7 Quadratic forms in n variables

In order to understand quadratic forms in n variables over $\mathbb{Z}$, one is let to study quadratic forms over various rings and fields such as $\mathbb{Q}$, $\mathbb{Q}_p$, $\mathbb{R}$ and $\mathbb{Z}_p$. This is consistent with the basic premise of algebraic number theory, which was the idea that to study solutions of a Diophantine equation in $\mathbb{Z}$, it is useful study the equation over other rings.

Definition 7.0.8. *Let R be a ring. A* **quadratic form** *in n variables (or n-ary quadratic form) over R is a homogenous polynomial of degree 2 in $R[x_1, x_2, \ldots, x_n]$.*

For example $x^2 - yz$ is a ternary (3 variable) quadratic form over any ring, since the coefficients ±1 live inside any ring R. On the other hand $x^2 - \frac{1}{2}yz$ is not a quadratic form over $\mathbb{Z}$, since $-\frac{1}{2} \notin \mathbb{Z}$, but it can be viewed as a quadratic form over $\mathbb{Q}$, $\mathbb{Z}_p$ for $p \neq 2$, $\mathbb{Q}_2$, $\mathbb{R}$ or $\mathbb{C}$ since $-\frac{1}{2}$ lies in each of those rings. In fact it can be viewed as a quadratic form over $\mathbb{Z}/n\mathbb{Z}$ for any odd n, as -2 is invertible mod n whenever n is odd.

The subject of quadratic forms is vast and central to many parts of mathematics, such as linear algebra and Lie theory, algebraic topology, and Riemannian geometry, as well as number theory. One cannot hope to cover everything about quadratic forms, even just in number theory, in a single course, let alone one or two chapters. I will describe the classification of quadratic forms over $\mathbb{Q}_p$ and $\mathbb{R}$ without proof, explain how one can use this to study forms over $\mathbb{Z}_p$ and $\mathbb{Z}$, subsequently prove Gauss' and Lagrange's theorems on sums of 3 and 4 squares, and then briefly explain some of the general theory of representation of numbers by quadratic forms. In particular, we will describe how studying forms over $\mathbb{Z}_p$ generalizes Gauss's genus theory and lead to Siegel's mass formula, which is a generalization of Dirichlet's mass formula to n-ary quadratic forms.

The main "algebraic" question about quadratic forms is how they can be classified, up to equivalence.

Definition 7.0.9. *Let $Q_1(x) = Q_1(x_1, \ldots, x_n)$ and $Q_2(x) = Q_2(x_1, \ldots, x_n)$ be n-ary quadratic forms over a ring R. We say Q_1 and Q_2 are* **equivalent over** *R denoted $Q_1 \sim Q_2$, or $Q_1 \sim_R Q_2$ when we want to specify R, if there exists*

$$\sigma \in \mathrm{GL}_n(R)$$

such that

$$Q_2(x) = Q_1(\sigma x).$$

In other words, two forms will be equivalent over R if one is obtained from the other by an invertible (linear) change of variables over R. This is the same as our definition of equivalence (not proper equivalence) for binary quadratic forms over $\mathbb{Z}$. Note that equivalent forms over R will represent the same numbers.

References for this chapter are [Serre], [Cassels], [Gerstein] and [Iwaniec].

7.1 Quadratic forms over fields

The main question about quadratic forms over fields is how they can be classified, and we start with fields because the classification over fields is much simpler than the classification over rings.

