Contents

9 The Geometry of Numbers **1**

9.1 Minkowski's Convex-Body Theorem and Applications 1

9.1.1 Minkowski's Convex-Body Theorems 2

9.1.2 Sums of Two and Four Squares 4

9.1.3 Sums of Three Squares 6

9.2 Ideal Class Groups of Quadratic Integer Rings 8

9.2.1 The Ideal Class Group 8

9.2.2 Minkowski's Bound 11

9.3 Binary Quadratic Forms 14

9.3.1 Representations of Integers by Binary Quadratic Forms 15

9.3.2 Equivalence of Binary Quadratic Forms 16

9.3.3 Composition of Binary Quadratic Forms 21

9.3.4 Dirichlet's Composition of Quadratic Forms 23

9.3.5 Quadratic Forms and Ideal Class Groups 24

9 The Geometry of Numbers

In this chapter, we discuss some applications of geometric ideas in number theory, which is referred to under the broad heading of the "geometry of numbers": our general theme will be to use properties of lattices in $\mathbb{R}^n$ (or $\mathbb{C}$) to study questions of number-theoretic interest. We will first establish some geometric preliminaries that lead to our main geometric result, Minkowski's convex-body theorem, which we then apply to classify integers that are the sum of two, three, and four squares.

Next, we introduce the class group of a quadratic integer ring, which quantifies quite precisely the degree to which $\mathcal{O}_{\sqrt{D}}$ fails to have unique factorization. We then establish Minkowski's bound on ideal class representatives, which provides one of the most effective ways to compute the structure of the class group, and give examples of how to use it to compute ideal class groups.

Afterwards, we detour somewhat to discuss representations of integers by binary quadratic forms, and show how binary quadratic forms also possess a composition law that is closely tied to the structure of the ideal class group of the associated quadratic integer ring.

9.1 Minkowski's Convex-Body Theorem and Applications

- First, we review some basic terminology for sets in $\mathbb{R}^n$.
    - We will denote the set of all points in $\mathbb{R}^n$ all of whose coordinates are integers by $\mathbb{Z}^n$.
- Definition: A set B in $\mathbb{R}^n$ is convex if, for any x and y in B, all points on the line segment joining x and y are also in B.

- Example: The n-ball of radius r centered at the origin in $\mathbb{R}^n$, given by the points $(x_1, x_2, \ldots, x_n)$ with $x_1^2 + x_2^2 + \cdots + x_n^2 \leq r^2$, is a convex set.
- Example: The unit cube, given by the points $(x_1, x_2, \ldots, x_n)$ with $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$, is a convex set.

- We may distinguish three different classes of points in $\mathbb{R}^n$ relative to B, based on their behaviors when we draw balls around them.

1. If we can draw a ball around P that is entirely contained in B, then P is called an interior point of B.
2. If we can draw a ball around P that is entirely contained in B^c, the complement of B, then we call P an exterior point of B. (Equivalently, it is an interior point of B^c.)
3. Otherwise, no matter what size of ball we draw, it will always contain some points in B and some points in B^c. Points with this property are called boundary points.

- Definition: The interior of the set B, denoted $\text{int}(B)$, is the set of its interior points. A set B is open if all its points are interior points. A set B is closed if its complement is open.
  - There are various equivalent conditions for closed and open sets. For example, B is closed if and only if it contains all of its boundary points if and only if any convergent sequence $\{a_n\}_{n\geq 1}$ of points in B has its limit in B.
  - Example: The open unit n-ball in $\mathbb{R}^n$, given by the points $(x_1, x_2, \ldots, x_n)$ with $x_1^2 + x_2^2 + \cdots + x_n^2 < 1$, is indeed an open set, since any point in this set is an interior point.
  - Example: The closed unit n-ball in $\mathbb{R}^n$, given by the points $(x_1, x_2, \ldots, x_n)$ with $x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1$, is not an open set, since any point with $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$ is a boundary point, rather than an interior point. It is in fact a closed set, since its complement is the open set with $x_1^2 + x_2^2 + \cdots + x_n^2 > 1$.
  - It is fairly straightforward to see that if B is an n-dimensional convex set in $\mathbb{R}^n$, then its interior is also convex.

- Definition: A set B in $\mathbb{R}^n$ is symmetric about the origin if, for any x in B, the point $-x$ is also in B.
  - Example: The n-ball of radius r centered at the origin in $\mathbb{R}^n$ is symmetric about the origin.
  - Non-Example: The unit cube, given by the points $(x_1, x_2, \ldots, x_n)$ with $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$, is not symmetric about the origin.

9.1.1 Minkowski's Convex-Body Theorems

- Our goal now is to prove that if a convex set is sufficiently nice and has a sufficiently large n-measure (i.e., n-volume), it must contain a lattice point.
- We first show the following result, which is sometimes called Blichfeldt's principle:
- Proposition (Blichfeldt's Principle): If S is a bounded measurable set in $\mathbb{R}^n$ whose n-measure is greater than 1, then there exist two points x and y in S such that $x - y$ has integer coordinates.
  - We will also remark that if we add the assumption that S is closed, then the conclusion holds also if the measure of S is equal to 1.
  - Proof: The idea is essentially to use the pigeonhole principle.
  - For each lattice point $a = (a_1, \cdots, a_n)$, let B_a be the "box" consisting of the points $(x_1, \cdots, x_n)$ whose coordinates satisfy $a_i \leq x_i < a_{i+1}$, and let $S_a = S \cap B_a$ be the intersection of S with B_a.
  - Since each point of S lies in exactly one box B_a, we have $\sum_{a\in\mathbb{Z}^n} \text{vol}(S_a) = \text{vol}(S)$.
  - Now let S_a^* be the set S_a translated by the vector $-a$: this translation preserves measure and moves B_a to B_0.
  - Then $\sum_{a\in\mathbb{Z}^n} \text{vol}(S_a^*) = \text{vol}(S) > 1$. But since all of the sets S_a^* lie inside B_0 which has volume 1, there must be some overlap.

- If $S^*_{a_1} \cap S^*_{a_2}$ contains some point P, then $P + a_1 \in S$ and also $P + a_2 \in S$. Taking $x = P + a_1$ and $y = P + a_2$ we see that $x - y = a_1 - a_2$ has integer coordinates, as claimed.
- Remark: This proof can also be formulated analytically in terms of the characteristic function $\chi_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \notin B \end{cases}$, which is integrable by the hypothesis that B is a measurable set. If we write $\psi(x) = \sum_{v \in \mathbb{Z}^n} \chi_B(x + v)$, then ψ is bounded because B is bounded so there are only finitely many nonzero terms for any $v \in \mathbb{Z}^n$. We may then integrate both sides and change the order of integration and summation (because the sum is a finite sum of nonnegative terms) and use the translation-invariance of the measure on $\mathbb{R}^n$ to see that $\int_{[0,1]^n} \psi(x)\,dx = \int_{[0,1]^n} \sum_{v \in \mathbb{Z}^n} \chi_B(x+v)\,dx = \sum_{v \in \mathbb{Z}^n} \int_{[0,1]^n} \chi_B(x+v)\,dx = \sum_{v \in \mathbb{Z}^n} \int_{[0,1]^n + v} \chi_B(x)\,dx = \int_{\mathbb{R}^n} \chi_B(x)\,dx$, and this last integral is simply the measure of B.

- Now we may prove our first main result:

- Theorem (Minkowski's Convex Body Theorem): Let B be a convex open set in $\mathbb{R}^n$ that is symmetric about the origin and whose n-measure is greater than 2^n. Then B contains a nonzero point all of whose coordinates are integers.

  - We will remark that the bound here is sharp, in the sense that we cannot lower the bound to any number less than 2^n. We will also remark that if we replace "open" with "closed", then we can weaken the inequality to "n-measure greater than or equal to 2^n".
  - Proof: Suppose B is a convex open set symmetric about 0 of measure $> 2^n$ and let $\frac{1}{2}B = \{\frac{1}{2}x : x \in B\}$.
  - Notice that since $\text{vol}(B) > 2^n$, we have $\text{vol}(\frac{1}{2}B) > 1$. Apply Blichfeldt's principle to the set $\frac{1}{2}B$: we obtain distinct points $x, y \in \frac{1}{2}B$ such that $x - y$ has integer coordinates.
  - Then $2x \in B$ and $2y \in B$. Furthermore, since B is symmetric about the origin, $-2y \in B$.
  - Then because B is convex, the midpoint of the line segment joining $2x$ and $-2y$ lies in B.
  - This midpoint $x - y$ yields the desired nonzero point in B whose coordinates are integers.

- The result of Minkowski's theorem does not apply merely to the lattice $\mathbb{Z}^n$ of points having integer coordinates.

  - If $v_1, \ldots, v_n$ are ($\mathbb{R}$-)linearly independent vectors in $\mathbb{R}^n$, the set Λ of vectors of the form $c_1v_1 + \cdots + c_nv_n$, where each $c_i \in \mathbb{Z}$, is called a lattice.
  - A fundamental region for this lattice can be obtained by drawing all of the vectors $v_1, \ldots, v_n$ outward from the origin, and then filling them in to create a "skew box". The points in this fundamental region give unique representatives for the quotient group $\mathbb{R}^n/\Lambda$, up to an appropriate choice of representatives on the boundary of the region.
  - A basic fact from linear algebra, which essentially amounts to the definition of the determinant, says that the n-volume of the fundamental domain is equal to the absolute value of the determinant of the matrix whose columns are the vectors $v_1, \ldots, v_n$ when expressed in terms of the standard basis of $\mathbb{R}^n$.
  - One may prove this fact by direct manipulations, or (more structurally), one may do it by observing that the signed n-volume of the fundamental domain satisfies the same properties as the determinant: interchanging two vectors scales the signed volume by -1, scaling a vector scales the signed volume by the same amount, adding a multiple of one vector to another does not change the signed volume, and the signed volume for the standard basis is 1. The determinant can be proven to be the only function satisfying these four properties, and so the signed volume is equal to the determinant.

- By changing basis, we obtain a version of Minkowski's theorem for general lattices:

- Theorem (Minkowski's Theorem for General Lattices): Let Λ be any lattice in $\mathbb{R}^n$ whose fundamental domain has volume V. If B is any open convex centrally-symmetric region in $\mathbb{R}^n$ whose volume is $> 2^nV$, then B contains a nonzero point of Λ.

  - Proof: Apply the linear transformation T sending the basis vectors of Λ to the standard basis of $\mathbb{R}^n$.
  - Linear transformations preserve open sets, convex sets, and central symmetry, so the image of B under this map is still open, convex, and centrally symmetric.

- The volume of $T(B)$ is equal to $1/V$ times the volume of B by the observation made about determinants above, so this new open convex centrally-symmetric set $T(B)$ has volume $> 2^n$.
- Applying the previous version of Minkowski's theorem to $T(B)$ yields that $T(B)$ contains a nonzero point all of whose coordinates are integers. This immediately implies that B contains a nonzero point of Λ, as required.

9.1.2 Sums of Two and Four Squares

- As our first application of Minkowski's convex body theorem, we will prove that every prime p congruent to 1 modulo 4 can be expressed as the sum of two squares.
  - We have previously established this result as a consequence of studying factorizations in $\mathbb{Z}[i]$.
  - The argument we will give using Minkowski's theorem is quite different.
- Theorem (Fermat's Christmas Theorem): If p is any prime congruent to 1 modulo 4, then there exist integers a and b such that $p = a^2 + b^2$.
  - This result was first explicitly noted by Girard in 1625, about 15 years before Fermat observed it (famously, in a letter to Mersenne on December 25th). Fermat also did not provide a proof; the first actual proof was given by Euler.
  - Proof: First, we observe that if $p \equiv 1 \pmod 4$ then -1 is a square modulo p: this follows by Euler's criterion $\left(\frac{-1}{p}\right) \equiv (-1)^{(p-1)/2} \equiv 1 \pmod p$. Alternatively, we could note that the group of nonzero residue classes modulo p is cyclic and has order $p-1$, and so since 4 divides $p-1$, there exists an element r of order 4: then r^2 has order 2, but the only element of order 2 modulo p is -1.
  - Now suppose $r^2 \equiv -1 \pmod p$, and consider the lattice Λ be the lattice in $\mathbb{R}^2$ spanned by the two vectors $\langle 1, r\rangle$ and $\langle 0, p\rangle$.
  - The determinant of these two vectors is p, so the volume (area) of the fundamental domain is p.
  - Let B be the interior of the disc $x_1^2 + x_2^2 < 2p$ in $\mathbb{R}^2$, and observe that B is open, convex and centrally-symmetric. From elementary geometry, the area of this disc is $2\pi p$.
  - Since $2\pi > 4$, the volume of B is larger than 2^2 times the volume of the fundamental domain of Λ, and so by Minkowski's theorem we conclude that there is a nonzero element
    $$\langle x_1, x_2\rangle = a\langle 1, r\rangle + b\langle 0, p\rangle$$
    of Λ in B.
  - But then
    $$\begin{aligned} x_1^2 + x_2^2 &= a^2 + (ra + bp)^2 \\ &\equiv a^2(1 + r^2) \pmod p \\ &\equiv 0 \pmod p \end{aligned}$$
    and since $\langle x_1, x_2\rangle$ is nonzero and has $x_1^2 + x_2^2 < 2p$, the only possibility is that $x_1^2 + x_2^2 = p$.
  - Thus, p is the sum of two squares, and we are done.
- We can give a similar kind of argument to establish that every positive integer n can be expressed as the sum of four squares, which is a result first proven by Lagrange.
  - This result was known, in most respects, to the ancient Greeks, and was stated explicitly by Bachet in 1621 in his translation notes of the works of Diophantus.
  - The first actual proof was given by Lagrange in 1770, and in 1834 Jacobi extended the result to give a formula for the number of representations of n as a sum of four squares.
  - Jacobi's result is as follows: if $\sigma(n)$ represents the sum of the divisors of n and $r_4(n)$ is the number of ways of writing n as the sum of four squares, then $r_4(n) = 8\sigma(n)$ if n is odd and $r_4(n) = 24\sigma(d)$ if $n = 2^k d$ (d odd) is even.

- We first show that if a, b are the sum of four squares, then so is ab:
- Lemma 1 (Products of Sums of Four Squares): If a and b are the sum of four squares, then so is ab.
  - Proof: This follows from the following identity:

$$\begin{aligned}(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2) &= (x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2+(x_1y_2-x_2y_1+x_3y_4-x_4y_3)^2 \\ &\quad +(x_1y_3-x_2y_4-x_3y_1+x_4y_2)^2+(x_1y_4+x_2y_3-x_3y_2-x_4y_1)^2\end{aligned}$$

  which can be verified simply by multiplying out and verifying that all of the cross-terms cancel.
- We will remark that, like the corresponding identity $(a^2+b^2)(c^2+d^2)=(ac+bd)^2+(ad-bc)^2$ for sums of two squares, which arises from the fact that the norm map on $\mathbb{Z}[i]$ is multiplicative, the four-squares identity also arises from a norm map on a ring.
  - Specifically, the ring in question is the ring $\mathbb{H}$ of (real) quaternions, which is a noncommutative ring. (The letter $\mathbb{H}$ is used because the quaternions were first described by Hamilton.)
  - Explicitly, $\mathbb{H}$ is the set of elements of the form $a+bi+cj+dk$, where a, b, c, d are real numbers, subject to the multiplication rules $i^2=j^2=k^2=ijk=-1$. (From these relations one can deduce explicitly that $ij=-ji=k$, $jk=-kj=i$, and $ki=-ik=j$.)
  - The conjugation map on $\mathbb{H}$ is $\overline{a+bi+cj+dk}=a-bi-cj-dk$, and the norm map is $N(q)=q\overline{q}$. One may compute explicitly that $N(a+bi+cj+dk)=a^2+b^2+c^2+d^2$, and the fact that the norm map is multiplicative (which is not obvious from its definition because the multiplication of quaternions is not commutative) amounts to the four-squares identity[1].
  - In fact, since the norm of a nonzero quaternion is nonzero, the multiplicativity of the norm map implies that every nonzero quaternion has a multiplicative inverse, which is to say, the quaternions form a division ring (which is the noncommutative analogue of a field).
  - Multiplication in this noncommutative manner using the letters i, j, and k might be familiar from the algebra of the cross product of vectors in 3-space: often the notation $\mathbf{i}=\langle 1,0,0\rangle$, $\mathbf{j}=\langle 0,1,0\rangle$, $\mathbf{k}=\langle 0,0,1\rangle$ is used for the basis vectors, and then for example one has $\mathbf{i}\times\mathbf{j}=\mathbf{k}$.
  - As a historical note, the development of quaternions actually predates the modern language of vectors by about 50 years, and so many of the classical results in physics (e.g., Maxwell's equations) predating the 20th century were originally written in terms of quaternions rather than vectors.
  - Due to their connection with geometry in 3 dimensions, the quaternions are often used in computer graphics, applied physics, and engineering, since they can be used to represent spatial rotations in 3-dimensional space far more efficiently than matrices.
- We now show that every prime can be written as the sum of four squares. To do this we require a simple lemma about sums of squares modulo a prime:
- Lemma 2: For any prime p, there exist integers r and s such that $r^2+s^2\equiv -1 \pmod p$. In other words, -1 is the sum of two squares modulo p.
  - Proof: If $p=2$ the result is obvious (take $r=1$, $s=0$). Now suppose p is odd.
  - From our results on quadratic residues, the set S of squares r^2 modulo p contains $(p+1)/2$ elements. Thus, the set T of elements of the form $-1-s^2$ also contains $(p+1)/2$ elements.
  - Since there are only p residue classes modulo p, the sets S and T must have a nontrivial intersection: then we have $r^2\equiv -1-s^2 \pmod p$ and so $r^2+s^2\equiv -1 \pmod p$, as required.
- We can now establish our main result:
- Theorem (Lagrange's Four-Square Theorem): If n is any positive integer, then n can be written as the sum of four squares.

