Chapter 10
What about geometry?

In this chapter we present a geometric theorem of Minkowski, and use it to prove Theorem 9.8. We start with the basic theory of lattices in $\mathbb{R}^n$, discuss the volume of a lattice, and explain the connection of the volume of a lattice to the determinants of certain matrices. We then prove two foundational results of Minkowski, Proposition 10.8 and Theorem 10.10. The remainder of the chapter is devoted to studying sums of squares using the results of Minkowski just mentioned. The Two and Four Square Theorems are relatively easy to prove using Theorem 10.10, but the Three Square Theorem is hard. The proof of the Three Square Theorem occupies §10.5. In the Notes, we discuss Waring's Problem, introduce the functions $g(k)$ and $G(k)$, and explain some recent results obtained using the Circle Method (we also include an explanation of the Circle Method). At the end of the Notes, we say a few words about geometry of numbers.

10.1 Lattices in $\mathbb{R}^n$

Definition 10.1. Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis of $\mathbb{R}^n$, i.e., a set of n $\mathbb{R}$-linearly independent vectors in $\mathbb{R}^n$. The lattice generated by $\mathscr{B}$, denoted by $\Lambda_{\mathscr{B}}$, is the set of all linear combinations

$$c_1 v_1 + \cdots + c_n v_n$$

with $c_i \in \mathbb{Z}$. We define the *fundamental parallelogram* $\mathscr{P}_{\mathscr{B}}$ by

$$\mathscr{P}_{\mathscr{B}} = \{\alpha_1 v_1 + \cdots + \alpha_n v_n \mid 0 \leq \alpha_1, \ldots, \alpha_n < 1\}.$$

We define Vol $\mathscr{P}_{\mathscr{B}}$ to be the n-dimensional volume of the fundamental parallelogram $\mathscr{P}_{\mathscr{B}}$.

In Figure 10.1, $n = 2$, $\mathscr{B} = \{(2, 1), (1, 3)\}$, and the marked points are the elements of the lattice $\Lambda_{\mathscr{B}}$. Here, the fundamental parallelogram is painted yellow. In this case, Vol $\Lambda_{\mathscr{B}} = 5$.

R. Takloo-Bighash, *A Pythagorean Introduction to Number Theory*,
Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-02604-2_10

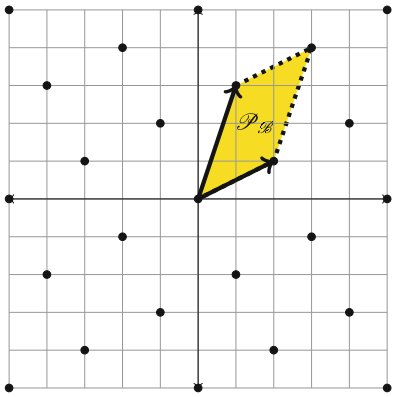

Fig. 10.1 A lattice in $\mathbb{R}^2$. The fundamental parallelogram is painted yellow

Note that the set $\Lambda_{\mathscr{B}}$ does not uniquely identify the basis $\mathscr{B}$. In fact it is easy to construct distinct bases $\mathscr{B}$ and $\mathscr{B}'$ of $\mathbb{R}^n$ such that

$$\Lambda_{\mathscr{B}} = \Lambda_{\mathscr{B}'};$$

see Exercise 10.1.

Definition 10.2. By a lattice in $\mathbb{R}^n$, we understand a set of the form $\Lambda_{\mathscr{B}}$ for some basis $\mathscr{B}$ of $\mathbb{R}^n$.

The quintessential example of a lattice in $\mathbb{R}^n$ is $\mathbb{Z}^n$ built from the standard basis:

$$e_1 = (1, 0, \ldots, 0),$$

$$e_2 = (0, 1, \ldots, 0),$$

$$\ldots$$

$$e_n = (0, 0, \ldots, 1).$$

The fundamental parallelogram associated to this basis is the unit cube in $\mathbb{R}^n$ whose volume is 1.

Proposition 10.3. *Suppose*

$$v_1 = (a_{11}, a_{12}, \ldots, a_{1n}),$$

$$v_2 = (a_{21}, a_{22}, \ldots, a_{2n}),$$

$$\ldots$$

$$v_n = (a_{n1}, a_{n2}, \ldots, a_{nn})$$

are n linearly independent vectors in $\mathbb{R}^n$. *Set* $\mathscr{B} = \{v_1, \ldots, v_n\}$. *Then*

$$\text{Vol}\,\mathscr{P}_{\mathscr{B}} = |\det(a_{ij})_{ij}| \neq 0$$

Proof. It is well known, e.g., [25, Ch. 6, §9], that the determinant $\det(a_{ij})_{ij}$ is non-zero if and only if the vectors $v_1, \ldots, v_n$ are linearly independent. For the volume statement, see Exercise 10.2. □