Let F be a field of characteristic not 2, and Q be an n-ary quadratic form over F. We can write

$$Q(x_1, \ldots, x_n) = \sum_{i \le j} c_{ij} x_i x_j = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} x_1 & \\ x_2 & \\ \vdots & x_n \end{pmatrix}$$

where A is a symmetric matrix in $M_n(F)$. Precisely let $A = (a_{ij})$ where

$$a_{ij} = \begin{cases} c_{ii} & i = j \\ \frac{1}{2} c_{ij} & i < j \\ \frac{1}{2} c_{ji} & i > j \end{cases}.$$

For example if $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 5x_1x_3 + 6x_2x_3$, then we may write

$$Q(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 1 & \frac{5}{2} \\ 1 & 2 & 3 \\ \frac{5}{2} & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

In this way, quadratic forms in n variables correspond to symmetric $n \times n$ matrices. Symmetric $n \times n$ matrices A correspond to symmetric bilinear forms $B(x, y) = x^T A y$ on F^n, hence quadratic forms $Q(x)$ are essentially the same as symmetric bilinear forms $B(x, y)$ (just set $Q(x) = B(x, x)$), which is how they arise in linear algebra and Lie groups.

We say Q is **nondegenerate** if the determinant of the associated matrix is nonzero. This essentially means that Q is not equivalent to a quadratic form in less than n variables. We will always assume this.

The first classification results for quadratic forms were in the cases $F = \mathbb{R}$ and $F = \mathbb{C}$. Let's first go through these.

Theorem 7.1.1. (Sylvester) *Let Q be be a nondegenerate quadratic form in n-variables over $\mathbb{R}$. Then Q is equivalent to $x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$ for some $1 \le k \le n$. Further no two of these are equivalent.*

Proof. Since any symmetric matrix is diagonalizable over R, up to equivalence we may assume the matrix for Q is $\mathrm{diag}(a_1, \ldots, a_n)$, i.e., Q is the *diagonal* form $a_1x_1^2 + a_2x_2^2 + \cdots a_nx_n^2$. What Q being nondegenerate means is that no $a_i = 0$ (or else the determinant of the diagonal matrix would be 0). Thus we can make the (invertible) change of variables which replaces each x_i with $\frac{1}{\sqrt{|a_i|}} x_i$. Under this transformation, Q becomes

$$Q(x_1, \ldots, x_n) = sgn(a_1)x_1^2 + sgn(a_2)x_2^2 + \cdots + sgn(a_n)x_n^2,$$

where $sgn(a_i) = \frac{a_i}{|a_i|}$ is the sign of a_i. Since we can permute the x_i's, we can in fact assume the first k a_i's are positive and the remaining a_i's are negative.

This shows any Q is equivalent to some $x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$. Note that the a_i's are the eigenvalues of the matrix A for Q. Sylvester showed that the number of positive and negative eigenvalues of $S^T A S$ is the same for any invertible matrix S. (This known as Sylvester's law of inertia.) This proves the classification theorem. □

If we think back to the notion of definite and indefinite forms, only two forms (up to equivalence) are definite (represent only positive or negative values), namely the positive definite form $x_1^2+\cdots+x_n^2$ and the negative definite form $-x_1^2-\cdots-x_n^2$.

Contrast what happens over $\mathbb{R}$ with what happens over $\mathbb{Q}$.

Theorem 7.1.2. *Let Q be a nondegenerate quadratic form in n-variables over $\mathbb{C}$. Then Q is equivalent to $x_1^2+x_2^2+\cdots x_n^2$.*

Proof. As in the real case, we may assume Q is of the form $a_1x_1^2+\cdots+a_nx_n^2$. But now $\sqrt{a}_i \in \mathbb{C}$ for all i, so making the change of variables $\frac{1}{\sqrt{a}_i}x_i$, proves the theorem. □

Over $\mathbb{C}$, there is no real notion of definite or indefinite since squares may be positive or negative. In any case, there is only one form over $\mathbb{C}$, up to equivalence.

Note that both over $\mathbb{R}$ and $\mathbb{C}$, the classification of quadratic forms is much simpler than the classification of binary quadratic forms over $\mathbb{Z}$. For one, there are infinitely many equivalence classes of binary quadratic forms (with no restriction on the discriminant), and even for a fixed discriminant the structure is rather complicated (though surprisingly beautiful, in that we have Gauss's composition law) In particular, while the discriminant is an invariant of the form over $\mathbb{Z}$, this is not true over $\mathbb{R}$ or $\mathbb{C}$. Over $\mathbb{R}$, there is a single invariant of a quadratic form, called the *signature* of the form, which is the number of $+1$ coefficients minus the number of -1 coefficients, assuming the form is written as $x_1^2+\cdots+x_k^2-x_{k+1}^2-\cdots-x_n^2$.