[1] In fact, the conjugation map on $\mathbb{H}$ is an anti-involution: $\overline{q\cdot r}=\overline{r}\cdot\overline{q}$. From this one has $N(qr)=qr\overline{qr}=q(r\overline{r})\overline{q}=(q\overline{q})(r\overline{r})=N(q)N(r)$ using the fact that $r\overline{r}$ is a real number, so it commutes with q.

- Proof: By Lemma 1, it suffices to prove that every prime p can be written as the sum of four squares, so let p be a prime. By Lemma 2, there exist integers r and s such that $r^2 + s^2 \equiv -1 \pmod{p}$.
- Now let Λ be the lattice in $\mathbb{R}^4$ spanned by the four vectors $\langle p, 0, 0, 0\rangle$, $\langle 0, p, 0, 0\rangle$, $\langle r, s, 1, 0\rangle$, and $\langle s, -r, 0, 1\rangle$. It is a simple computation to see that the determinant of these four vectors is p^2, so the volume of the fundamental domain is p^2.
- Let B be the convex, centrally-symmetric open set in $\mathbb{R}^4$ defined by $x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2p$. The volume of this ball can be computed (either directly or via a more enlightened use of cylindrical coordinates[2], or spherical coordinates[3]) to be $2\pi^2 p^2$.
- Since the volume of B is larger than 2^4 times the volume of the fundamental domain of Λ (since $2\pi^2 p^2 > 16p^2$), we conclude that there is a nonzero element

$$\langle x_1, x_2, x_3, x_4\rangle = a\langle p, 0, 0, 0\rangle + b\langle 0, p, 0, 0\rangle + c\langle r, s, 1, 0\rangle + d\langle s, -r, 0, 1\rangle$$

  of Λ in B.
- But then

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 + x_4^2 &= (ap + cr + ds)^2 + (bp + cs - dr)^2 + c^2 + d^2 \\ &\equiv (c^2 + d^2)(1 + r^2 + s^2) \pmod{p} \\ &\equiv 0 \pmod{p} \end{aligned}$$

  and since $\langle x_1, x_2, x_3, x_4\rangle$ is nonzero and has $x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2p$, the only possibility is that $x_1^2 + x_2^2 + x_3^2 + x_4^2 = p$.
- Thus, p is the sum of four squares, and we are done.

9.1.3 Sums of Three Squares

- As another application of Minkowski's convex-body theorem, we can characterize the integers that are the sum of three squares.
  - By testing small examples, one is rapidly led to the conjecture that n may be written as the sum of three squares if and only if n is not a power of 4 times an integer that is 7 modulo 8 (i.e., when $n \neq 4^a(8b+7)$ for some a, b).
  - Indeed, we showed in our discussion of Diophantine equations that if $n = 4^a(8b + 7)$ then n is not the sum of three squares: for $a = 0$ this follows immediately by considering n modulo 8, and then one may induct on a.
  - It remains to establish that integers not of this form can be written as the sum of three squares.
- We will establish one case of this theorem and then remark on the modifications necessary to establish the other cases.
  - Unlike in the case for sums of two squares and sums of four squares, the set of integers that are a sum of three squares is not closed under multiplication: both $3 = 1^2 + 1^2 + 1^2$ and $5 = 2^2 + 1^2 + 0^2$ are the sum of three squares, but $15 = 3 \cdot 5$ is not.
  - Therefore, we cannot simply reduce to the case of considering representations of primes, as we did for the case of sums of two and four squares.
  - Our approach will be to use Minkowski's theorem along with our characterization of integers that are expressible as the sum of two squares.

[2]For $x_1^2 + x_2^2 + x_3^2 + x_4^2 = R^2$ let r denote the distance between a point in the plane $x_1 = x_2 = 0$ and θ be the azimuthal angle. Then the intersection of the ball with the 2-dimensional plane obtained by fixing r and θ is a disc of radius $\sqrt{R^2 - r^2}$, so the 4-volume of the ball is given by the integral $\int_0^{2\pi}\int_0^R \pi(R^2 - r^2) \cdot r\, dr\, d\theta = \pi^2 R^4/2$.

[3]For $x_1^2 + x_2^2 + x_3^2 + x_4^2 = R^2$ take radial coordinate $\rho \in [0, R]$ and angular coordinates $\theta \in [0, 2\pi]$, $\varphi_1, \varphi_2 \in [0, \pi]$ with $x_1 = \rho\cos\theta\sin\varphi_1\sin\varphi_2$, $x_2 = \rho\sin\theta\sin\varphi_2\sin\varphi_2$, $x_3 = \rho\cos\varphi_1\sin\varphi_2$, $x_4 = \rho\cos\varphi_2$. The Jacobian is then $\rho^3\sin^2\varphi_2\sin\varphi_1$ and so the 4-volume is $\int_0^{2\pi}\int_0^{\pi}\int_0^{\pi}\int_0^R \rho^3\sin^2\varphi_2\sin\varphi_1\, d\rho d\varphi_2 d\varphi_1 d\theta = \pi^2 R^4/2$.

- We will treat the case where $n \equiv 3 \pmod 8$, since the exposition is easiest to give there. This argument was originally given by Ankeny in 1956:
- Theorem (Sums of 3 Squares, 3 Mod 8 Case): If n is a positive integer congruent to 3 modulo 8, then n is the sum of three squares.
  - Proof: First, we observe that there exists a prime $q \equiv 1 \pmod 4$ such that $-2q$ is a quadratic residue modulo n. This follows from Dirichlet's theorem on primes in arithmetic progressions, since saying $-2q$ is a quadratic residue modulo n is simply a congruence condition modulo n.
  - So, since $-2q$ is a quadratic residue modulo n, its reciprocal is also: say with $-1/(2q) \equiv t^2 \pmod n$.
  - Next, we observe that $\left(\frac{-2q}{n}\right) = \left(\frac{-2}{n}\right)\left(\frac{q}{n}\right) = \left(\frac{-2}{n}\right)\left(\frac{n}{q}\right) = \left(\frac{-n}{q}\right)$ by quadratic reciprocity for Jacobi symbols and the fact that $n \equiv 3 \pmod 8$ so that $\left(\frac{-2}{n}\right) = +1$ and that $q \equiv 1 \pmod 4$ so that $\left(\frac{n}{q}\right) = \left(\frac{q}{n}\right)$ and $\left(\frac{-1}{q}\right) = +1$.
  - Therefore, $-n$ is a quadratic residue modulo q, say with $-n \equiv b^2 \pmod q$ where we may assume that b is odd. This means $b^2 + n = qh'$ for some $h' \in \mathbb{Z}$, but now since $n \equiv 3 \pmod 8$, reducing both sides modulo 4 yields $h' \equiv 0 \pmod 4$, and so $h' = 4h$.
  - To summarize, we have integers q, b, and h such that $-1/(2q) \equiv t^2 \pmod n$ and $b^2 + n = 4qh$.
  - Let Λ be the lattice in $\mathbb{R}^3$ spanned by the vectors $\langle 2tq, \sqrt{2q}, 0\rangle$, $\left\langle tb, b/\sqrt{2q}, \sqrt{n/(2q)}\right\rangle$, and $\langle n, 0, 0\rangle$. It is a simple computation to see that the determinant of these three vectors is $n^{3/2}$, so the volume of the fundamental domain is $n^{3/2}$.
  - Now let B be the convex, centrally-symmetric open set in $\mathbb{R}^3$ defined by $x_1^2 + x_2^2 + x_3^2 < 2n$, whose volume is $\frac{4}{3}\pi(2n)^{3/2}$ since it is merely a sphere of radius $\sqrt{2n}$.
  - Since the volume of B is larger than 2^3 times the volume of the fundamental domain of Λ (since $\frac{4}{3}\pi \cdot 2^{3/2} > 8$), we conclude that there is a nonzero element
$$\langle R, S, T\rangle = x\left\langle 2tq, \sqrt{2q}, 0\right\rangle + y\left\langle tb, b/\sqrt{2q}, \sqrt{n/(2q)}\right\rangle + z\langle n, 0, 0\rangle$$
of Λ in B.
  - Then $R = 2tqx + tby + nz$, $S = x\sqrt{2q} + by/\sqrt{2q}$, $T = y\sqrt{n/(2q)}$, and so $R^2 + S^2 + T^2 = (2tqx + tby + nz)^2 + (\sqrt{2q}x + b/\sqrt{2q}y)^2 + (\sqrt{n/(2q)}y)^2 \equiv (t^2 + 1/(2q))(2qx + by)^2 \equiv 0 \pmod n$.
  - Notice also that $R^2 + S^2 + T^2$ is an integer, because it equals $R^2 + 2qx^2 + 2bxy + 2hy^2$, and all of these quantities are integers. Thus, $R^2 + S^2 + T^2 = n$.
  - Now we will show that the integer $N = S^2 + T^2 = 2qx^2 + 2bxy + 2hy^2$ is actually the sum of two integer squares, which will complete the proof because then $n = R^2 + N$ is then the sum of three squares.
  - So suppose p is an odd prime dividing N to an odd power, meaning that p^{2a+1} divides N but p^{2a+2} does not: we wish to show that $p \equiv 1 \pmod 4$.
  - First suppose p does not divide n: then $n \equiv R^2 \pmod p$ and so $\left(\frac{n}{p}\right) = +1$.
  - Also, if $p = q$ then since $-2q$ is a quadratic residue modulo n we have $\left(\frac{-n}{p}\right) = +1$. Otherwise if $p \neq q$ then $2qN = 4q^2x^2 + 4bqxy + 4qhy^2 = (2qx + by)^2 + ny^2$, and the only way that this quantity can be divisible by an odd power of p is if there is a nonzero solution to $e^2 + nf^2 \equiv 0 \pmod p$, which forces $-n$ to be a quadratic residue modulo p.
  - In both cases we have $\left(\frac{n}{p}\right) = +1$ and $\left(\frac{-n}{p}\right) = +1$, so $\left(\frac{-1}{p}\right) = +1$ and so $p \equiv 1 \pmod 4$.

- Now suppose p does divide n. Then $R^2 + N = n$, so since p divides N it must also divide R. Rewriting the equation as $R^2 + \frac{1}{2q}\left[(2qx+by)^2 + ny^2\right] = n$, we see that p must also divide $2qx + by$. Dividing through by p and then reducing modulo p yields $\frac{1}{2q} \cdot \frac{n}{p} y^2 \equiv \frac{n}{p} \pmod{p}$, so since n/p is nonzero modulo p as n is squarefree, we get $y^2 \equiv 2q \pmod{p}$ and thus $\left(\frac{2q}{p}\right) = +1$. Since we assumed at the very beginning that $\left(\frac{-2q}{p}\right) = +1$, this implies $\left(\frac{-1}{p}\right) = +1$ and so $p \equiv 1 \pmod{4}$ once again.
- Thus, all odd primes that exactly divide N to an odd power are congruent to 1 modulo 4, so by our characterization of sums of two squares, this means N is the sum of two squares, and so $n = R^2 + N$ is the sum of three squares.

- We will remark that the proof for the case $n \equiv 3 \pmod{8}$ can be adapted to establish the other cases $n \equiv 1, 2, 5, 6 \pmod{8}$ as well, by suitable minor modifications on the conditions taken at the beginning.

9.2 Ideal Class Groups of Quadratic Integer Rings

- We now discuss some additional properties of the ideals in quadratic integer rings: we will introduce the ideal class group, and then use Minkowski's convex-body theorem to establish that the ideal class group of any quadratic integer ring is finite. We will then use our results to compute explicitly the ideal class groups of various quadratic integer rings.

9.2.1 The Ideal Class Group

- As we have already discussed, a quadratic integer ring $\mathcal{O}_{\sqrt{D}}$ has unique factorization if and only if it is a principal ideal domain, and (thus) any examples of non-unique factorization necessarily arise from nonprincipal ideals.
    - Our goal now is to quantify more precisely how "non-unique" the non-unique factorization in $\mathcal{O}_{\sqrt{D}}$ can be, which is (in a sense we will make precise) the same as asking about the various possible classes of nonprincipal ideals.

- To motivate the ideas, consider the quadratic integer ring $\mathbb{Z}[\sqrt{-5}]$, which we have shown not to be a PID by constructing explicit nonprincipal ideals $I_2 = (2, 1+\sqrt{-5})$, $I_3 = (3, 1+\sqrt{-5})$, and $I_3' = (3, 1-\sqrt{-5})$.
    - Notice, however, that many of the possible products of these nonprincipal ideals are principal: for instance as we have already computed, we have $I_2^2 = (4, 2(1+\sqrt{-5}), -4+2\sqrt{-5}) = (2)$, $I_2 I_3 = (6, 2(1+\sqrt{-5}), 3(1+\sqrt{-5}), -4+2\sqrt{-5}) = (1+\sqrt{-5})$, $I_2 I_3' = (1-\sqrt{-5})$, and $I_3 I_3' = (3)$
    - With a bit more work we can also see $I_3^2 = (9, 3+3\sqrt{-5}, -4+2\sqrt{-5}) = (2-\sqrt{-5})$ since $2-\sqrt{-5}$ is in this ideal since $2-\sqrt{-5} = 9 - (3+3\sqrt{-5}) + (-4+2\sqrt{-5})$, and also $2-\sqrt{-5}$ divides each of the generators.
    - Conjugating this calculation then shows $(I_3')^2 = (2+\sqrt{-5})$.
    - By searching for factorizations of other integer primes in $\mathbb{Z}[\sqrt{-5}]$ we can construct additional nonprincipal ideals, such as $I_7 = (7, 3+\sqrt{-5})$ and its conjugate $I_7' = (7, 3-\sqrt{-5})$.
    - If we then try computing various products, like $I_2 I_3$ or $I_3 I_7'$, we will discover that no matter which pair of nonprincipal ideals we multiply together, the result will always be principal.
    - For example, for $I_2 I_7 = (14, 6+2\sqrt{-5}, 7+7\sqrt{-5}, -2+4\sqrt{-5})$, we see this ideal contains $3+\sqrt{-5} = 2(6+2\sqrt{-5}) - (7+7\sqrt{-5}) - (-2+4\sqrt{-5})$ and also each element in the ideal is divisible by $3+\sqrt{-5}$, so in fact $I_2 I_7 = (3+\sqrt{-5})$.
    - Similarly, for $I_3 I_7' = (3, 1+\sqrt{-5})(7, 3-\sqrt{-5}) = (21, 9-3\sqrt{-5}, 7+7\sqrt{-5}, 8+2\sqrt{-5})$, we see this ideal contains $4+\sqrt{-5} = 21 + (7+7\sqrt{-5}) - 3(8+2\sqrt{-5})$ and also each element in the ideal is divisible by $4+\sqrt{-5}$, so in fact $I_3 I_7' = (4+\sqrt{-5})$.

