophantine equation $x^2 - Dy^2 = N$, $D > 1$, in integers, *Expositiones Mathematicae*, **18** (2000), 323-331. Gives the LMM algorithm for solving $x^2 - Dy^2 = N$ for any nonzero N. Additional material at http://www.numbertheory.org/papers.html.

[25] Howard H. Mitchell, On Classes of Ideals in a Quadratic Field, *The Annals of Mathematics*, 2nd Ser., Vol. 27, No. 4. (Jun., 1926), pp. 297-314.

[26] Richard Mollin, *Algebraic Number Theory*, Chapman & Hall/CRC, Boca Raton, 1999.

[27] Richard Mollin, *Fundamental Number Theory with Applications*, CRC Press, Boca Raton, 1998.

[28] Richard Mollin, *Quadratics*, CRC Press, Boca Raton, 1996.

[29] Richard Mollin, Simple Continued Fraction Solutions for Diophantine Equations, *Expositiones Mathematicae*, **19** (2001), pp. 55–73. Gives the LMM algorithm for solving $x^2 - Dy^2 = N$ for any nonzero N.

[30] Richard Mollin, *Advanced Number Theory with Applications*, to appear.

[31] W. Narkiewicz, *Elementary and Analytic Theory of Algebraic Numbers*, 3rd edition, Springer-Verlag, 2004.

[32] Phost and Zassenhaus, *Algorithmic Algebraic Number Theory*, Cambridge University Press, 1983.

[33] E. V. Podsypanin, Length of the period of a quadratic irrational (in Russian), Studies in Number Theory, 5, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI)* 82 (1979) 95-99, 166; Engl. transl. in *J. Soviet Math.* 18 (1982) 919-923; MR0537024 (80h:12002).

[34] Keith Ramsay, Re: eˆ(pi*sqrt(163)) and all that (was: Re: "Almost" an Integer – eˆ?), post to sci.math, December 8, 2003.

[35] Legh Wilber Reid, *The Elements of the Theory of Algebraic Numbers*, The Macmillan Company, 1910. Available at http://historical.library.cornell.edu/cgi-bin/cul.math/docviewer?did=01200001

[36] Paulo Ribenboim, *Classical Theory of Algebraic Numbers*, Springer, New York, 2001.

[37] Paulo Ribenboim, *My Numbers, My Friends*, Springer–Verlag, New York, 2000.

[38] John Robertson, "Matthews' Method for Solving $ax^2 + bxy + cy^2 = N$," at www.jpr2718.org.

[39] John Robertson, "Solving the generalized Pell equation," at www.jpr2718.org.

[40] Pierre Samuel, *Algebraic Theory of Numbers*, translated by Allan J. Silberger, Houghton Mifflin Company, Boston, 1970.

[41] H. M. Stark, A complete determination of the complex quadratic fields of class-number one, *Michigan Math. J.* **14** 1967, pp. 1-27.

[42] Ian Stewart and David Tall, *Algebraic Number Theory and Fermat's Last Theorem*, Third Edition, A. K. Peters, 2002.

[43] H. P. F. Swinnerton-Dyer, *A Brief Guide to Algebraic Number Theory*, London Mathematical Society Student Texts, 50, Cambridge Universtiy Press, 2001.

[44] David C. Terr, A Modification of Shanks' Baby-Step Giant-Step Algorithm, *Mathematics of Computation* **69**, No. 230, 2000, pp. 767-773.

[45] André Weil, *Number Theory, An approach through history from Hammurapi to Legendre*, Birkhäuser, Boston, 2001.

[46] André Weilert, Two Efficient Algorithms for the Computation of Ideal Sums in Quadratic Orders, *Mathematics of Computation* **75**, No. 254, 2006, pp. 941-981.