In general, for a quadratic form Q over any field F (characteristic not 2), we may make a change of variables to write Q as a diagonal form

$$Q(x_1,\ldots,x_n)=a_1x_1^2+a_2x_2^2+\cdots+a_nx_n^2.$$

Then the question of classification becomes simply a question of whether each $\sqrt{a_i}\in F$. If so, then we can make a change of variables $x_i \mapsto \frac{1}{\sqrt{a_i}}x_i$ to see Q is equivalent to $x_1^2+\cdots+x_n^2$. In particular, if F is algebraically closed, $\sqrt{a_i}$ is always in F, so there is only one (nondegenerate) quadratic form in n-variables up to equivalence.

In light of the above, the following result should be fairly evident.

Proposition 7.1.3. *Any n-ary form Q over F is equivalent to*

$$a_1x_1^2+a_2x_2^2+\cdots+a_nx_n^2,$$

where each a_i lies in a set of representatives for $F^\times/F^{\times(2)}$. Here $F^{\times(2)}$ denotes the subgroup of squares of $F^\times$.

One can show there are three invariants for a quadratic form $Q=a_1x_1^2+\cdots a_nx_n^2$, the **rank** (or number of variables) n, the **discriminant** $\mathrm{disc}(Q)=a_1a_2\cdots a_n$, and the **Hasse invariant** $\epsilon(Q)=\prod_{i<j}\left(\frac{a_i,a_j}{F}\right)=\pm1$. Here $\left(\frac{a,b}{F}\right)$ is the **Hilbert symbol** which is defined to be $+1$ if $ax^2+by^2=z^2$ has a nonzero solution over F and -1 otherwise.

Proposition 7.1.4. *For p odd, a set of representatives for $\mathbb{Q}_p^\times/\mathbb{Q}_p^{\times(2)}$ is $\{1,p,u,up\}$ where $u\in\mathbb{Z}$ satisfies $\left(\frac{u}{p}\right)=-1$. This quotient group is isomorphic to $C_2\times C_2$.*

Proposition 7.1.5. *A set of representatives for $\mathbb{Q}_2^\times/\mathbb{Q}_2^{\times(2)}$ is $\{\pm1,\pm2,\pm5,\pm10\}$. This quotient group is isomorphic to C_2^3.*

Theorem 7.1.6. *Let Q_1 and Q_2 be quadratic forms over $\mathbb{Q}_p$. They are equivalent if and only if they have the same rank, discriminant and Hasse invariant.*

See [Serre] for proofs.

Things are much more complicated over $\mathbb{Q}$ since the quotient group $\mathbb{Q}^\times/\mathbb{Q}^{\times(2)}$ is infinite. For instance, the primes $2, 3, 5, 7, 11, 13, \ldots$ are all distinct in $\mathbb{Q}^\times/\mathbb{Q}^{\times(2)}$, as can be easily checked by the exercise below.

Exercise 7.1. *Suppose p and q are two distinct primes p and q. Show p and q do not differ by a rational square. Consequently $\mathbb{Q}^\times/\mathbb{Q}^{\times(2)}$ is infinite.*

The way to study forms over the *global field* $\mathbb{Q}$ is by reducing the question to studying forms over the *local fields* $\mathbb{Q}_p$. To be a little more precise, the philosophy is that we can study problems over $\mathbb{Q}$ by studying the associated problems in *all* of its completions (w.r.t. nontrivial absolute values), in other words in each $\mathbb{Q}_p$ and $\mathbb{R}$. This notion is called **Hasse's local-to-global principle**.