Example 10.4. Define a set $\Lambda \subset \mathbb{Z}^2$ as follows:

$$\Lambda = \{(x, y) \in \mathbb{Z}^2 \mid x \equiv y \bmod 2\}.$$

Let $v_1 = (2, 0)$, $v_2 = (1, 1)$, and set $\mathscr{B} = \{v_1, v_2\}$. We will show that $\Lambda = \Lambda_{\mathscr{B}}$. It is clear that $v_1, v_2 \in \Lambda$, and consequently, $\Lambda_{\mathscr{B}} \subset \Lambda$. Now we show the opposite inclusion. Observe that $2v_2 - v_1 = (2, 2) - (2, 0) = (0, 2)$. Now suppose $(x, y) \in \Lambda$. Since $x \equiv y \bmod 2$, there are two possibilities:

- x, y are even. In this case, $(x, y) = (2k, 2l)$ for integers k, l. Then

$$(x, y) = k(2, 0) + l(0, 2) = kv_1 + l(2v_2 - v_1) = (k - l)v_1 + 2lv_2 \in \Lambda_{\mathscr{B}};$$

- x, y are odd. In this case, $(x, y) = (2k + 1, 2l + 1)$ for integers k, l. Then

$$(x, y) = k(2, 0) + l(0, 2) + (1, 1) = (k - l)v_1 + (2l + 1)v_2 \in \Lambda_{\mathscr{B}}.$$

The fundamental domain $\mathscr{P}_{\mathscr{B}}$ is painted yellow in Figure 10.2. Proposition 10.3 shows that

$$\text{Vol}\,\mathscr{P}_{\mathscr{B}} = \left|\det \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\right| = 2.$$

Another relevant basis here is $\mathscr{B}' = \{v_1', v_2\}$ with $v_1' = (1, 3)$ and v_2 as above. The associated fundamental domain $\mathscr{P}_{\mathscr{B}'}$ is painted green in Figure 10.2. One easily

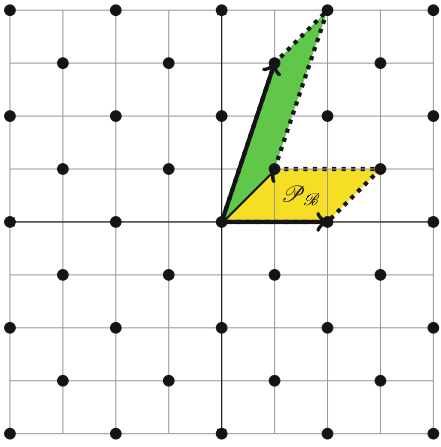

Fig. 10.2 The regions painted yellow and green are both fundamental domains

checks that $\Lambda = \Lambda_{\mathscr{B}'}$. Then we have

$$\operatorname{Vol} \mathscr{P}_{\mathscr{B}'} = \left|\det \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}\right| = |-2| = 2.$$

Even though the bases $\mathscr{B}$ and $\mathscr{B}'$ are different, the fundamental parallelograms $\mathscr{P}_{\mathscr{B}}$ and $\mathscr{P}_{\mathscr{B}'}$ have the same volume.

In general, for a lattice $\Lambda \subset \mathbb{R}^n$ there are infinitely many bases $\mathscr{B}$ such that $\Lambda = \Lambda_{\mathscr{B}}$. We will see in Exercise 10.3 that even though the set $\mathscr{P}_{\mathscr{B}}$ depends on the choice of $\mathscr{B}$, its volume, $\operatorname{Vol} \mathscr{P}_{\mathscr{B}}$, is independent of it, and that it depends only on the lattice Λ. This statement inspires the following definition:

Definition 10.5. If $\Lambda \subset \mathbb{R}^n$ is a lattice, then we define $\operatorname{Vol} \Lambda$ to be $\operatorname{Vol} \mathscr{P}_{\mathscr{B}}$ for any basis $\mathscr{B}$ such that $\Lambda = \Lambda_{\mathscr{B}}$.

10.2 Minkowski's Theorem

Let's start with a question:

Question 10.6. Suppose $\Lambda \subset \mathbb{R}^n$ is a lattice and let $S \subset \mathbb{R}^n$ be some subset . Under what conditions on S and Λ does S contain a non-zero point of Λ?

It is impossible to give exact necessary and sufficient conditions in this generality. In this section we state and prove an important theorem of Minkowski from 1896, Theorem 10.10, that gives necessary conditions for the existence of a point as asked in the question in some fairly narrow special cases. The surprising thing is that this theorem has some powerful applications in number theory. Our discussion of Minkowski's Theorem, while not particularly complicated, is, unfortunately, fairly abstract. It is only in the later parts of this chapter, starting with §10.3, that the relevance of what we do in this section to our concrete problems becomes clear.