- These calculations suggest that there might actually be only one type of nonprincipal ideal in $\mathbb{Z}[\sqrt{-5}]$, up to an appropriate notion of equivalence of ideals.
- We can also see that there is a natural composition operation on ideals, given by ideal multiplication, which seems to behave nicely with respect to this equivalence.

- We will now make all of this precise:

- Definition: Let $R = \mathcal{O}_{\sqrt{D}}$ be a quadratic integer ring. We define a relation $\sim$ on the set of nonzero ideals of R by saying $I \sim J$ if $(a)I = (b)J$ for some nonzero principal ideals (a) and (b).

  - Intuitively, we declare two ideals to be equivalent if they differ by a principal ideal factor.
  - Example: Inside $\mathbb{Z}[i]$, since every nonzero ideal I is principal, we have $I \sim (1)$ for all nonzero I.
  - Example: Inside $\mathcal{O}_{\sqrt{-5}}$, with the notation as above, we have $I_2 \sim I_3$: since $I_2^2 = (2)$ and $I_2 I_3 = (1+\sqrt{-5})$, we see that $(1+\sqrt{-5})I_2^2 = (2)I_2 I_3$ and thus cancelling I_2 gives $(1+\sqrt{-5})I_2 = (2)I_3$.

- The relation $\sim$ is, perhaps unsurprisingly, an equivalence relation, and we can also use it to detect whether $\mathcal{O}_{\sqrt{D}}$ is a principal ideal domain:

- Proposition (Properties of Ideal Classes): Suppose $R = \mathcal{O}_{\sqrt{D}}$ is a quadratic integer ring. Then the following properties hold for the relation $I \sim J$ if $(a)I = (b)J$ for some nonzero $a, b \in R$:

  1. The relation is an equivalence relation on nonzero ideals. The equivalence classes under this relation are called ideal classes.
     - Proof: Clearly $I \sim I$ since $(1)I = (1)I$.
     - Also, if $I \sim J$ then $(a)I = (b)J$, and then by interpreting this as $(b)J = (a)I$ we see $J \sim I$.
     - Finally, if $I \sim J$ and $J \sim K$ then $(a)I = (b)J$ and $(c)J = (d)K$, and so $(ac)I = (bc)J = (bd)K$ meaning $I \sim K$.
  2. We have $I \sim (1)$ if and only if I is principal. Thus, $\mathcal{O}_{\sqrt{D}}$ is a principal ideal domain if and only if $I \sim (1)$ for all nonzero ideals I of $\mathcal{O}_{\sqrt{D}}$.
     - Proof: If $I \sim (1)$ then $(a)I = (b)$ for some nonzero a and b. This equality requires that a divides b, say with $b = ka$. Then cancelling (a) yields $I = (k)$, so I is principal.
     - The second statement follows immediately from the first, since the zero ideal (0) is principal and thus being a PID only requires having all nonzero ideals be principal.
  3. Multiplication of ideals respects ideal classes: if $I \sim I'$ and $J \sim J'$, then $IJ \sim I'J'$.
     - Proof: Suppose $(a)I = (b)I'$ and $(c)J = (d)J'$. Multiplying these relations yields $(ac)IJ = (bd)I'J'$, so $IJ \sim I'J'$.

- We have a natural multiplication operation on ideals, which makes the set of nonzero ideals into a semigroup. Because the multiplication of ideals respects ideal classes, the set of ideal classes inherits this multiplication operation; even better, it becomes an actual group:

- Proposition (Group Operation on Ideals): Let $R = \mathcal{O}_{\sqrt{D}}$ be a quadratic integer ring and let $[I]$ represent the ideal class of an ideal I of R. Then the operation $[I] \cdot [J] = [IJ]$ makes the set of ideal classes into an abelian group. This group is called the ideal class group of R (or often, just the class group of R).

  - By a mild abuse of terminology, it is also very common to refer to "the ideal class group of the quadratic field $\mathbb{Q}(\sqrt{D})$" as meaning the ideal class group of its ring of integers $\mathcal{O}_{\sqrt{D}}$.
  - Proof: First, the operation is well-defined by (3) from the proposition above.
  - The operation is associative and commutative because multiplication of ideals in $\mathcal{O}_{\sqrt{D}}$ is associative and commutative: $([I][J])[K] = [IJ][K] = [IJK] = [I][JK] = [I]([J][K])$ and $[I][J] = [IJ] = [JI] = [J][I]$.
  - The ideal class of (1) is a multiplicative identity, since $(1)I = I$ and so $[(1)][I] = [I]$ for all I.
  - Finally, every ideal class has an inverse: as we proved, for any ideal I the product $I \cdot \overline{I}$ is a principal ideal (a), and so $[I][\overline{I}] = [(a)] = [(1)]$.

- We see that the ideal classes have the structure of an abelian group under multiplication. However, by itself, this fact does not yield very much useful information about the ideal classes: what we really want to do is compute the structure of the ideal class group.
- Our first major result in this direction is that the class group is always finite:
- Proposition (Properties of the Class Group): Suppose $R = \mathcal{O}_{\sqrt{D}}$ is a quadratic integer ring and let $[I]$ denote the ideal class of an ideal I of R. Then the following are true:

  1. If I is a nonzero ideal of R, then I contains a nonzero element α such that $|N(\alpha)| \leq (|D|+1)N(I)$.
     - Proof: Let $m = \lfloor \sqrt{N(I)} \rfloor$ so that $m^2 \leq N(I) < (m+1)^2$.
     - Then since the cardinality of R/I is $N(I) < (m+1)^2$, by the pigeonhole principle at least two of the $(m+1)^2$ elements $\{a + b\sqrt{D} : 0 \leq a, b \leq m\}$ in R must be congruent modulo I, so their difference is in I.
     - Thus, there exists a nonzero element $\gamma \in I$ of the form $a + b\sqrt{D}$ where $-m \leq a, b \leq m$.
     - Then $|N(\gamma)| = \left|(a+b\sqrt{D})(a-b\sqrt{D})\right| = \left|a^2 - Db^2\right| \leq \left|a^2\right| + \left|Db^2\right| = m^2(|D|+1) \leq (|D|+1)N(I)$, as claimed.
     - Remark: When $D \equiv 1 \pmod 4$ this bound can be improved by working instead with the elements of the form $a + b\omega$ where ω is a generator of the quadratic integer ring. However, since we will be improving this result shortly, we will not bother making this calculation.
  2. Every ideal class of R contains an ideal J such that $N(J) \leq |D| + 1$.
     - Proof: Let $\mathcal{C}$ be an ideal class and let I be any ideal in the inverse class $\mathcal{C}^{-1}$.
     - By (1), there exists a nonzero element $\alpha \in I$ such that $N(\alpha) \leq (|D|+1)N(I)$. Because $\alpha \in I$, by the equivalence of divisibility and containment we see that I divides (α) and so $(\alpha) = IJ$ for some ideal J.
     - Taking norms yields $N(\alpha) = N(I)N(J)$, so $N(J) = \dfrac{N(\alpha)}{N(I)} \leq |D| + 1$. Finally, taking ideal classes gives $[1] = [(\alpha)] = [I][J]$ so $J \in [I]^{-1} = (\mathcal{C}^{-1})^{-1} = \mathcal{C}$, as required.
  3. The ideal class group of $\mathcal{O}_{\sqrt{D}}$ is finite.
     - Proof: By (2), every ideal class contains some ideal J with $N(J) \leq |D| + 1$.
     - But there are only finitely many possible ideals J with $N(J) \leq |D| + 1$: there are only finitely many possible prime ideals that could occur in the prime factorization of J (namely, the primes of norm at most $|D| + 1$) and the power to which each such ideal can occur is bounded (since a prime power P^a has norm $N(P)^a$, we must have $a \leq \log_{N(P)}(|D| + 1)$ for all such P).
     - Thus, the ideal classes are all represented by a finite list of ideals, so there are finitely many ideal classes.

- Definition: If D is a squarefree integer not equal to 1, the class number of the quadratic integer ring $\mathcal{O}_{\sqrt{D}}$ is the order of the ideal class group of $\mathcal{O}_{\sqrt{D}}$. The class number is often written as $h(D)$.
  - As we noted earlier, the class number of $\mathcal{O}_{\sqrt{D}}$ is equal to 1 if and only if $\mathcal{O}_{\sqrt{D}}$ is a principal ideal domain. A larger class number corresponds to having more inequivalent types of non-unique factorizations.
  - Our proof of (3) in the proposition above gives us an explicit way to calculate the ideal class group of $\mathcal{O}_{\sqrt{D}}$: we need only compute all of the possible prime ideals having norm at most $|D| + 1$, and then determine the resulting structure of these ideals under multiplication.
- Example: Show that the class group of $\mathbb{Z}[\sqrt{2}]$ is trivial and deduce that $\mathbb{Z}[\sqrt{2}]$ is a principal ideal domain.
  - From the proposition, we know that any ideal class contains an ideal J of norm at most 3.
  - Then the only possible prime divisors of the norm are 2 and 3, so the only possible prime ideal divisors of J are the primes lying above 2 and 3.
  - Using the Dedekind-Kummer factorization theorem shows that in $\mathbb{Z}[\sqrt{2}]$ we have $(2) = (\sqrt{2})^2$ while the ideal (3) is inert and has norm 9, so the only possible ideals J are (1) of norm 1 and $(\sqrt{2})$ of norm 2.

- Since both of these ideals are principal, we conclude that every ideal of $\mathbb{Z}[\sqrt{2}]$ is principal and so $\mathbb{Z}[\sqrt{2}]$ is a principal ideal domain.

- Example: Show that the class group of $\mathbb{Z}[\sqrt{-5}]$ has order 2.
  - From the proposition, we know that any ideal class contains an ideal J of norm at most 6.
  - Then the only possible prime divisors of the norm are 2, 3, and 5 so the only possible prime ideal divisors of J are the primes lying above 2, 3, and 5.
  - Using the Dedekind-Kummer factorization theorem (or appealing to our analysis from earlier) shows that in $\mathbb{Z}[\sqrt{-5}]$ we have $(2) = (2, 1+\sqrt{-5})^2$, $(3) = (3, 1+\sqrt{-5})(3, 1-\sqrt{-5})$, and $(5) = (\sqrt{-5})^2$.
  - Thus, the possible prime ideals dividing J are $I_2 = (2, 1+\sqrt{-5})$ of norm 2, $I_3 = (3, 1+\sqrt{-5})$ and $I_3' = (3, 1-\sqrt{-5})$ both of norm 3, and $I_5 = (\sqrt{-5})$ of norm 5.
  - As we have previously shown, the ideal I_2 is not principal, so since $I_2^2 = (2)$ we see that $[I_2]$ is an element of order 2 in the class group.
  - We have also previously shown that $I_2 I_3 = (1+\sqrt{-5})$, so $[I_3] = [I_2]^{-1} = [I_2]$, and then since $I_3 I_3' = (3)$ we see $[I_3'] = [I_2]$ as well.
  - Thus, since I_5 is principal, we see that all of the nonprincipal ideals lie in the same class (namely, the class $[I_2]$) and so the class group of $\mathbb{Z}[\sqrt{-5}]$ has order 2.

- Example: Determine the class group of $\mathbb{Z}[\sqrt{6}]$ and decide whether it is a principal ideal domain.
  - From the proposition, we know that any ideal class contains an ideal J of norm at most 7.
  - Then the only possible prime divisors of the norm are 2, 3, 5, and 7, so the only possible prime ideal divisors of J are the primes lying above 2, 3, 5, and 7.
  - Using the Dedekind-Kummer factorization theorem shows that in $\mathbb{Z}[\sqrt{6}]$ we have $(2) = (2, \sqrt{6})^2$, $(3) = (3, \sqrt{6})^2$, $(5) = (5, 1+\sqrt{6})(5, 1-\sqrt{6})$, and (7) is inert.
  - Thus the possible prime ideals dividing J are $I_2 = (2, \sqrt{6})$ of norm 2, $I_3 = (3, \sqrt{6})$ of norm 3, and $I_5 = (5, 1+\sqrt{6})$ and $I_5' = (5, 1-\sqrt{6})$ both of norm 5. (Note that $I_7 = (7)$ cannot divide J since its norm is 49.)
  - In fact we can see I_2 is principal, since it contains $2-\sqrt{6}$ and both 2 and $\sqrt{6}$ are divisible by $2-\sqrt{6}$. Likewise, I_3 is principal since it contains $3-\sqrt{6}$ and both 3 and $\sqrt{6}$ are divisible by $3-\sqrt{6}$, and also I_5 (hence also its conjugate I_5') is principal since $1+\sqrt{6}$ divides 5.
  - Thus, no matter what the ideal J is, it is principal, and so the class group of $\mathbb{Z}[\sqrt{6}]$ is $\boxed{\text{trivial}}$, and $\mathbb{Z}[\sqrt{6}]$ is a PID.

9.2.2 Minkowski's Bound

- Our ability to compute the class group relies upon being able to get a good estimate on the norm of the smallest nonzero element in an ideal I.
  - If D is negative, then the elements of the quadratic integer ring $\mathcal{O}_{\sqrt{D}}$ naturally form a lattice in the complex plane. Then any nonzero ideal I will form a sublattice, to which we can then apply Minkowski's convex-body theorem to obtain an element of small norm.
  - If D is positive, we will have to take a slightly different approach to embed $\mathcal{O}_{\sqrt{D}}$ into $\mathbb{R}^2$ as a lattice, but we will be able to do essentially the same thing. The idea in this case is instead to map an element $\alpha \in \mathcal{O}_{\sqrt{D}}$ to the point $(\alpha, \overline{\alpha}) \in \mathbb{R}^2$.

- In order to pose the results consistently, we will introduce a little bit more terminology:

- Definition: If $\mathcal{O}_{\sqrt{D}}$ is a quadratic integer ring, the discriminant of $\mathcal{O}_{\sqrt{D}}$ is defined as $\Delta = \begin{cases} 4D & \text{if } D \equiv 2, 3 \pmod 4 \\ D & \text{if } D \equiv 1 \pmod 4 \end{cases}$.
  - This quantity is the discriminant of the minimal polynomial of the generator of $\mathcal{O}_{\sqrt{D}}$ (whence the name).