The simplest precise form of the local-to-global principle is the following theorem of Hasse and Minkowski. For a quadratic form Q over a field F, we always have $Q(0) = 0$, and the simplest representation question is whether $Q(x) = 0$ for any nonzero $x \in F^n$. If $Q(x) = 0$ for some $0 \neq x \in F^n$, we say Q **represents** 0 (nontrivially), or Q is **isotropic**.

Theorem 7.1.7. (Hasse–Minkowski) *Let Q be a quadratic form over $\mathbb{Q}$. Then Q represents 0 (nontrivially) over $\mathbb{Q}$ if and only if it does over $\mathbb{Q}_p$ for each p and over $\mathbb{R}$.*

We remark that this statement makes sense because any form over $\mathbb{Q}$ can be regarded as a form over $\mathbb{Q}_p$ or $\mathbb{R}$ since $\mathbb{Q} \subseteq \mathbb{Q}_p$ and $\mathbb{Q} \subseteq \mathbb{R}$.

Proof. (Sketch) Let Q be a quadratic form of rank n over $\mathbb{Q}$. By the above we can write $Q(x_1, \ldots, x_n) = a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2$ with each $a_i \in \mathbb{Q}$. Since Q represents 0 (over $\mathbb{Q}$, $\mathbb{Q}_p$ or $\mathbb{R}$) if and only if the form $\frac{1}{a_1}Q$ does, we may replace Q with $\frac{1}{a_1}Q$ to assume that $a_1 = 1$. Also, by replacing x_i with an appropriate multiple c_ix_i, we may assume each $a_i \in \mathbb{Z}$ and squarefree. Further it is clear that if Q represents 0 over $\mathbb{Q}$, it also will over the completions $\mathbb{Q}_p$ and $\mathbb{R}$. Hence it suffice to show that Q represents 0 over $\mathbb{Q}$ if it does over each $\mathbb{Q}_p$ and $\mathbb{R}$. We consider various cases.

$n = 1$. If $n = 1$, then we have $Q(x_1) = x_1^2$, so Q does not represent 0 (nontrivially) over any field, and there is nothing to prove.

$n = 2$. If $n = 2$, write $Q(x, y) = x^2 - ay^2$ (here $a = -a_2$). Then Q represents 0 over a field F if and only if $x^2 = ay^2$ has a solution in F, i.e., if and only if $a = (\frac{x}{y})^2$ has a solution in F, i.e., if and only if a is a square in F. So we want to prove that if a is a square in $\mathbb{Q}_p$ and a is a square in $\mathbb{R}$, then a is a square in $\mathbb{Q}$. The condition that $a \in \mathbb{R}^{\times(2)}$ just means $a > 0$. Note that $a \in \mathbb{Q}_p^{\times(2)}$ means that $\mathrm{ord}_p(a)$ is even for each p (since $a = b^2$ implies $\mathrm{ord}_p(a) = 2\mathrm{ord}_p(b)$).

Write $a = \frac{r}{s}$ where $r, s \in \mathbb{Z}$ in reduced form. If p is a prime dividing r or s, then $\mathrm{ord}_p(a)$ even means that p occurs to an even power in the prime factorization of r and s (it will be positive for one of r and s, and 0 for the other). Hence $\frac{r}{s} = a$ is a square in $\mathbb{Q}$.

$n \geq 3$. One can treat the cases $n = 3$ (due to Legendre) and $n = 4$ separately, and then prove the theorem for $n \geq 5$ by induction on n by breaking the form up into the sum of a binary form with a form of rank $n - 2$. This is done with fairly elementary p-adic analysis. □

Now one might wonder if the Hasse–Minkowski theory unnecessarily complicates the problem by requiring us to check things over infinitely many fields $\mathbb{Q}_p$. In practice however, one only needs things for finitely many primes. This can even be seen in our proof of the $n = 2$ case: to check if $a = \frac{r}{s}$ is a square in $\mathbb{Q}$, it suffices to check it over $\mathbb{R}$ and $\mathbb{Q}_p$ for just the primes p dividing r and s.

One reason representing 0 is a basic question is the following.

Proposition 7.1.8. *Suppose Q represents* 0 *over F. Then Q is* **universal***, i.e., Q represents every element of F.*

The proof is fairly simple: A nontrivial representation $Q(x_1, \ldots, x_n) = 0$ implies Q "contains" a product of linear forms. For example, $x_1^2 + x_2^2 - x_3^2 = x_1^2 + (x_2 + x_3)(x_2 - x_3) = x_1^2 + yz$ where $y = x_2 + x_3$, $z = x_2 - x_3$. Setting $x_1 = 0$, $y = 1$ and letting z vary, we see this form is universal. The general argument is similar, but we will not go through the details—in any event, this example is essentially the case we will be considering in the next section.

In fact, the Hasse–Minkowski theorem really contains information about a form representing any $a \in \mathbb{Q}$.

Exercise 7.2. *Consider a form $Q(x_1, \ldots, x_n)$ over $\mathbb{Q}$ and set $Q_a(x_1, \ldots, x_{n+1}) = Q(x_1, \ldots, x_n) - ax_{n+1}^2$ for $a \in \mathbb{Q}$. Show Q represents a if and only if Q_a represents* 0*. (Hint: use Proposition 7.1.8.)*

Exercise 7.3. *Let Q be a quadratic form over $\mathbb{Q}$. Deduce from Hasse–Minkowski that Q represents some $a \in \mathbb{Q}$ over $\mathbb{Q}$ if and only if it does over $\mathbb{R}$ and each $\mathbb{Q}_p$. (Hint: use the previous exercise.)*

One can show that *any* quadratic form of rank ≥ 4 over $\mathbb{Q}_p$ represents all p-adic numbers. Then from previous exercise, one can deduce that for any Q of rank ≥ 4, Q represents $a \in \mathbb{Q}$ over $\mathbb{Q}$ if and only if it does over $\mathbb{R}$. With this you should easily be able to convince yourself that form of rank ≥ 4 over $\mathbb{Q}$ either represents (i) all nonnegative rationals, (ii) all nonpositive rationals, or (iii) all rationals, just based on the signs of the coefficients of the form.

This suggests the following phenonemon—it is easy to determine what numbers are represented by a form Q with rank ≥ 4 (at least over $\mathbb{Q}$), and it is also fairly easy to determine what numbers are represented by a form of rank 2 (or 1), but the case of rank 3 is considerably more subtle. This phenomenon persists when restricting to forms over $\mathbb{Z}$ as well. This notion of some problems being easy in low dimensions and high dimensions, but very subtle in middle (often 3 or 4) dimensions, occurs in other areas of mathematics also, a famous example being the classification of n-manifolds, which is "simple" in dimensions ≤ 2 or ≥ 5.

7.2 Sums of Squares

Ideally, one would like to use the Hasse–Minkowski theorem to reduce representation problems over $\mathbb{Z}$ to problems over $\mathbb{Z}_p$. The general situation is rather complicated, so for simplicity and completeness, we will show how to apply these ideas to the cases of sums of three and four squares, following [Serre] and [Gerstein].

Let's start with the sum of 3 squares over a field F. Recall the Hilbert symbol $\left(\frac{a,b}{F}\right)$ is 1 if $ax^2 + by^2 - z^2$ represents 0 and is -1 else. Hence $x^2 + y^2 + z^2$ represents 0 over F if and only if $\left(\frac{-1,-1}{F}\right) = 1$. We claim that this is the case if $F = \mathbb{Q}_p$, p odd. One can treat specific cases via simple applications of quadratic reciprocity and Hensel's lemma.