- Definition: Suppose D is a squarefree integer not equal to 1. We define the Minkowski embedding $\varphi : \mathcal{O}_{\sqrt{D}} \to \mathbb{R}^2$ as follows: if $D < 0$, we map the element $a+b\sqrt{D} \in \mathcal{O}_{\sqrt{D}}$ to $(a, b\sqrt{|D|})$, and if $D > 0$, we map the element $a + b\sqrt{D} \in \mathcal{O}_{\sqrt{D}}$ to $(a + b\sqrt{D}, a - b\sqrt{D})$.
  - It is easy to see that the Minkowski map φ is a homomorphism of additive groups (i.e., it is $\mathbb{Z}$-linear), and so the image of $\mathcal{O}_{\sqrt{D}}$ will be a 2-dimensional lattice spanned by the vectors $\varphi(1)$ and $\varphi(\omega)$ where ω is a generator of $\mathcal{O}_{\sqrt{D}}$.
  - If $D < 0$, the Minkowski embedding is simply the result of identifying the elements of $\mathcal{O}_{\sqrt{D}}$ as points in the complex plane. The image of the embedding is a lattice spanned by $\varphi(1) = (1, 0)$ and $\varphi(\omega)$, which is either $(0, \sqrt{|D|})$ or $(1/2, \sqrt{|D|}/2)$ according to whether $D \equiv 2, 3$ or $D \equiv 1 \pmod 4$.
  - If $D > 0$, the image of the Minkowski embedding is still a lattice, since it is spanned by the linearly-independent vectors $\varphi(1) = (1, 1)$ and $\varphi(\omega) = (\omega, \overline{\omega})$, which is either $(\sqrt{D}, -\sqrt{D})$ or $(\frac{1+\sqrt{D}}{2}, \frac{1-\sqrt{D}}{2})$.
- Now we can prove Minkowski's bound:
- Theorem (Minkowski's Bound): Suppose D is a squarefree integer not equal to 1, let Δ be the discriminant of $\mathcal{O}_{\sqrt{D}}$, and let $\varphi : \mathcal{O}_{\sqrt{D}} \to \mathbb{R}^2$ be the Minkowski embedding. Then the following hold:
  1. The area of the fundamental domain for $\Lambda = \varphi(O_{\sqrt{D}})$ is equal to $\begin{cases} \sqrt{\Delta} & \text{if } D > 0 \\ \frac{1}{2}\sqrt{|\Delta|} & \text{if } D < 0 \end{cases}$.
     - Proof: The area of the fundamental domain for this lattice is equal to the determinant of $\varphi(1)$, $\varphi(\omega)$, where ω is a generator for the quadratic integer ring.
     - If $D < 0$, we have $\varphi(1) = (1, 0)$ and $\varphi(\omega) = (\mathrm{Re}(\omega), \mathrm{Im}(\omega))$ is either $(0, \sqrt{|D|})$ or $(1/2, \sqrt{|D|}/2)$ according to whether $D \equiv 2, 3$ or $D \equiv 1 \pmod 4$. Then the determinant is $\sqrt{|D|}$ or $\sqrt{|D|}/2$ respectively, and in both cases we see that the area equals $\frac{1}{2}\sqrt{|\Delta|}$.
     - If $D > 0$, we have $\varphi(1) = (1, 1)$ and $\varphi(\omega) = (\omega, \overline{\omega})$ is either $(\sqrt{D}, -\sqrt{D})$ or $(\frac{1+\sqrt{D}}{2}, \frac{1-\sqrt{D}}{2})$. Then the determinant is $2\sqrt{D}$ or $\sqrt{D}$ respectively, and in both cases, we see that the area equals $\sqrt{\Delta}$.
  2. If I is any nonzero ideal of R and $\Lambda_I = \varphi(I)$ is the image of I under the Minkowski embedding, then the fundamental domain for Λ_I has area equal to $N(I)$ times the fundamental domain for Λ.
     - Proof: Note that Λ_I is a sublattice (i.e., an additive subgroup) of $\Lambda = \varphi(\mathcal{O}_{\sqrt{D}})$.
     - Since φ is an isomorphism of additive abelian groups that maps $\mathcal{O}_{\sqrt{D}}$ to Λ and I to Λ_I, we see that $\Lambda/\Lambda_I \cong \mathcal{O}_{\sqrt{D}}/I$. Taking cardinalities then yields $\#(\Lambda/\Lambda_I) = \#(\mathcal{O}_{\sqrt{D}}/I) = N(I)$.
     - Geometrically, this means that the fundamental domain for Λ_I consists of $N(I)$ copies of the fundamental domain for Λ. Thus, the fundamental domain for Λ_I has area $N(I)$ times the area of the fundamental domain for Λ, as claimed.
  3. Every nonzero ideal I of R contains a nonzero element α with $|N(\alpha)| \leq \mu \cdot N(I)$, where $\mu = \begin{cases} \frac{1}{2}\sqrt{\Delta} & \text{if } D > 0 \\ \frac{2}{\pi}\sqrt{|\Delta|} & \text{if } D < 0 \end{cases}$.
     - Proof: Let $\Lambda_I = \varphi(I)$ be the image of I under the Minkowski embedding. By (1) and (2), the fundamental domain of Λ_I has area $\begin{cases} N(I) \cdot \sqrt{\Delta} & \text{if } D > 0 \\ N(I) \cdot \frac{1}{2}\sqrt{|\Delta|} & \text{if } D < 0 \end{cases}$.
     - First suppose $D > 0$ and let B be the convex, centrally-symmetric closed set in $\mathbb{R}^2$ defined by $|x_1| + |x_2| \leq N(I)^{1/2}\Delta^{1/4}\sqrt{2}$. It is simply a square of side length $2N(I)^{1/2}\Delta^{1/4}$ so its area is $4N(I)\sqrt{\Delta}$.
     - By Minkowski's theorem, since the area of B equals 2^2 times the area of the fundamental domain of Λ_I, there necessarily exists some nonzero element $\varphi(\alpha) = (\alpha, \overline{\alpha})$ of Λ_I in B.
     - Then $|N(\alpha)| = |\alpha|\,|\overline{\alpha}| \leq \left[\dfrac{|\alpha| + |\overline{\alpha}|}{2}\right]^2 \leq N(I) \cdot \dfrac{1}{2}\sqrt{\Delta}$ where we used the arithmetic-geometric mean inequality $xy \leq \left[\frac{x+y}{2}\right]^2$ which holds for any nonnegative x, y. This is the desired inequality.
     - Now suppose $D < 0$ and let B be the convex, centrally-symmetric closed set in $\mathbb{R}^2$ defined by $x_1^2 + x_2^2 \leq \dfrac{2}{\pi}N(I)\sqrt{|\Delta|}$, which is simply a circle of area $2N(I)\sqrt{|\Delta|}$.

- By Minkowski's theorem, since the area of B equals 2^2 times the area of the fundamental domain of Λ_I, there necessarily exists some nonzero element $\varphi(\alpha) = (\mathrm{Re}(\alpha), \mathrm{Im}(\alpha))$ of Λ_I in B.
- Then $N(\alpha) = \mathrm{Re}(\alpha)^2 + \mathrm{Im}(\alpha)^2$ is the sum of the squares of the coordinates of $\varphi(\alpha)$, which by the hypotheses on B is at most $\frac{2}{\pi}\sqrt{|\Delta|} \cdot N(I)$, as claimed.

4. Every ideal class of R contains an ideal J with $N(J) \leq \begin{cases} \frac{1}{2}\sqrt{\Delta} & \text{if } D > 0 \\ \frac{2}{\pi}\sqrt{|\Delta|} & \text{if } D < 0 \end{cases}$.
   - Proof: This follows identically as in our earlier proof, merely with the constant $1 + |D|$ replaced by the constant μ from (3) above.

- Minkowski's bound is quite a lot better than the estimate we obtained earlier, since it is asymptotic to $\sqrt{\Delta} \sim D^{1/2}$ rather than to D itself, so for large D we have far fewer ideals to examine in order to compute the class group.
- Example (again): Show that the class group of $\mathbb{Z}[\sqrt{-5}]$ has order 2.
  - Since $-5 \equiv 3 \pmod 4$, we have $\Delta = -20$, and so Minkowski's bound says that every ideal class of R contains an ideal of norm at most $\frac{2}{\pi}\sqrt{20} \approx 2.8471 < 3$, so the only nontrivial ideals we need to consider are ideals of norm 2.
  - Since (2) splits as $(2) = (2, 1 + \sqrt{-5})^2$, and we have previously shown that $(2, 1 + \sqrt{-5})$ is nonprincipal, we conclude that the class group is generated by the nonprincipal ideal $I_2 = (2, 1 + \sqrt{-5})$. Since I_2 has order 2 as $I_2^2 = (2)$, the class group has order 2 as claimed.
- Example: Show that the class group of $\mathcal{O}_{\sqrt{-19}}$ is trivial and deduce that it is a principal ideal domain.
  - Since $-19 \equiv 1 \pmod 4$, we have $\Delta = -19$, and so Minkowski's bound says that every ideal class of R contains an ideal of norm at most $\frac{2}{\pi}\sqrt{19} \approx 2.7750 < 3$, so the only nontrivial ideals we need to consider are ideals of norm 2.
  - The minimal polynomial of the generator is $x^2 - x + 5$, which is irreducible modulo 2. Therefore, (2) is inert, and so there are no ideals of norm 2 in $\mathcal{O}_{\sqrt{-19}}$.
  - Therefore, the only ideal class is the trivial class, so the class group is trivial and $\mathcal{O}_{\sqrt{-19}}$ is a PID.
  - Remark: It can be shown that $\mathcal{O}_{\sqrt{-19}}$ is not Euclidean with respect to any norm (though this is not quite so easy to do), so it provides an example of a PID that is not a Euclidean domain.
- Example: Determine the class group of $\mathbb{Z}[\sqrt{5}]$.
  - Since $5 \equiv 1 \pmod 4$, we have $\Delta = 5$, and so Minkowski's bound says that every ideal class of R contains an ideal of norm at most $\frac{1}{2}\sqrt{5} \approx 1.1180 < 2$, so there can be no nontrivial ideal classes.
  - Thus, the class group of $\mathbb{Z}[\sqrt{5}]$ is trivial.
- Example: Determine the class group of $\mathbb{Z}[\sqrt{6}]$.
  - Since $6 \equiv 2 \pmod 4$, we have $\Delta = 24$, and so Minkowski's bound says that every ideal class of R contains an ideal of norm at most $\frac{1}{2}\sqrt{24} \approx 2.4495 < 3$, so the only nontrivial ideals we need to consider are ideals of norm 2.
  - The minimal polynomial of the generator is $x^2 - 6$, which has a repeated root $r = 0$ modulo 2, so (2) is ramified: $(2) = (2, \sqrt{6})^2$. This ideal $I_2 = (2, \sqrt{6})$ is in fact principal as we showed earlier (it is generated by $2 + \sqrt{6}$).
  - Therefore, the only ideal class is the trivial class, so the class group is trivial.
- Example: Determine the class group of $\mathbb{Z}[\sqrt{10}]$.

- Since $10 \equiv 2 \pmod 4$, we have $\Delta = 40$, and so Minkowski's bound says that every ideal class of R contains an ideal of norm at most $\frac{1}{2}\sqrt{40} \approx 3.1623 < 4$, so the only nontrivial ideals we need to consider are ideals of norm 2 and norm 3.
- The minimal polynomial of the generator is $x^2 - 10$, which has a repeated root $r = 0$ modulo 2, so (2) is ramified: $(2) = (2, \sqrt{10})^2$. This ideal $I_2 = (2, \sqrt{10})$ is not principal, since any generator would necessarily have norm ± 2, but there are no elements of norm ± 2 since $x^2 - 10y^2 = \pm 2$ has no solutions modulo 5. Thus, $[I_2]$ is an element of order 2 in the class group since I_2 is not principal but I_2^2 is.
- For 3, since $x^2 - 10$ has roots ± 1 modulo 3, we see (3) splits: $(3) = (3, 1 + \sqrt{10})(3, 1 - \sqrt{10})$. The ideal $I_3 = (3, 1 + \sqrt{10})$ and its conjugate $I_3' = (3, 1 - \sqrt{10})$ are both nonprincipal, since any generator would necessarily have norm ± 3, but there are no elements of norm ± 3 since $x^2 - 10y^2 = \pm 3$ has no solutions modulo 5.
- We can then compute $I_3^2 = (9, 3 + 3\sqrt{10}, 11 + 2\sqrt{10})$. To test for principality we can look for elements of norm 9, and looking at such elements (e.g., $1 \pm \sqrt{10}$) will reveal this ideal is in fact principal and generated by $(1 + \sqrt{10})$. Explicitly, $1 + \sqrt{10} = 9 + (3 + 3\sqrt{10}) - (11 + 2\sqrt{10}) \in I_3^2$ and each generator is divisible by $1 + \sqrt{10}$. Then $(I_3')^2 = (1 - \sqrt{10})$, so $[I_3]$ and $[I_3']$ are both ideal classes of order 2 and they are equal.
- It remains to determine the relationship between I_2 and I_3. Indeed, $I_2 I_3 = (6, 2+2\sqrt{10}, 3\sqrt{10}, 10+\sqrt{10})$. To test for principality we can look for elements of norm 6, and looking at such elements (e.g., $4 \pm \sqrt{10}$) will reveal this ideal is in fact principal and generated by $(4 + \sqrt{10})$, since $4 + \sqrt{10} = (10 + \sqrt{10}) - 6$ and each generator is divisible by $4 + \sqrt{10}$. Thus since $[I_2][I_3] = (1) = [I_2]^2$, we see $[I_2] = [I_3]$.
- Thus, we conclude that there is one nonprincipal ideal class of order 2, so the class group is isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

- Example: Determine the class group of $\mathcal{O}_{\sqrt{-31}}$.
  - Since $-31 \equiv 1 \pmod 4$, we have $\Delta = -31$, and so Minkowski's bound says that every ideal class of R contains an ideal of norm at most $\frac{2}{\pi}\sqrt{31} \approx 3.5445 < 4$, so the only nontrivial ideals we need to consider are ideals of norm 2 and 3.
  - The minimal polynomial of the generator $\omega = \frac{1+\sqrt{-31}}{2}$ is $x^2 - x + 8$.
  - For (2) we see the polynomial has roots 0 and 1 so we get $(2) = (2, \frac{1+\sqrt{-31}}{2})(2, \frac{1-\sqrt{-31}}{2})$. If the ideal $I_2 = (2, \frac{1+\sqrt{-31}}{2})$ or its conjugate $I_2' = (2, \frac{1-\sqrt{-31}}{2})$ were principal then it would be generated by an element of norm 2, but there are no elements of norm 2 since this would require $N(\frac{x+y\sqrt{-31}}{2}) = 2$, but there are no solutions to $x^2 + 31y^2 = 8$.
  - The ideal I_2^2 cannot be principal either, since it would have to be generated by an element of norm 4, but the only such elements are ± 2 and we already have the ideal factorization $(2) = I_2 I_2'$ and $I_2 \neq I_2'$ since 2 is not ramified.
  - On the other hand, I_2^3 has norm 8, and there are elements of norm 8, namely, $\frac{1 \pm \sqrt{-31}}{2}$. Indeed, we can see that $I_2^3 = (8, 2 + 2\sqrt{-31}, -15 + \sqrt{-31}, \frac{-23-7\sqrt{-31}}{2})$ so this ideal contains $8 + 2(2 + 2\sqrt{-31}) + \frac{-23-7\sqrt{-31}}{2} = \frac{1+\sqrt{-31}}{2}$. Thus $I_2^3 = (\frac{1+\sqrt{-31}}{2})$ is principal, and so $[I_2]$ is an element of order 3 in the class group with inverse $[I_2'] = [I_2]^2$.
  - For (3) we see the polynomial $x^2 - x + 8$ is irreducible modulo 3, so (3) is inert of norm 9 and it does not yield a nontrivial element of the class group.
  - Therefore, the class group is generated by $[I_2]$ and is isomorphic to $\mathbb{Z}/3\mathbb{Z}$.

9.3 Binary Quadratic Forms

- We will now discuss representations of integers by binary quadratic forms, which are expressions of the form $f(x, y) = ax^2 + bxy + cy^2$ for fixed integers a, b, c not all zero.
  - We have already classified the integers that are represented by the forms $x^2 + y^2$, $x^2 + 2y^2$, $x^2 + xy + y^2$, and $x^2 + 3y^2$ using unique factorization in the quadratic integer rings $\mathbb{Z}[i]$, $\mathbb{Z}[\sqrt{-2}]$, and $\mathcal{O}_{\sqrt{-3}}$.
  - Our goal is now to broaden our focus and analyze integers represented by other binary quadratic forms.