Exercise 7.4. *Suppose $p \equiv 1 \bmod 4$. Show -1 is a square in $\mathbb{Z}_p$. Deduce $x^2 + y^2 + z^2$ represents* 0 *over $\mathbb{Q}_p$ (in fact $\mathbb{Z}_p$), i.e., $\left(\frac{-1,-1}{\mathbb{Q}_p}\right) = 1$.*

Exercise 7.5. *Suppose $p \equiv 3 \bmod 8$. Show -2 is a square in $\mathbb{Z}_p$. Deduce $x^2+y^2+z^2$ represents 0 over $\mathbb{Q}_p$ (in fact $\mathbb{Z}_p$), i.e., $\left(\frac{-1,-1}{\mathbb{Q}_p}\right)=1$.*

For the general case (well, really we only need it for $p \equiv 7 \bmod 8$ after the above exercises) we will appeal to the following formula.

Proposition 7.2.1. *Suppose p is odd, $a, b \in \mathbb{Q}_p$, and write $a = p^\alpha u$, $b = p^\beta v$, where u, v are units of $\mathbb{Z}_p$. Then*

$$\left(\frac{a,b}{\mathbb{Q}_p}\right) = (-1)^{\alpha\beta\frac{p-1}{2}} \left(\frac{u}{p}\right)^\beta \left(\frac{v}{p}\right)^\alpha.$$

(One extends the Legendre symbol $\left(\frac{\cdot}{p}\right)$ to $\mathbb{Z}_p^\times$ by putting $\left(\frac{a}{p}\right) = \left(\frac{a_0}{p}\right)$ for $a = a_0 + a_1 p + a_2 p^2 + \cdots$. However, we will only apply the above formula in the case where $a, b \in \mathbb{Z}$.)

Exercise 7.6. *Let p be odd. Compute $\left(\frac{-1,-1}{\mathbb{Q}_p}\right)$. Deduce that $x^2+y^2+z^2$ is universal over $\mathbb{Q}_p$.*

Lemma 7.2.2. *Let $\alpha \in \mathbb{Q}^\times$. Then α is a sum of 3 rational squares if and only if $\alpha > 0$ and $-\alpha \in \mathbb{Q}_2^{\times(2)}$.*

Proof. By Hasse–Minkowski, α is represented by $Q = x^2+y^2+z^2$ over $\mathbb{Q}$ if and only if α is represented by Q over each $\mathbb{Q}_p$ and $\mathbb{R}$. The representation condition over $\mathbb{R}$ is equivalent to $\alpha > 0$. By the above exercise, we know $x^2+y^2+z^2$ represents all α in $\mathbb{Q}_p$ for p odd, so it suffices to show $x^2+y^2+z^2$ represents α in $\mathbb{Q}_2$ if and only if $-\alpha \notin \mathbb{Q}_2^{\times(2)}$.

By an earlier exercise $x^2+y^2+z^2$ represents α in $\mathbb{Q}_2$ if and only if $x^2+y^2+z^2-\alpha w^2$ represents 0. One can show a rank 4 quadratic form $a_1x^2+a_2y^2+a_3z^2+a_4w^2$ over $\mathbb{Q}_p$ does *not* represent 0 if and only if the discriminant is a square and the Hasse symbol $\epsilon = \prod_{i<j}\left(\frac{a_i,a_j}{\mathbb{Q}_p}\right) = -\left(\frac{-1,-1}{\mathbb{Q}_p}\right)$. When $p=2$, we have $\left(\frac{-1,-1}{\mathbb{Q}_p}\right) = -1$ so this Hasse symbol condition holds if the discriminant $a_1a_2a_3a_4$ is a square, which in our case is just $-\alpha$. This proves the lemma. □

To pass to representations over $\mathbb{Z}$, we need the following.

Lemma 7.2.3. (Davenport–Cassels) *Let Q be a positive definite quadratic form of rank n over $\mathbb{Q}$ given by a symmetric matrix $A = (a_{ij}) \in M_n(\mathbb{Z})$. Suppose*

(DCH) for all $x \in \mathbb{Q}^n$, there is a $y \in \mathbb{Z}^n$ such that $Q(x-y) < 1$.