9.3.1 Representations of Integers by Binary Quadratic Forms

- Definitions: The discriminant of the binary quadratic form $f(x,y) = ax^2+bxy+cy^2$ is $\Delta = b^2-4ac$. If f takes both positive and negative values on $\mathbb{R}$ then we say f is indefinite. If f takes only nonnegative values we say f is positive semidefinite, and if in addition $f = 0$ only when $(x,y) = (0,0)$ we say f is positive definite. Finally, we say f is negative semidefinite (respectively, negative definite) if $-f$ is positive semidefinite (respectively, positive definite).
    - From elementary algebra, we know that the discriminant characterizes the behavior of the roots of f: if $D > 0$ then f has two real roots (and they are rational whenever Δ is a perfect square), if $\Delta = 0$ then f has a repeated (rational) root, and if $D < 0$ then f has no real roots.
    - Thus, f is indefinite precisely when $\Delta > 0$, f is definite precisely when $\Delta < 0$ (it is positive definite for $a > 0$ and negative definite for $a < 0$), and f is semidefinite but not definite when $\Delta = 0$ (it is positive semidefinite when $a + c > 0$ and negative semidefinite when $a + c < 0$).
    - Example: The forms $x^2 - y^2$ ($\Delta = 4$), xy ($\Delta = 1$), and $x^2 - 5xy + y^2$ ($\Delta = 21$) are all indefinite.
    - Example: The forms x^2+y^2 ($\Delta = -4$), $x^2+2xy+3y^2$ ($\Delta = -8$), x^2 ($\Delta = 0$), and $x^2+2xy+y^2$ ($\Delta = 0$) are all positive semidefinite. The first two are positive definite while the last two are not.
    - Example: The forms $-x^2 + 2xy - 2y^2$ ($\Delta = -4$) and $-4x^2 - 6xy - 9y^2$ ($\Delta = 0$) are both negative semidefinite. The first is negative definite while the second is not.
- We will observe that the discriminant Δ of any quadratic form is always congruent to 0 or 1 modulo 4, so it is always the discriminant of a quadratic integer ring up to a square factor.
    - Conversely, if Δ is 0 or 1 modulo 4 and is squarefree up to a factor of 4, then the norm $N(x+y\omega)$ where ω is the generator of the quadratic integer ring $\mathcal{O}_{\sqrt{D}}$ ($\sqrt{D}$ or $\frac{1+\sqrt{D}}{2}$) of the field $\mathbb{Q}(\sqrt{\Delta})$ gives a quadratic form of discriminant Δ. In this case, Δ is simply the discriminant of the ring $\mathcal{O}_{\sqrt{D}}$ itself.
- Now we can discuss representations of integers by quadratic forms:
- Definition: If f is a binary quadratic form and n is an integer, we say f represents n if there exist integers x and y such that $f(x,y) = n$, and we say f properly represents n if these x, y are also relatively prime.
    - Example: $f = x^2 + y^2$ represents 2, 9, and 13, but it does not properly represent 9 because there is no solution to $x^2 + y^2 = 9$ with x, y relatively prime.
    - Example: $f = x^2 + xy + y^2$ represents 3, 4, and 7, but it does not properly represent 4 because there is no solution to $x^2 + xy + y^2 = 4$ with x, y relatively prime.
- Ultimately, we would like to be able to classify the integers represented (or properly represented) by a given quadratic form. This turns out to be quite difficult, but we can establish some more basic results.
- Proposition (Representations by Forms of Discriminant Δ): Suppose Δ is a nonzero integer congruent to 0 or 1 modulo 4.
    1. If n is a nonzero integer, then there exists a binary quadratic form of discriminant Δ that properly represents n if and only if D is a quadratic residue modulo $4n$.
        - Proof: First suppose that Δ is a quadratic residue modulo $4n$, say with $\Delta \equiv b^2 \pmod{4n}$, so that $b^2 - \Delta = 4nc$ for some integer c. Then the quadratic form $f(x,y) = nx^2+bxy+cy^2$ has discriminant $b^2 - 4nc = \Delta$ and it properly represents n since $f(1,0) = n$.
        - Conversely, suppose $ax^2+bxy+cy^2 = n$ with x, y relatively prime and with $b^2-4ac = \Delta$. Multiplying by $4a$ and completing the square gives $4an = 4a^2x^2 + 4abxy + 4acy^2 = (2ax + by)^2 + (b^2 - 4ac)y^2$ so that $(b^2 - 4ac)y^2 \equiv (2ax + by)^2 \pmod{4n}$.
        - Therefore, $b^2 - 4ac$ is a quadratic residue modulo $4n/\gcd(y, 4n)$. By a symmetric argument, we see $b^2 - 4ac$ is also a quadratic residue modulo $4n/\gcd(x, 4n)$, and since x, y are relatively prime, this means $b^2 - 4ac$ is a quadratic residue modulo $4n$, as required.
    2. If p is an odd prime, then there exists a binary quadratic form of discriminant Δ that represents p if and only if Δ is a quadratic residue modulo p.

- Proof: Since p is squarefree, any representation of p must automatically be proper.
- Then by (1), we see that p is represented by a form of discriminant Δ if and only if Δ is a quadratic residue modulo $4p$.
- However, since p is odd and Δ is 0 or 1 modulo 4 (hence is a square modulo 4), by the Chinese remainder theorem this is equivalent to saying that Δ is a quadratic residue modulo p.

9.3.2 Equivalence of Binary Quadratic Forms

- The results above give an easy way to decide whether there is *some* quadratic form of discriminant Δ that represents a given prime p.
  - It therefore stands to reason that if we can understand the structure of the quadratic forms of a given discriminant Δ, then we might be able to determine whether p is represented by a *particular* quadratic form of discriminant Δ.
  - To begin, we can see that there are simple changes of variable we can perform that do not affect representability: for example, the binary quadratic forms $f(x,y) = x^2 + y^2$ and $g(x,y) = f(x-y,y) = x^2 - 2xy + 2y^2$ represent the same integers, since the pair $(x,y) \in \mathbb{Z}^2$ if and only if $(x-y,y) \in \mathbb{Z}^2$.
  - On the other hand, the binary quadratic forms $f(x,y) = x^2 + y^2$ and $g(x,y) = f(2x,y) = 4x^2 + y^2$ do not represent the same integers (e.g., f represents 2 while g does not).
  - More generally, we may identify any two quadratic forms that are obtained via a linear change of variables from one another, as long as the change of variables is invertible over $\mathbb{Z}$.
- The easiest way to keep track of such changes of basis is to use matrices:
- Definition: If $f(x,y) = ax^2 + bxy + cy^2$ is a binary quadratic form, its associated matrix is the symmetric matrix $M_f = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$.
  - The connection is that the quadratic form $f(x,y)$ is equal to $\mathbf{x}^T M_f \mathbf{x}$ where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ is the column vector of the two variables. We also observe that $\det(M_f) = ac - b^2/4 = -\Delta/4$.
- It is then easy to write down how a binary quadratic form f transforms under a change of coordinates $\mathbf{x} \mapsto A\mathbf{x}$: explicitly, we have $f(A\mathbf{x}) = (A\mathbf{x})^T M_f (A\mathbf{x}) = \mathbf{x}^T [A^T M_f A]\mathbf{x}$, and so the associated matrix of the new form is $A^T M_f A$.
  - For the purposes of representations of integers, we want only to consider changes of variables $\mathbf{x} \mapsto A\mathbf{x}$ that are a bijection from $\mathbb{Z}^2$ to itself, since this ensures that the possible input vectors $\mathbf{x}$ are the same for both forms. It is easy to see that this is equivalent to saying that A is an invertible matrix with integer entries whose inverse also has integer entries.
  - These conditions imply that $\det(A^{-1}) = 1/\det(A) \in \mathbb{Z}$, so A must have determinant ± 1. Conversely, by the adjugate inverse formula $A^{-1} = \frac{1}{\det(A)} A^\dagger$, which for 2×2 matrices reads as $\begin{bmatrix} e & f \\ g & h \end{bmatrix}^{-1} = \frac{1}{eh - fg} \begin{bmatrix} h & -f \\ -g & e \end{bmatrix}$, the condition of having integer entries and determinant ± 1 is sufficient for $\mathbf{x} \mapsto A\mathbf{x}$ to be a bijection from $\mathbb{Z}^2$ to itself.
- For various reasons (primarily, that the resulting theory is much nicer), we will restrict our attention to changes of coordinates with determinant $+1$ only, which yields the matrix group $SL_2(\mathbb{Z}) = \{M \in GL_2(\mathbb{Z}) : \det(M) = 1\}$.
  - From our discussion above, for any $A \in SL_2(\mathbb{Z})$, we see that the integers represented by the forms $f(\mathbf{x})$ and $f(A\mathbf{x})$ will be the same, as will the integers properly represented by these two forms.

- Definition: We define the relation $\sim$ on binary quadratic forms by writing $f \sim g$ if there exists a matrix $A \in SL_2(\mathbb{Z})$ such that $g(\mathbf{x}) = f(A\mathbf{x})$, which is to say that g is obtained from f by an invertible orientation-preserving linear change of variables with integer coefficients. Equivalently, $f \sim g$ if there exists $A \in SL_2(\mathbb{Z})$ such that $M_g = A^T M_f A$.
  - Example: The quadratic forms $f(x,y) = x^2 + y^2$ and $g(x,y) = f(x-y,y) = x^2 - 2xy + 2y^2$ have $f \sim g$, since we can take the matrix $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \in SL_2(\mathbb{Z})$. For the matrix calculation, we have $M_f = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $M_g = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ and indeed we have $M_g = A^T M_f A$.
  - Example: The quadratic forms $f(x,y) = x^2 + 2xy - y^2$ and $g(x,y) = 7x^2 + 22xy + 17y^2$ have $f \sim g$, since we can take the matrix $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \in SL_2(\mathbb{Z})$; one may check that $g(x,y) = f(2x+3y, x+2y)$.
  - It is not hard to see that $\sim$ is an equivalence relation:
    1. To see $f \sim f$, simply take $A = 1$.
    2. If $f \sim g$ then $M_g = A^T M_f A$ and so $(A^{-1})^T M_g (A^{-1}) = M_f$ so that $g \sim f$.
    3. If $f \sim g$ and $g \sim h$ let $M_g = A^T M_f A$ and $M_h = B^T M_g B$: then $M_h = (AB)^T M_f (AB)$ so $f \sim h$.
  - Also, by taking determinants, if $f \sim g$ then $\det(M_g) = \det(A^T)\det(M_f)\det(A) = \det(M_f)$ and so forms in the same equivalence class under $\sim$ will have the same discriminant.

- Since we are interested in representability of integers by quadratic forms, and representability is the same for different forms in the same equivalence class under $\sim$, our next task is to identify nice representatives for the equivalence classes under $\sim$.

- Definition: If $f(x,y) = ax^2 + bxy + cy^2$ is a binary quadratic form whose discriminant Δ is not a square, we say f is reduced when $|b| \le |a| \le |c|$, and if $|b| = |a|$ or $|a| = |c|$ we also insist that $b \ge 0$.
  - Example: The forms $x^2 + y^2$, $x^2 - y^2$, $-3x^2 + 3xy + 4y^2$, and $2x^2 + xy + 3y^2$ are all reduced.
  - Example: The forms $x^2 + 2xy$, $xy - 2y^2$, $2x^2 + 2xy + y^2$, $2x^2 - 2xy + 3y^2$, and $3x^2 - 2xy + 3y^2$ are not reduced.

- Theorem (Reduced Forms): Let Δ be a nonsquare integer congruent to 0 or 1 modulo 4 and suppose $f(x,y) = ax^2 + bxy + cy^2$ is a reduced form of discriminant Δ. Then the following hold:
  1. If $D < 0$ then a, c must have the same sign and $|a| \le \sqrt{-\Delta/3}$. If $\Delta > 0$ then a, c have opposite signs and $|a| \le \sqrt{\Delta}/2$. In either case, there are finitely many reduced forms of discriminant Δ.
     - Proof: If a, c have the same sign then $\Delta = b^2 - 4ac = |b|^2 - 4|a||c| \le 0$, while if a, c have opposite signs then $\Delta = b^2 - 4ac = b^2 + 4|ac| \ge 0$. Since $\Delta \ne 0$ we see that Δ has the same sign as $-ac$.
     - If $\Delta < 0$ then because $|a| \le |c|$ we see that $\Delta = b^2 - 4|a||c| \le a^2 - 4a^2 = -3a^2$ so $|a| \le \sqrt{-\Delta/3}$. If $\Delta > 0$ then again because $|a| \le |c|$ we see that $\Delta = b^2 + 4|ac| \ge 4a^2$ and so $|a| \le \sqrt{\Delta}/2$.
     - In either case there are finitely many values of a. For each of these values of a, there are only finitely many possible b since $|b| \le |a|$, and then $c = (b^2 - \Delta)/(4a)$ is determined. Thus, there are only finitely many reduced forms of discriminant Δ.
  2. Every equivalence class of quadratic forms of discriminant Δ contains at least one reduced form.
     - As motivation, note that $SL_2(\mathbb{Z})$ is generated[4] by the matrices $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.
     - The idea is then to show that if we have a non-reduced form, we must be able to apply either S or a power of T to obtain a "smaller" form, and so iterating this procedure must eventually yield a reduced form.

[4] We will not actually use this fact, but one may prove it by applying the Euclidean algorithm to the first column of a matrix M of $SL_2(\mathbb{Z})$: note that $T^{-q}M$ will subtract q times the second row from the first row, and SM will swap the rows and negate the first one: these (up to the scaling by -1) are precisely the operations in the Euclidean algorithm. Applying it will turn the first column into $[1\,0]^T$, and then the second column must be $[m\,1]^T$ since it is in $SL_2(\mathbb{Z})$, and this is just T^m. This gives a procedure for writing any matrix in $SL_2(\mathbb{Z})$ in terms of S and T.

- Proof: Suppose $f(x,y) = ax^2 + bxy + cy^2$ has discriminant Δ and has associated matrix M_f.
- Then $S^T M_f S = \begin{bmatrix} c & -b/2 \\ -b/2 & a \end{bmatrix}$ corresponds to the form $cx^2 - bxy + ay^2$, which swaps the x^2- and y^2-coefficients while leaving the absolute value of the xy-coefficient unchanged.
- Also, $(T^m)^T M_f T^m = \begin{bmatrix} a & a + mb/2 \\ a + mb/2 & m^2 a + mb + c \end{bmatrix}$ corresponds to the form $ax^2 + (b + 2am)xy + (am^2 + bm + c)y^2$, which leaves the x^2-coefficient unchanged and shifts the xy-coefficient by $2am$.
- Thus, starting with f, we may apply the following steps:
  (a) If b is not in the interval $(-|a|, |a|]$, let m be the unique integer such that $b + 2am \in (-|a|, |a|]$ and apply T^m to the quadratic form. This yields an equivalent form whose xy-coefficient is $b + 2am \in (-|a|, |a|]$ and whose x^2-coefficient is the same. Then go to step (b).
  (b) If b is in the interval $(-|a|, |a|]$, test if $|a| = |c|$. If so and if $b \geq 0$, the form is reduced; otherwise if $b < 0$ then applying S will yield a reduced form. Otherwise, test whether $|a| < |c|$. If so, the form is reduced, and if not, apply S to the quadratic form. This yields an equivalent form whose x^2-coefficient is smaller in absolute value, and return to step (a).
- After applying the steps once, the form is either reduced or has an x^2-coefficient that is strictly smaller in absolute value, and so iterating the procedure must eventually yield a reduced form. Since each application of S or T yields an equivalent form, we conclude that every equivalence class contains at least one reduced form.

3. There are finitely many equivalence classes of binary quadratic forms of discriminant Δ.
   - Proof: Each equivalence class contains at least one reduced form by (2), and there are finitely many reduced forms by (1).