Then if Q represents an integer m over $\mathbb{Q}$, it does over $\mathbb{Z}$.

As in the binary case, positive definite means $Q(x) \geq 0$ with equality only if $x = 0 \in \mathbb{Q}^n$.

Proof. Write $\langle u, v\rangle = u^T A v$ for $u, v \in \mathbb{Q}^n$, so that $\langle v, v\rangle = Q(v)$.

Suppose $Q(v) = \langle v, v\rangle = m$ where $v \in \mathbb{Q}^n$. Multiplying through by denominators in $v = (v_1, \ldots, v_n)$, there is a multiple $x = tv \in \mathbb{Z}^n$ of v (for some $t \in \mathbb{Z}$) such that $Q(x) = Q(tv) = t^2 m$. Choose v and t such that t is minimal. We want to show $t = 1$.

(DCH) tells us there is a $y \in \mathbb{Z}^n$ such that $z = \frac{x}{t} - y \in \mathbb{Q}^n$ satisfies $Q(z) = \langle z, z\rangle < 1$. If $\langle z, z\rangle = 0$, then $z = 0$ (since Q is positive definite), so $\frac{x}{t} = y \in \mathbb{Z}^n$ and $t = 1$.

Now suppose $\langle z, z\rangle \neq 0$. Set

$$a = \langle y, y\rangle - m,\ b = 2(mt - \langle x, y\rangle),\ t' = at + b,\ x' = ax + by.$$

Then $a, b, t' \in \mathbb{Z}$, and it is easy to compute that $\langle x', x'\rangle = mt'^2$ and $tt' = t^2\langle z, z\rangle$. Consequently $t' = t\langle z, z\rangle < t$, contradicting the minimality of t. □

Theorem 7.2.4. (Gauss) *A positive integer n is a sum of 3 squares if and only if $n \neq 4^j(8k+7)$.*

Proof. Let $x = (x_1, x_2, x_3) \in \mathbb{Q}^3$. Choose $y = (y_1, y_2, y_3) \in \mathbb{Z}^3$ such that $|x_i - y_i| \leq \frac{1}{2}$. Then

$$Q(x-y) = (x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y^3)^2 \leq \frac{1}{4} + \frac{1}{4} + \frac{1}{4} < 1,$$

i.e., the form $x^2+y^2+z^2$ satisfies (DCH). By the Davenport–Cassels Lemma and the previous lemma, n is a sum of 3 squares if and only if $-n \in \mathbb{Q}_2^{\times(2)}$. Since $-n \in \mathbb{Z} \subseteq \mathbb{Z}_2$, this is equivalent to $-n$ is a square in $\mathbb{Z}_2$ (since $x^2 = -n$ in $\mathbb{Q}_2$ implies $|x^2|_2 = |-n|_2 \leq 1$ which implies $|x|_2 \leq 1$).

Write

$$-n = 2^e(1 + a_1 \cdot 2 + a_2 \cdot 2^2 + \cdots) \in \mathbb{Z}_2.$$

If n is a square e must be even, and then n is a square in $\mathbb{Z}_2$ if and only if

$$\frac{-n}{2^e} = 1 + a_1 \cdot 2 + a_2 \cdot 2^2 + \cdots$$

is. Using a slight generalization of Hensel's lemma, we see $a \in \mathbb{Z}_2^\times$ is a square if and only if it is mod 8, i.e., if and only if $a \equiv 1 \bmod 8$. Hence $-n$ is a square in $\mathbb{Q}_p$ if and only if $-n = 4^j(8m+1)$, i.e., if and only if $n = 4^j(8k+7)$. □

Corollary 7.2.5. (Lagrange) *Every positive integer n is a sum of 4 squares.*

Proof. If $n \neq 4^j(8k+7)$, then it is a sum of 4 squares since it is a sum of 3 squares. If $n = 4^j(8k+7)$ then $m = 8k+6$ is the sum of 3 squares so $8k+7 = m+1^2$ is the sum of 4 squares, whence n is also. □

We remark that we can't use the Davenport–Cassels Lemma for sums of 4 squares because (DCH) fails.