- We may apply the algorithm in (2) to find equivalent reduced forms:
- Example: Find a reduced form equivalent to $f(x,y) = 17x^2 + 99xy - 46y^2$.
  - First, since $b \notin (-17, 17]$ we find m with $b + 34m \in (-17, 17]$, which gives $m = -3$.
  - Applying T^3 yields the equivalent form $g(x,y) = f(x - 3y, y) = 17x^2 - 3xy - 190y^2$.
  - Now because $|a| < |c|$, the resulting form $\boxed{17x^2 - 3xy - 190y^2}$ is reduced.
- Example: Find a reduced form equivalent to $f(x,y) = 119x^2 - 145xy + 17y^2$.
  - First, since $b \notin (-119, 119]$ we find m with $b + 238m \in (-119, 119]$, which gives $m = 1$.
  - Applying T yields the equivalent form $g(x,y) = f(x + y, y) = 119x^2 + 93xy - 9y^2$.
  - Now because $|a| > |c|$ the form is not reduced so we apply S to get the form $h(x,y) = f(-y, x) = -9x^2 - 93xy + 119y^2$.
  - Then since $b \notin (-9, 9]$ we find m with $b + 18m \in (-9, 9]$, which gives $m = 5$.
  - Applying T^5 yields the equivalent form $i(x,y) = f(x + 5y, y) = -9x^2 - 3xy + 359y^2$. Since $|a| < |c|$, this form $\boxed{-9x^2 - 3xy + 359y^2}$ is reduced.
- Example: Find a reduced form equivalent to $f(x,y) = 81x^2 - 65xy + 13y^2$.
  - First apply S, yielding $13x^2 + 65xy + 81y^2$. Then apply T^{-2}, yielding $13x^2 + 13xy + 3y^2$.
  - Next apply S, yielding $3y^2 - 13xy + 13y^2$. Then apply T^2, yielding $3y^2 - xy - y^2$.
  - Finally, applying S yields $\boxed{-x^2 + xy + 3y^2}$, which is reduced.
- For small values of Δ we can also use the partial description of reduced forms in (2) from the theorem to make a full list of reduced forms.
  - By deciding which of these are equivalent to one another, we can then determine the precise number of equivalence classes of forms.

- Example: Find all reduced forms of discriminant $\Delta = -4$ and show that there is only one equivalence class of positive-definite forms.
  - From the analysis in (2) we see that any reduced form $ax^2 + bxy + cy^2$ of discriminant $\Delta = -4$ must have $|a| \leq \sqrt{4/3}$, so since $a \neq 0$ this means $a = \pm 1$. Then since $|b| \leq |a|$ we have $b = 0, \pm 1$.
  - Also, since $c = (b^2 - \Delta)/(4a)$ must be an integer, b must be even. Thus the only possible forms have $a = \pm 1$ and $b = 0$, which yields the two forms $x^2 + y^2$ and its negative $-x^2 - y^2$.
  - Therefore, since only x^2+y^2 is positive-definite, it represents the only equivalence class of positive-definite forms.
- Example: Find all reduced forms of discriminant $\Delta = 13$ and determine the number of equivalence classes.
  - From the analysis in (2) we see that any reduced form $ax^2 + bxy + cy^2$ of discriminant $\Delta = 5$ must have $|a| \leq \sqrt{13}/2 < 2$, so since $a \neq 0$ this means $a = \pm 1$. Then since $b \in (-|a|, a]$ we must have $b = 0$ or $b = 1$.
  - However, since $c = (b^2 - \Delta)/(4a)$ must be an integer, b must be odd, and so $b = 1$. We then get two possible forms for $a = -1$ and $a = 1$ respectively: $f(x, y) = -x^2 + xy + 3y^2$ and $g(x, y) = x^2 + xy - 3y^2$.
  - Although both of these forms are reduced, they are in fact equivalent: if we take the matrix $A = \begin{bmatrix} 2 & -3 \\ 1 & -1 \end{bmatrix}$ then it is straightforward to check that $A^T M_f A = M_g$, and so $f \sim g$.
  - Therefore, there is only one equivalence class of forms of discriminant $\Delta = 13$.
- Example: Find all reduced forms of discriminant $\Delta = -40$ and determine the number of equivalence classes.
  - From the analysis in (2) we see that any reduced form $ax^2+bxy+cy^2$ of discriminant $\Delta = -40$ must have $|a| \leq \sqrt{40/3} < 4$, so since $a \neq 0$ this means $a = \pm 1, \pm 2, \pm 3$. Then since $b \in (-|a|, a]$ and $b^2 - 4ac = -40$ so that b is even, we must have $b = 0, \pm 2$.
  - If $a = 1$ then $b = 0$ and then $c = (b^2 + 40)/(4a) = 10$, and if $a = -1$ then $b = 0$ and $c = -10$.
  - If $a = \pm 2$ then $b = 0, 2$ so that $c = (b^2 + 40)/(4a) = \pm 5$ or $\pm 44/8$, but the second case yields non-integral c.
  - If $a = \pm 3$ then $b = 0, \pm 2$ so that $c = (b^2 + 40)/(4a) = \pm 40/12$ or $\pm 44/12$ but these are not integral either.
  - So in summary, we obtain two positive-definite forms $x^2 + 10y^2$ and $2x^2 + 5y^2$ along with their negatives (which are negative-definite) $-x^2 - 10y^2$ and $-2x^2 - 5y^2$.
  - The positive-definite and negative-definite forms are not equivalent to one another, and so we only have to consider equivalence of the two positive-definite forms. But they are not equivalent because they do not represent the same numbers: for example, $2x^2 + 5y^2$ represents both 2 and 5, while $x^2 + 10y^2$ does not.
  - Therefore, all these reduced forms are inequivalent, and so there are four inequivalent forms in total.
- Example: Find all reduced forms of discriminant $\Delta = -31$ and determine the number of equivalence classes.
  - From the analysis in (2) we see that any reduced form $ax^2+bxy+cy^2$ of discriminant $\Delta = -31$ must have $|a| \leq \sqrt{31/3} < 4$, so since $a \neq 0$ this means $a = \pm 1, \pm 2, \pm 3$. Then since $b \in (-|a|, a]$ and $b^2 - 4ac = -31$ so that b is odd, we must have $b = \pm 1, \pm 3$.
  - If $a = 1$ then $b = 1$ and then $c = (b^2 + 31)/(4a) = 8$, and if $a = -1$ then $b = 1$ and $c = -8$.
  - If $a = 2$ then $b = \pm 1$ and then $c = (b^2 + 31)/(4a) = 4$, and if $a = -2$ then $b = \pm 1$ and $c = -4$.
  - Finally, if $a = \pm 3$ then $b = \pm 1$ or 3 and then $c = (b^2 + 31)/(4a)$ is either $32/\pm 6$ or $40/\pm 6$, but none of these are integers.
  - So in summary, we obtain three positive-definite forms $x^2+xy+8y^2$, $2x^2+xy+4y^2$, $2x^2-xy+4y^2$ along with their negatives (which are negative-definite): $-x^2-xy-8y^2$, $-2x^2-xy-4y^2$, and $-2x^2+xy-4y^2$.
  - The positive-definite and negative-definite forms are not equivalent to one another, and so we only have to consider equivalence of the positive-definite forms.

- It is not hard to see that $x^2+xy+8y^2$ is not equivalent to either of the others, since it does not represent 2 whereas the other two do. The other two forms $2x^2+xy+4y^2$, $2x^2-xy+4y^2$ are also inequivalent, although this is harder to show. (In particular, we cannot use the approach from the last example: these two forms represent the same integers because they are obtained via a change of variables $(x,y)\mapsto(x,-y)$ of determinant -1.)
- We can show the inequivalence using associated matrices: so suppose we had $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\in SL_2(\mathbb{Z})$ such that $A^T\begin{bmatrix} 2 & 1/2 \\ 1/2 & 4 \end{bmatrix}A=\begin{bmatrix} 2 & -1/2 \\ -1/2 & 4 \end{bmatrix}$.
- Since A has determinant 1, we have $A^{-1}=\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ and so we equivalently must solve $A^T\begin{bmatrix} 2 & 1/2 \\ 1/2 & 4 \end{bmatrix}=\begin{bmatrix} 2 & -1/2 \\ -1/2 & 4 \end{bmatrix}A^{-1}$, so that $\begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} 2 & 1/2 \\ 1/2 & 4 \end{bmatrix}=\begin{bmatrix} 2 & -1/2 \\ -1/2 & 4 \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
- This yields an explicit linear system that reduces to $2a-2d=a+2b+4c=2b+4c+d=-4a+4d=0$, so that $a=d=-2b-4c$. But then $ad-bc=(2b+4c)^2-bc=4b^2+15bc+16c^2$ cannot equal 1, as can be seen by completing the square: $(2b+15c/4)^2+(31/16)c^2=1$ implies $(8b+15c)^2+31c^2=16$ and this has no integer solutions.
- Therefore, all three of these reduced forms are inequivalent, and so there are six inequivalent forms in total.

- In the examples above, we identified reduced forms that were equivalent when $\Delta>0$, but all of our examples of reduced forms with $\Delta<0$ were inequivalent. In fact, reduced forms of negative discriminant always yield distinct equivalence classes:

- Proposition (Inequivalence of Reduced Forms for $\Delta<0$): Suppose $\Delta<0$ is the discriminant of a quadratic integer ring.

  1. If $f(x,y)=ax^2+bxy+cy^2$ is a reduced positive-definite form of discriminant Δ, then a, c, and $a-|b|+c$ are the smallest nonzero integers properly represented by f.
     - Proof: Since the definiteness of a form is preserved by equivalence, without loss of generality we may work only with the positive-definite forms.
     - Now suppose $f(x,y)=ax^2+bxy+cy^2$ is reduced, so that $|b|\le a\le c$. If $x^2\ge y^2$ then $f(x,y)\ge ax^2-|b|\,xy+cy^2\ge(a-|b|+c)y^2$, and similarly if $y^2\ge x^2$ then $f(x,y)\ge(a-|b|+c)x^2$. Therefore, $f(x,y)\ge(a-|b|+c)\min(x^2,y^2)$.
     - We have $f(x,0)=ax^2$ and $f(0,y)=cy^2$, so the only integers with $xy=0$ properly represented by f are $f(\pm1,0)=a$ and $f(0,\pm1)=c$. Otherwise, $f(x,y)\ge a-|b|+c$, and since $f(1,1)=f(-1,-1)=a+b+c$ and $f(1,-1)=f(-1,1)=a-b+c$, the value $a-|b|+c$ is also properly represented by f.
     - Any other value represented by f necessarily has $\min(x^2,y^2)\ge4$ and is larger than these three values.
     - Therefore, the smallest nonzero integers properly represented by f are a, c, and $a-|b|+c$.
  2. If f and g are reduced positive-definite forms of discriminant Δ and $f\sim g$, then in fact $f=g$.
     - Proof: Suppose $f(x,y)=ax^2+bxy+cy^2$ and $g(x,y)=a'x^2+b'xy+c'y^2$ are reduced positive-definite forms of discriminant Δ.
     - Since the forms are reduced, we have $a\le c\le a-|b|+c$ and also $a'\le c'\le a'-|b'|+c'$.
     - Since $f\sim g$, as we have shown, the integers properly represented by g are the same as those properly represented by f. Therefore, by (1), we must have $a=a'$, $c=c'$, and $|b|=|b'|$, so $b=\pm b'$.
     - In the case where $a=c$ or where $|b|=a$, since both forms are reduced we must also have b and $b'\ge0$, so $b=b'$.
     - Otherwise, suppose $|b|<a<c$ and that $g(x,y)=f(px+qy,rx+sy)$ where $ps-qr=1$. Then $a=g(1,0)=f(p,q)$ and $c=g(0,1)=f(r,s)$ are proper representations of a and c respectively, and it is easy to see that f properly represents a only at $(x,y)=(\pm1,0)$ and f properly represents c only at $(x,y)=(0,\pm1)$.
     - This forces $(p,q)=(\pm1,0)$ and $(r,s)=(0,\pm1)$, and then the determinant condition requires $(p,q,r,s)=(1,0,0,1)$ or $(-1,0,0,-1)$, and in both cases this yields $g(x,y)=f(x,y)$.

3. Every equivalence class of binary quadratic forms of discriminant Δ is represented by a unique reduced form.
    - Proof: For positive-definite forms this follows immediately from (2), since as we showed earlier, every equivalence class contains at least one reduced form.
    - For negative-definite forms we can simply scale everything by -1 and note that equivalence preserves the definiteness type of a form.

- If we can classify all of the binary quadratic forms of a given discriminant, we can often identify exactly which primes may be represented by a given form.
    - As we proved earlier, a prime p is (properly) represented by a form of discriminant Δ if and only if Δ is a square modulo p.
    - If the equivalence classes of the forms of discriminant Δ are represented by $f_1, f_2, \ldots, f_k$, then (at least) one of the f_i represents p if and only if Δ is a square modulo p.
    - For $\Delta = -4$, for instance, this result says that p is represented by the unique equivalence class representative $f(x, y) = x^2 + y^2$ if and only if -4 is a square modulo p, which is in turn equivalent to saying that -1 is a square modulo p, which (as we have already noted numerous times) is equivalent to saying that $p \equiv 1 \pmod 4$.
    - For $\Delta = 13$, we see that p is represented by the unique equivalence class representative $f(x, y) = x^2 + xy - 3y^2$ if and only if 13 is a square modulo p, which (by quadratic reciprocity) is equivalent to saying that p is a quadratic residue modulo 13, which is to say, when $p = 13$ or when $p \equiv 1, 3, 4, 9, 10, 12 \pmod{13}$.

9.3.3 Composition of Binary Quadratic Forms

- We now investigate the composition of binary quadratic forms, which we will motivate first via some examples.
    - As we have already seen during our discussion of solutions to Pell's equation, the product $(a+b\sqrt{D})(c+d\sqrt{D}) = (ac + Dbd) + (ad + bc)\sqrt{D}$ is also an element of $\mathbb{Z}[\sqrt{D}]$.
    - Taking norms on both sides yields $(a^2 + Db^2)(c^2 + Dd^2) = (ac + Dbd)^2 + D(ad + bc)^2$, which shows that the product of two integers represented by the quadratic form $x^2 + Dy^2$ is also represented by this same quadratic form.
    - More generally, if $D \equiv 1 \pmod 4$, because the quadratic integer ring $\mathcal{O}_{\sqrt{D}}$ is a ring, the product of two elements $a + b\omega$ and $c + d\omega$ is again an element of this ring, and so the product of two elements represented by the norm form $x^2 + xy + \frac{1-D}{4}y^2$ is also represented by that form.
    - For some values of Δ, we have found several inequivalent forms of discriminant Δ, only one of which necessarily corresponds to a norm form $x^2 + Dy^2$ or $x^2 + xy + \frac{1-D}{4}y^2$.

- For example, for $D = -10$ corresponding to $\Delta = -40$, we identified two inequivalent positive-definite forms $x^2 + 10y^2$ and $2x^2 + 5y^2$.
    - Some small values represented by $x^2 + 10y^2$ are 0, 1, 4, 9, 10, 11, 14, 16, 19, 25, 26, 35, ... , while some small values represented by $2x^2 + 5y^2$ are 0, 2, 5, 7, 8, 13, 18, 20, 22, 23, 28, 32, ...
    - Aside from 0, these lists are disjoint. The first list is closed under multiplication (as we showed above), but the second visibly is not: indeed, 2, 5, and 7 are all on the second list, but their pairwise products 10, 14, and 35 actually all appear on the first list.
    - In fact, this holds for any product of two elements from the second list: we can see that if we multiply out $(2a^2 + 5b^2)(2c^2 + 5d^2) = 4a^2c^2 + 10(a^2d^2 + b^2c^2) + 25b^2d^2$, the result is of the form $x^2 + 10y^2$ for $x = 2ac + 5bd$ and $y = bc - ad$.
    - If we multiply an element on the first list by an element on the second list, we seem always to obtain something on the second list: for example, $10 \cdot 2 = 20$, $14 \cdot 2 = 28$, $4 \cdot 5 = 20$, and so forth.
    - Like above, this holds in general: if we multiply out $(2a^2+5b^2)(c^2+10d^2) = 2a^2c^2+5b^2c^2+20a^2d^2+50b^2d^2$, the result is of the form $2x^2 + 5y^2$ for $x = ac + 5bd$ and $y = bc - 2ad$.

- All of this together shows that the equivalence classes of positive-definite quadratic forms of discriminant $\Delta = -40$ have a group structure isomorphic to $\mathbb{Z}/2\mathbb{Z}$ under multiplication, with the form $x^2 + 10y^2$ as the identity and the form $2x^2 + 5y^2$ as the nontrivial element in the group.

- We can find similar patterns with the quadratic forms for other discriminants, although in some cases they are harder to identify.

  - For example, for discriminant $\Delta = -84$, one may show that there are four positive-definite reduced forms: $x^2 + 21y^2$, $2x^2 + 2xy + 11y^2$, $3x^2 + 7y^2$, and $5x^2 + 4xy + 5y^2$. Here are the integers less than 100 represented by each form:

	Form	Integers
e	$x^2 + 21y^2$	0, 1, 4, 9, 16, 21, 22, 25, 30, 36, 37, 46, 49, 57, 64, 70, 81, 84, 85, 88, 93
a	$2x^2 + 2xy + 11y^2$	0, 2, 8, 11, 15, 18, 23, 32, 35, 42, 44, 50, 51, 60, 71, 72, 74, 92, 95, 98, 99
b	$3x^2 + 7y^2$	0, 3, 7, 10, 12, 19, 27, 28, 31, 34, 40, 48, 55, 63, 66, 75, 76, 82, 90
c	$5x^2 + 4xy + 5y^2$	0, 5, 6, 14, 17, 20, 24, 33, 38, 41, 45, 54, 56, 62, 68, 69, 77, 80, 89, 96

  - If we hypothesize that the reduced forms up to equivalence form a group under multiplication, then since $x^2 + 21y^2$ is the norm form on $\mathbb{Z}[\sqrt{-21}]$, its set of represented integers is closed under multiplication, so it corresponds to the identity element.
  - If we label the other three classes as a, b, and c, then it is not hard to verify that $a \cdot a = e$ for small entries in the table (e.g., $8 \cdot 11 = 88$, $2 \cdot 23 = 46$, etc.), and also $b \cdot b = e$ (e.g., $3 \cdot 7 = 21$, $7 \cdot 12 = 84$) and $c \cdot c = e$ (e.g., $5 \cdot 6 = 30$, $6 \cdot 14 = 84$).
  - This suggests the group structure is isomorphic to the Klein 4-group, and so we should also have $a \cdot b = c$ ($2 \cdot 3 = 6$, $11 \cdot 7 = 77$, etc.), $a \cdot c = b$ ($2 \cdot 5 = 10$, $11 \cdot 6 = 66$), and $b \cdot c = a$ ($3 \cdot 14 = 42$, $10 \cdot 5 = 50$), which all do seem to hold.
  - We would expect, as above, that all of these should arise from algebraic identities. This is in fact the case, although it is not so easy to find and verify all of them.
  - But for example, we have $(a^2+21b^2)(2c^2+2cd+11d^2) = 2x^2+2xy+11y^2$ for $x = ac-bc+ad+10bd$ and $y = 2bc - ad + bd$, and also $(2a^2 + 2ab + 11b^2)(3c^2 + 7d^2) = 5x^2 + 4xy + 5y^2$ for $x = ac - ad + 2bc + 4bd$ and $y = -ac - ad + bc - 4bd$.

- Part of the difficulty is that in some cases, the integers represented by inequivalent forms are the same, so we cannot use tables to identify the group structure, nor can we necessarily identify the composition structure by searching for algebraic identities.

  - For example, consider the case $\Delta = -31$, where we showed that there are three equivalence classes of positive-definite forms represented by $x^2 + xy + 8y^2$, $2x^2 + xy + 4y^2$, and $2x^2 - xy + 4y^2$. Here are the integers less than 50 represented by each form:

	Form	Integers
e	$x^2 + xy + 8y^2$	0, 1, 4, 8, 9, 10, 14, 16, 20, 25, 28, 31, 32, 35, 36, 38, 40, 47, 49
a	$2x^2 + xy + 4y^2$	0, 2, 4, 5, 7, 8, 10, 14, 16, 18, 19, 20, 25, 28, 32, 35, 36, 38, 40, 41, 45, 49
b	$2x^2 - xy + 4y^2$	0, 2, 4, 5, 7, 8, 10, 14, 16, 18, 19, 20, 25, 28, 32, 35, 36, 38, 40, 41, 45, 49

  - We can see that the forms $2x^2 + xy + 4y^2$ and $2x^2 - xy + 4y^2$ represent the same integers, since they are related via an improper change of variables $(x, y) \mapsto (x, -y)$ of determinant -1.
  - As in the examples above, we can write down algebraic identities that yield a group structure on the equivalence classes of these forms. However, because of the presence of the improper change of variables relating two of the forms, we can also generate composition relations that do not yield a group structure.
  - For example, despite the fact that we want the norm form $x^2 + xy + 8y^2$ to be the identity element of the group, so that $e \cdot a = a$, we have the identity $(a^2 + ab + 8b^2)(2c^2 + cd + 4d^2) = 2x^2 \pm xy + 4y^2$ for $x = \pm(ac - 4bd)$ and $y = ad + 4bc - 2bc + bd$, so depending on our choice of sign, we could either take $e \cdot a = a$ (which is the identity we want if we are to have a group structure) or $e \cdot a = b$ (which would not give a group structure).

- Many of the properties of binary quadratic forms we have discussed were first treated by Legendre: for example, he gave the definition of a reduced form, discussed the equivalence of forms, and described a procedure for computing the composition of two forms.

- However, Legendre's treatment also allowed what we now call improper equivalence of forms (i.e., changes of coordinates with determinant -1), which collapses the equivalence classes further and makes it very difficult to identify the right composition structure.
- To give an example in the case where Δ is negative and even, Legendre observed that for forms $f(x,y) = ax^2 + 2bxy + cy^2$ and $g(x,y) = a'x^2 + 2b'xy + c'y^2$ with a, a' relatively prime, then if $B \equiv \pm b \pmod{a}$ and also $B \equiv \pm b' \pmod{a'}$, then $B^2 - \Delta/4 \equiv b^2 + (ac - b^2) \equiv 0 \pmod{a}$ and similarly $B^2 - \Delta/4 \equiv 0 \pmod{a'}$, so $B^2 - \Delta/4 \equiv 0 \pmod{aa'}$.
- One can then write down an appropriate linear change of variables to show that the product $f(x,y)g(x',y')$ is equal to $aa'X^2 + 2BXY + \dfrac{B^2 - \Delta/4}{aa'}Y^2$ for X and Y appropriate bilinear forms in x, y and x', y'.
- However, because of the choice of $\pm$ signs in Legendre's composition above, there are multiple different possible results of composing two forms, and (as with the example for $\Delta = -31$) these need not actually yield forms lying in the same equivalence class.
- The resolution of this quite tricky issue was first accomplished by Gauss, who introduced the notion of proper equivalence[5] (which is our relation $\sim$) and identified a consistent procedure for composing quadratic forms that does give them the structure of a group: this is known as Gauss direct composition. However, his treatment is fairly complicated, owing to the necessity of identifying the correct choice of compositions whenever there is more than one option, although it is quite remarkable how much of the general theory he was able to characterize, given that the notion of an abstract group was still decades away from being developed.

9.3.4 Dirichlet's Composition of Quadratic Forms

- We will describe a simplified composition law that is due to Dirichlet.
- Definition: Let $f(x,y) = ax^2 + bxy + cy^2$ and $g(x,y) = a'x^2 + b'xy + c'y^2$ be binary quadratic forms of discriminant Δ. Suppose that $\gcd(a, a', (b+b')/2) = 1$. Then the Dirichlet composition of $f(x,y)$ and $g(x,y)$ is the binary quadratic form $h(x,y) = Ax^2 + Bxy + Cy^2$ where $A = aa'$, B is the unique integer in $(-A, A]$ satisfying $B \equiv b \pmod{2a}$, $B \equiv b' \pmod{2a'}$, and $B^2 \equiv \Delta \pmod{4aa'}$, and $C = \dfrac{B^2 - \Delta}{4aa'}$.
  - Clearly the new form also has discriminant Δ, since $C = \dfrac{B^2 - \Delta}{4aa'} = \dfrac{B^2 - \Delta}{4A}$, and the coefficients A, B, C are integers since the assumptions on B indicate that $B^2 - \Delta$ is divisible by $4aa'$. We also remark that b and b' have the same parity since they are both congruent to Δ mod 2, so the gcd condition is well-posed.
  - It is less obvious why there is a unique value of B in $(-A, A]$ satisfying the simultaneous congruences $B \equiv b \pmod{2a}$, $B \equiv b' \pmod{2a'}$, and $B^2 \equiv \Delta \pmod{4aa'}$, but this can be shown[6] to hold in general.
  - The Dirichlet composition does yield composition identities like the ones we described earlier: by hypothesis, B is congruent to $b \pmod{2a}$ and to $b' \pmod{2a'}$, so by applying the appropriate power of T we see that $f(x,y) = ax^2 + bxy + cy^2$ and $g(x,y) = a'x^2 + b'xy + c'y^2$ are equivalent to the forms $f^*(x,y) = ax^2 + Bxy + a'Cy^2$ and $g^*(x,y) = a'x^2 + Bxy + aC'y^2$ respectively.
  - Then one has $f^*(x_1,y_1)g^*(x_2,y_2) = AX^2 + BXY + CY^2$ where $X = x_1x_2 - Cy_1y_2$ and $Y = ax_1x_2 + a'y_1y_2 + By_1x_2$. (Changing variables back yields an analogous identity for f and g.)
  - Example: For $\Delta = -40$, to compute the Dirichlet composition of $x^2 + 10y^2$ with itself, we see $A = 1 \cdot 1 = 1$, $B \equiv 0 \pmod{2}$, $B \equiv 0 \pmod{2}$, and $B^2 \equiv -40 \pmod{4}$, so that $B = 0$, and then $C = (B^2 - \Delta)/(4A) = 10$. Thus, the Dirichlet composition of $x^2 + 10y^2$ with itself is again $x^2 + 10y^2$.

[5] In fact, Gauss's notion of proper equivalence is the origin of the term "equivalence relation".

[6] First, if B is a solution to the first two congruences, then $(B-b)(B-b')$ is divisible by $2a \cdot 2a' = 4aa'$, which means that $B^2 - (b+b')B + bb' \equiv 0 \pmod{4aa'}$. Then the third congruence is equivalent to $(b+b')B \equiv bb' + \Delta \pmod{4aa'}$, and cancelling a factor of 2 yields the equivalent congruence $\frac{b+b'}{2}B \equiv \frac{bb'+\Delta}{2} \pmod{2aa'}$. Then by scaling the first two congruences, we see that the full system is equivalent to $a'B \equiv a'b \pmod{2A}$, $aB \equiv ab' \pmod{2A}$, and $\frac{b+b'}{2}B \equiv \frac{bb'+\Delta}{2} \pmod{2A}$. Finally, because $\gcd(a, a', (b+b')/2) = 1$, we can write $1 = pa + qa' + r\frac{b+b'}{2}$ for $p, q, r \in \mathbb{Z}$: then one may verify that $B = qa'b + pab' + r\frac{bb'+\Delta}{2}$ is a solution to this system and that the solution is unique modulo $2A$.

- Example: For $\Delta = -40$, to compute the Dirichlet composition of $x^2 + 10y^2$ with $2x^2 + 5y^2$, we see $A = 1 \cdot 2 = 2$, $B \equiv 0 \pmod 2$, $B \equiv 0 \pmod 4$, and $B^2 \equiv -40 \pmod 8$, so that $B = 0$, and then $C = (B^2 - \Delta)/(4A) = 5$. Thus, the Dirichlet composition of $x^2 + 10y^2$ with $2x^2 + 5y^2$ is $2x^2 + 5y^2$.
- Example: For $\Delta = -84$, to compute the Dirichlet composition of $2x^2 + 2xy + 11y^2$ with $3x^2 + 7y^2$, we see $A = 2 \cdot 3 = 6$, $B \equiv 2 \pmod 4$, $B \equiv 0 \pmod 6$, and $B^2 \equiv -84 \pmod{24}$, so that $B = 6$, and then $C = (B^2 - \Delta)/(4A) = 5$. Thus, the Dirichlet composition of $x^2 + 10y^2$ with $2x^2 + 2xy + 11y^2$ with $3x^2 + 7y^2$ is $6x^2 + 6xy + 5y^2$. This form is not reduced, but applying S yields $5y^2 - 5xy + 6y^2$ and then applying T yields the reduced form $5x^2 + 4xy + 5y^2$.
- It can be shown using direct manipulations that Dirichlet composition is well-defined on equivalence classes of forms (we will omit this argument, since the calculations are fairly involved).
- Thus, in situations where the condition for evaluating the Dirichlet composition is not met (i.e., when $\gcd(a, a', (b + b')/2) > 1$) we may instead use equivalent non-reduced forms for computing compositions.
- Example: For $\Delta = -40$, to compute the Dirichlet composition of $2x^2 + 5y^2$ with itself, we cannot use the composition formula directly since $\gcd(a, a', (b + b')/2) = 2$. Instead, if we compute the composition of $2x^2 + 5y^2$ with the equivalent form $5x^2 + 2y^2$ obtained by applying S, we get $A = 2 \cdot 5 = 10$, $B \equiv 0 \pmod 4$, $B \equiv 0 \pmod{10}$, and $B^2 \equiv -40 \pmod{40}$, so that $B = 0$, and then $C = (B^2 - \Delta)/(4A) = 1$. Thus, the Dirichlet composition of $2x^2 + 5y^2$ with $2x^2 + 5y^2$ is $10x^2 + y^2$. This form is not reduced, but applying S yields the reduced form $x^2 + 10y^2$.

- Dirichlet's composition law makes the collection of equivalence classes of forms of discriminant Δ into an abelian group:

- Theorem (Composition of Quadratic Forms): Suppose Δ is the discriminant of a quadratic integer ring and let $\mathcal{F}$ be the set of equivalence classes of quadratic forms of discriminant Δ. Then $\mathcal{F}$ has the structure of an abelian group under Dirichlet composition. The identity of $\mathcal{F}$ is the norm form on the quadratic integer ring Δ and the inverse of the class containing $ax^2 + bxy + cy^2$ is the class containing $ax^2 - bxy + cy^2$.

  - Proof: As we remarked on above, Dirichlet composition is well-defined on equivalence classes. To show that it is an abelian group, we must show that the operation is associative, commutative, and that the identity and inverses are as claimed.
  - Associativity is a direct (albeit quite tedious) calculation, which we will omit.
  - Commutativity is immediate from the definition, because the definition of the Dirichlet composition is symmetric in f and g.
  - For the identity, we want to compose the norm form on $\mathcal{O}_{\sqrt{D}}$ with a reduced form $ax^2 + bxy + cy^2$. Clearly the gcd condition is satisfied, since the norm form has leading coefficient 1. Then $A = a$ and we require B to be the unique integer in $(-A, A]$ satisfying $B \equiv b \pmod{2a}$ with $B^2 \equiv \Delta \pmod{4a}$, but clearly $B = b$ satisfies this condition so since $b \in (-|a|, |a|]$ is reduced, we simply have $B = b$. Then $C = \dfrac{B^2 - \Delta}{4a} = c$, and so the result of the composition is again just $ax^2 + bxy + cy^2$.
  - For inverses, we want to compose $ax^2 + bxy + cy^2$ with $ax^2 - bxy + cy^2$. If we instead apply S to the second form to obtain $cx^2 + bxy + ay^2$, we then have $\gcd(a, c, (b + b)/2) = \gcd(a, b, c) = 1$ because Δ is squarefree except for a factor of 4. Then the Dirichlet composition of $ax^2 + bxy + cy^2$ with $cx^2 + bxy + ay^2$ has $A = ac$, $B \equiv b \pmod{2a}$, $B \equiv b \pmod{2c}$, and $B^2 \equiv \Delta \pmod{4ac}$, but clearly $B = b$ satisfies this condition. Then $C = \dfrac{b^2 - (b^2 - 4ac)}{4ac} = 1$, so the resulting composition is $acx^2 + bxy + y^2$. Applying S yields $x^2 - bxy + acy^2$, and then applying the appropriate power of T reduces this to a form $x^2 + (\Delta/4)y^2$ if Δ is even, or $x^2 + xy + \frac{1-\Delta}{4}y^2$ if Δ is odd, and this is precisely the norm form, as claimed.