Corollary 7.2.6. (Gauss) *Ever positive integer n is a sum of 3 triangular numbers.*

(Recall a triangular number is one of the form $\frac{m(m+1)}{2}$.)

Proof. Applying the 3 squares theorem to $8n+3$, we see $8n+3 = x^2+y^2+z^2$ for some $x, y, z \in \mathbb{Z}$. But since the only squares mod 8 are $0, 1, 4$, we must have $x^2 \equiv y^2 \equiv z^2 \equiv 1 \bmod 8$, so x, y and z are odd. Write $x = 2a+1$, $y = 2b+1$, $z = 2c+1$. Then

$$\frac{a(a+1)}{2} + \frac{b(b+1)}{2} + \frac{c(c+1)}{2} = \frac{1}{8}\left((2a+1)^2 + (2b+1)^2 + (2c+1)^2 - 3\right) = \frac{1}{8}(8n+3-3) = n.$$

□

7.3 Siegel's mass formula

Here we give a brief summary of Siegel's mass formula, following [Iwaniec].

Let Q be a positive definite quadratic form over $\mathbb{Z}$ of rank r. The **genus** of Q is the set forms over $\mathbb{Z}$ which are equivalent to Q over each $\mathbb{Q}_p$ and $\mathbb{R}$. The group of **automorphs** $\mathrm{Aut}(Q)$ of Q is the set of $\sigma \in \mathrm{GL}_r(\mathbb{Z})$ such that $\sigma^T A \sigma = A$, where A is the symmetric matrix associated to Q. We say solutions $Q(x) = n$ and $Q(y) = n$ are equivalent if $y = \sigma x$ for some automorph σ of Q.

The number of automorphs $\mathrm{Aut}(Q)$ in general can be different for different forms in the same genus. Let $gen(Q)$ denote the set of equivalence classes of forms in the genus of Q. The **genus mass** of Q is

$$w(Q) = \frac{1}{|\mathrm{Aut}(Q)| m(gen(Q))}$$

where

$$m(gen(Q)) = \sum_{Q_j \in gen(Q)} \frac{1}{|\mathrm{Aut}(Q_j)|}.$$

The total mass is

$$\sum_{Q_j \in gen(Q)} w(Q_j) = 1.$$

Then the number of ways n can be represented by some form in the genus of Q is

$$r_{gen(Q)}(n) = \sum_{Q_j \in gen(Q)} w(Q_j) r_{Q_j}(n).$$

Siegel's mass formula then states

$$r_{gen(Q)}(n) = \prod_p \delta_p(n, Q) \cdot \delta_\infty(n, Q)$$

where $\delta_p(n, Q)$ is the "density" of solutions $Q(x) = n$ in $\mathbb{Z}_p^r$. When $r = 2$ this is essentially Dirichlet's mass formula.

So as in the binary case, when we have one class per genus (e.g., for sums of 3 or 4 squares), one knows the individual $r_Q(n)$'s. But this only happens finitely often, and in general it is hard to separate out the information about individual forms Q.

To attempt to do this, one approach is to associate to Q a **modular form**

$$\Theta_Q(z) = \sum_{n \geq 0} r_Q(n) e^{2\pi i n z},$$

which is a meromorphic function. Notice the $r_Q(n)$'s are Fourier coefficients for Θ_Q. Consequently, one can apply analytic methods to study the $r_Q(n)$'s and obtain beautiful formulas in many cases. Iwaniec uses analytic number theory to show an *asymptotic* formula for $r_Q(n)$ (as $n \to \infty$) for *individual* Q's.

We will not introduce modular forms or discuss other results in this direction here, but the study of quadratic forms and modular forms is a rich area, and there are many interesting open questions still out there.