9.3.5 Quadratic Forms and Ideal Class Groups

- In fact, the abelian group we obtain by composing binary quadratic forms of discriminant Δ is essentially just the ideal class group of the quadratic integer ring $\mathcal{O}_{\sqrt{D}}$:

- Theorem (Quadratic Forms and Ideal Class Groups): Suppose $\Delta < 0$ is the discriminant of a quadratic integer ring $\mathcal{O}_{\sqrt{D}}$. Then the group $\mathcal{F}$ of equivalence classes of binary quadratic forms of discriminant Δ

under composition is isomorphic to the group $\mathcal{I} \times \{\pm 1\}$ of equivalence classes of ideals of $\mathcal{O}_{\sqrt{D}}$ under ideal multiplication, together with a sign ± 1. More explicitly, if $\varphi_{FI} : \mathcal{F} \to \mathcal{I} \times \{\pm 1\}$ is the map that sends a quadratic form $ax^2 + bxy + cy^2$ to the pair $((a, \frac{-b+\sqrt{\Delta}}{2}), \text{sign}(a))$, then φ_{FI} is a group isomorphism, and its inverse is the map $\varphi_{IF} : \mathcal{I} \times \{\pm 1\} \to \mathcal{F}$ that takes an ideal $I = (n, \omega)$ of norm n (with $\omega \notin \mathbb{Z}$) along with a sign $s \in \{\pm 1\}$ to the quadratic form $s \cdot \frac{N(nx - s\omega y)}{N(I)}$.

- What this result says is that, up to some minor business with $\pm$ signs, we obtain an isomorphism between the group of binary quadratic forms under composition with the ideal class group.
- In particular, when $\Delta < 0$, the positive-definite forms have a $+$ sign and the negative-definite forms have a $-$ sign: thus, the theorem gives an isomorphism between classes of positive-definite forms (which are uniquely represented by reduced positive-definite forms) and classes in the ideal class group.
- For positive Δ, the nonuniqueness of reduced forms in a given equivalence class causes issues with the sign ± 1. However, one can essentially salvage this result by restricting the definition of reduced forms and accounting for the sign ambiguity, which depends on whether the fundamental unit of $\mathcal{O}_{\sqrt{D}}$ has positive or negative norm.
- Proof: First we show that the maps are well defined.
- If we start with a quadratic form $ax^2 + bxy + cy^2$ of discriminant Δ, then as we have observed numerous times, $b \equiv \Delta \bmod 2$, and so $\frac{-b+\sqrt{\Delta}}{2} \in \mathcal{O}_{\sqrt{D}}$. Then the result of applying φ_{FI} to $ax^2 + bxy + cy^2$ is indeed an ideal of $\mathcal{O}_{\sqrt{D}}$.
- Furthermore, if we apply S or T to the quadratic form, the resulting ideal class is not changed: applying T changes b to $b + 2a$, and as ideals we have $(a, \frac{-b+\sqrt{\Delta}}{2}) = (a, \frac{-(b+2a)+\sqrt{\Delta}}{2})$, while applying S changes $(a, \frac{-b+\sqrt{\Delta}}{2})$ to $(c, \frac{b+\sqrt{\Delta}}{2})$, and these are the same ideal class because $\frac{b+\sqrt{\Delta}}{2} \cdot (a, \frac{-b+\sqrt{\Delta}}{2}) = a \cdot (c, \frac{b+\sqrt{\Delta}}{2})$.
- Therefore φ_{FI} is a well-defined map from $\mathcal{F}$ to $\mathcal{I} \times \{\pm 1\}$.
- For φ_{IF}, suppose that (n, ω) is an ideal of $\mathcal{O}_{\sqrt{D}}$ and $s \in \{\pm 1\}$. Then $N(nx + s\omega y) = (nx + s\omega y)(nx + s\overline{\omega} y) = n^2x^2 + sn(\omega + \overline{\omega})xy + s^2\omega\overline{\omega}y^2$ is a quadratic form. Furthermore, all of its coefficients are divisible by the norm of I since $(N(I)) = I \cdot \overline{I} = (n^2, n(\omega + \overline{\omega}), \omega\overline{\omega})$, so the quotient $s\frac{N(nx - s\omega y)}{N(I)} = snx^2 + (\omega + \overline{\omega})xy + s\frac{\omega\overline{\omega}}{n}y^2$ is still a quadratic form with integer coefficients. Its discriminant is $\frac{n^2(\omega + \overline{\omega})^2 - 4n^2\omega\overline{\omega}}{N(I)^2} = (\omega - \overline{\omega})^2 = \Delta$ since we may assume ω is of the form $\frac{-b+\sqrt{\Delta}}{2}$ by an appropriate linear change of basis for I.
- Furthermore, if we scale the ideal I by a principal factor, the resulting quadratic form is not changed, since the ratio $s\frac{N(nx - s\omega y)}{N(I)}$ is invariant if we scale n, ω by the same constant. Therefore, φ_{IF} is a well-defined map from $\mathcal{I} \times \{\pm 1\}$ to $\mathcal{F}$.
- If we apply φ_{IF} to the ideal $\varphi_{FI}(ax^2 + bxy + cy^2) = ((a, \frac{-b+\sqrt{\Delta}}{2}, s)$ where $s = \text{sign}(a)$, we obtain the quadratic form $s \cdot \frac{1}{|a|}(ax - s\frac{-b+\sqrt{\Delta}}{2}y)(ax - s\frac{-b-\sqrt{\Delta}}{2}y) = \frac{1}{a}(a^2x^2 - sabxy + s^2acy^2) = ax^2 + bxy + cy^2$. Thus, $\varphi_{IF} \circ \varphi_{FI}$ is the identity.
- Oppositely, if we apply φ_{FI} to the ideal $I = (n, \omega)$ and sign s, we obtain $snx^2 + (\omega + \overline{\omega})xy + s\frac{\omega\overline{\omega}}{n}y^2$ as noted above. Then since $n > 0$, we see $\text{sign}(sn) = s$, and the underlying ideal is generated by sn and $\frac{-(\omega+\overline{\omega})+\sqrt{\Delta}}{2} = \omega$, hence is simply $I = (n, \omega)$. Thus, $\varphi_{FI} \circ \varphi_{IF}$ is the identity as well, so the maps φ_{FI} and φ_{IF} are inverses.
- Finally, φ_{IF} is multiplicative on ideals, since both the numerator and denominator are multiplicative. Thus, φ_{IF} is a group isomorphism and its inverse is φ_{FI}, which is the desired result. (It is harder to check directly that φ_{FI} is an isomorphism since the description of Dirichlet composition is more complicated, but this fact is embedded in the results we have already shown.)

- Here are some examples using forms and ideal class groups we have previously computed:

- ○ Example: For $D = -1$, with $\Delta = -4$, we have two reduced binary quadratic forms x^2+y^2 and $-x^2-y^2$. Applying the map φ_{FI} to them yields the same ideal $(1, i) = \mathbb{Z}[i]$ along with a sign ± 1. Conversely, applying φ_{IF} to $I = (1, i)$ and the sign $+1$ yields the quadratic form $\dfrac{N(x+iy)}{N(1)} = x^2+y^2$, while applying φ_{IF} to $I = (1, i)$ and the sign -1 yields the quadratic form $-\dfrac{N(x-iy)}{N(1)} = -x^2-y^2$.
- ○ Example: For $D = -10$, with $\Delta = -40$, we have two reduced positive-definite forms x^2+10y^2 and $2x^2+5y^2$. Applying the map φ_{FI} to x^2+10y^2 yields the ideal $(1, \sqrt{-10}) = (1)$ representing the trivial ideal class, while $2x^2+5y^2$ yields the ideal $(2, \sqrt{-10})$ which as we have seen represents the nontrivial ideal class. Conversely, applying φ_{IF} to $I = (1, \sqrt{-10})$ and the sign $+1$ yields the quadratic form $\dfrac{N(x+y\sqrt{-10})}{1} = x^2+10y^2$, while applying φ_{IF} to $I = (2, \sqrt{-10})$ and the sign $+1$ yields the quadratic form $\dfrac{N(2x+y\sqrt{-10})}{2} = 2x^2+5y^2$.
- ○ Example: For $D = \Delta = -31$, we have three reduced positive-definite forms $x^2+xy+8y^2$, $2x^2+xy+4y^2$, and $2x^2-xy+4y^2$. Applying the map φ_{FI} to $x^2+xy+8y^2$ yields the ideal $(1, \frac{-1+\sqrt{-31}}{2}) = (1)$ representing the trivial ideal class, while $2x^2+xy+4y^2$ yields the ideal $(2, \frac{-1+\sqrt{-31}}{2})$ which is one of the two ideal factors of (2), and $2x^2+xy+4y^2$ yields $(2, \frac{1+\sqrt{-31}}{2})$ which is the other ideal factor of (2).

- As an application of computing the class group, we can give characterizations (in certain cases) of the primes represented by a quadratic form.
  - ○ To illustrate, consider the two inequivalent positive-definite forms x^2+10y^2 and $2x^2+5y^2$ with discriminant $\Delta = -40$.
  - ○ From our results earlier, we see that a prime p is represented by one of these forms if and only if -40 is a quadratic residue modulo p. Excluding $p = 2, 5$ (which are both clearly represented by the second form and not the first), by quadratic reciprocity we have $(\frac{-40}{p}) = (\frac{-10}{p}) = (\frac{-2}{p})(\frac{5}{p}) = (\frac{-2}{p})(\frac{p}{5})$. Since $(\frac{-2}{p}) = +1$ for $p \equiv 1, 3 \pmod 8$ and $(\frac{p}{5}) = +1$ for $p \equiv 1, 4 \pmod 5$, we see that $(\frac{-40}{p}) = +1$ for $p \equiv 1, 9, 11, 19 \pmod{40}$ (both symbols are $+1$) and for $p \equiv 7, 13, 23, 37 \pmod{40}$ (both symbols are -1).
  - ○ Thus, a prime $p \neq 2, 5$ is represented by one of x^2+10y^2 and $2x^2+5y^2$ if and only if $p \equiv 1, 7, 9, 11, 13, 19, 23, 37 \pmod{40}$.
  - ○ But if p is represented by x^2+10y^2, then $p \equiv x^2 \pmod 5$, so p must be a quadratic residue modulo 5 and so $p \equiv 1, 9, 11, 19 \pmod{40}$.
  - ○ Likewise, if p is represented by $2x^2+5y^2$, then $p \equiv 2x^2 \pmod 5$ so p must be a quadratic nonresidue modulo 5 and so $p \equiv 7, 13, 23, 37 \pmod{40}$.
  - ○ Since these cases partition the primes, we conclude that the primes represented by x^2+10y^2 are precisely the primes $p \equiv 1, 9, 11, 19 \pmod{40}$, while the primes represented by $2x^2+5y^2$ are precisely 2, 5, and the primes $p \equiv 7, 13, 23, 37 \pmod{40}$.
  - ○ By composing these forms, we can then classify all integers represented by these forms: they may have arbitrary square factors times a product of primes 2, 5, and $p \equiv 1, 7, 9, 11, 13, 19, 23, 37 \pmod{40}$. Since x^2+10y^2 represents the trivial element of the class group, we also see that the form representing n will be x^2+10y^2 when the total number of primes dividing n to an odd power among $2, 5$ and $p \equiv 7, 13, 23, 37 \pmod{40}$ is even, while the form will be $2x^2+5y^2$ when the total number of such primes is odd.

- There are many open problems regarding class groups of quadratic fields. A natural immediate question is to determine which quadratic integer fields have class number n for each fixed n.
  - ○ The case $n = 1$ is known as the class number 1 problem, and (per our earlier discussion) it is equivalent to asking which quadratic integer rings have unique factorization. It was conjectured by Gauss in 1801 that there are only finitely many imaginary quadratic fields of class number 1.
  - ○ It was proven by Heilbronn in 1934 that there are only finitely many imaginary quadratic fields of any fixed class number (so that in particular the class number $h(-d) \to \infty$ as $d \to \infty$). This result was sharpened by Siegel to obtain an ineffective bound $h(-d) \geq c\sqrt{d}$ for a positive constant c whose value was not computable.

- Heilbronn also showed that there were at most 10 imaginary quadratic fields of class number 1; since 9 such fields, corresponding to $D = -1$, -2, -3, -7, -11, -19, -43, -67, and -163 were known, this meant there could exist at most one more.
- The nonexistence of this 10th field was essentially proven by Heegner in 1952 using modular forms, but his proof had some minor gaps and it was not accepted[7] until Stark gave a full proof of the result in 1967. Baker also gave a proof, using an entirely different method (linear forms in logarithms), in 1966.
- These results have subsequently been extended to classify all imaginary quadratic fields with a given small class number: for instance, there are 18 fields of class number 2, corresponding to $D = -5, -6, -10, -13, -15, -22, -35, -37, -51, -58, -91, -115, -123, -187, -235, -267, -403, -427$, there are 16 fields of class number 3, and so forth.

- For real quadratic fields, the results are quite different: Gauss conjectured in this case that there are infinitely many real quadratic fields of class number 1.
  - This problem of determining whether there actually are infinitely many real quadratic fields of class number 1 is still open (as of 2021). It can be shown, however, that any field whose discriminant has at least 3 distinct prime factors has even class number.
  - Many small values of D do yield real quadratic fields of class number 1. In fact, the only values of $D < 100$ that do not are $D = 10, 15, 26, 30, 34, 35, 39, 42, 51, 55, 58, 65, 66, 70, 74, 78, 85, 87, 91, 95$ which all have class number 2, along with $D = 79$ of class number 3 and $D = 82$ of class number 4.

- There are various conjectures about various aspects of the class groups of real and imaginary quadratic fields.
  - One set of predictions are the Cohen-Lenstra heuristics, which posit, for odd primes p, the density with which any given abelian p-group will appear as the p-power torsion part of a class group (i.e., the Sylow p-subgroup) of a real or imaginary quadratic field.
  - For the prime $p = 2$, the structures of p-power torsion subgroups of class groups are fully understood, and are consequences of what is called genus theory, which is a name due to Gauss (as is the term "equivalence class", which first appeared in Gauss's treatment of quadratic forms) that has nothing to do with other uses of the word "genus", e.g., in topology.
  - Intuitively, the Cohen-Lenstra heuristics say that the probability, in an appropriate sense, that a given abelian p-group P will occur as the p-part of the class group of an imaginary quadratic field should be proportional to $1/\#\mathrm{Aut}(P)$. This may initially seem to be a rather unnatural weighting, but in fact it is quite sensible in the appropriate context: given a group acting on a set X, if we wish to select an orbit of the group uniformly at random, we should weight each of the elements of X by 1 over the size of its orbit and then pick an element of X at random with that weighting.
  - By summing $1/\#\mathrm{Aut}(P)$ over all finite abelian p-groups P, one obtains a constant μ_P, which can be computed (though not easily). Then the Cohen-Lenstra heuristics predict that the proportion of imaginary quadratic fields whose p-power torsion subgroup is isomorphic to P is equal to $\frac{1/\#\mathrm{Aut}(P)}{\mu_P}$.
  - Some various results for other primes: the probability that the class number is divisible by 3 (i.e., that the 3-part of the class group is not trivial) is approximately 43.99%, the probability that it is divisible by 5 is approximately 23.97%, and the probability that it is divisible by 7 is approximately 16.32%.
  - A similar heuristic holds for real quadratic fields, although the weighting is slightly different. For real quadratic fields, the probability that a prime p divides the class number is predicted to be $1 - \prod_{k\geq 2}(1 - p^{-k})$, which for $p = 3$ is approximately 15.98%, for $p = 5$ is approximately 4.96%, and for $p = 7$ is approximately 2.37%.
  - All of these results agree extremely well with the available numerical data.

Well, you're at the end of my handout. Hope it was helpful.
Copyright notice: This material is copyright Evan Dummit, 2014-2026. You may not reproduce or distribute this material without my express permission. No aspect of this material was created or edited by a large-language model or similar technology ("generative AI"), and this material may not be used for training of such technology.

[7] Heegner was not a professional mathematician (he was in fact a radio engineer and high school teacher), which certainly contributed to the lack of belief in his claim to have settled a 150-year-old conjecture of Gauss by the broader mathematical community. Sadly, he died in 1965, before his results gained general acceptance.