First some preparation. If x is a vector in $\mathbb{R}^n$ and $S \subset \mathbb{R}^n$, we define

$$x + S = \{x + s \mid s \in S\}.$$

Hence $x + S$ is obtained from shifting the whole set S by the vector x. For example, if S is the disk of radius r centered at the origin in $\mathbb{R}^2$, $x + S$ will be the disk of radius r centered at x.

Lemma 10.7. *Let Λ be a lattice in $\mathbb{R}^n$, and let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis of $\mathbb{R}^n$ such that $\Lambda = \Lambda_{\mathscr{B}}$. Then,*

$$\mathbb{R}^n = \bigcup_{\lambda \in \Lambda} (\lambda + \mathscr{P}_{\mathscr{B}}), \tag{10.1}$$

a disjoint union.

Proof. Let $v \in R^n$. Since $\mathscr{B}$ is a basis, we can write

$$v = r_1 v_1 + \ldots r_n v_n,$$

for $r_i \in \mathbb{R}$. Next, for each i, write $r_i = [r_i] + \{r_i\}$. We obtain

$$v = \sum_{i=1}^{n} [r_i] v_i + \sum_{i=1}^{n} \{r_i\} v_i.$$

It is clear that $\sum_{i=1}^{n} [r_i] v_i \in \Lambda$ and $\sum_{i=1}^{n} \{r_i\} v_i \in \mathscr{P}_{\mathscr{B}}$. Now we show the union in (10.1) is disjoint. Suppose for vectors $\lambda, \lambda' \in \Lambda$, we have

$$(\lambda + \mathscr{P}_{\mathscr{B}}) \cap (\lambda' + \mathscr{P}_{\mathscr{B}}) \neq \varnothing. \tag{10.2}$$

Write

$$\lambda = \sum_{i=1}^{n} c_i v_i, \quad \lambda' = \sum_{i=1}^{n} c_i' v_i$$

for integers $c_1, c_1', \ldots, c_n, c_n'$. Equation (10.2) means that there are real numbers $\alpha_1, \alpha_1', \ldots, \alpha_n, \alpha_n'$ with $0 \leq \alpha_i, \alpha_i' < 1$ for each i such that

$$\sum_{i=1}^{n} c_i v_i + \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} c_i' v_i + \sum_{i=1}^{n} \alpha_i' v_i.$$

Consequently,

$$\sum_{i=1}^{n} (c_i + \alpha_i) v_i = \sum_{i=1}^{n} (c_i' + \alpha_i') v_i.$$

Since $\mathscr{B}$ is a basis of $\mathbb{R}^n$, this last identity implies that for all i we have

$$c_i + \alpha_i = c_i' + \alpha_i',$$

from which it immediately follows that $c_i = c_i'$ for all i. Hence, $\lambda = \lambda'$, and we are done. □

If we consider the example considered earlier where $n = 2$, $\mathscr{B} = \{(2, 1), (1, 3)\}$, we get the partition of $\mathbb{R}^2$ as a union of parallelograms as in Figure 10.3. (Care is needed about the boundary of each parallelogram!)

The following simple proposition is fundamental:

Proposition 10.8 (Minkowski). *Let Λ be a lattice in $\mathbb{R}^n$. Suppose U is an open set in $\mathbb{R}^n$ such that* Vol $U >$ Vol Λ. *Then there are distinct vectors $u_1, u_2 \in U$ such that $u_1 - u_2 \in \Lambda$.*

Proof. The starting point is Equation (10.1). Intersecting with the open set U gives

$$U = \bigcup_{\lambda \in \Lambda} \{(\lambda + \mathscr{P}_{\mathscr{B}}) \cap U\},$$

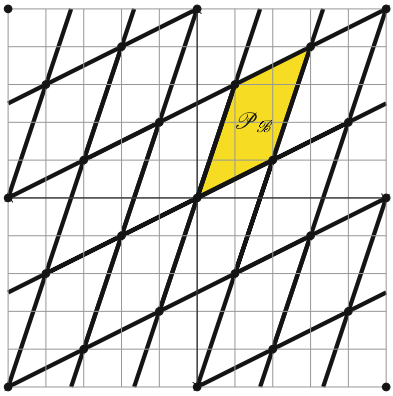

Fig. 10.3 The partition of $\mathbb{R}^2$ as a union of the translates of the fundamental domain as in Lemma 10.7

a disjoint union. Now we consider the volume of the set U. Since the right-hand side of the above equation is a disjoint union we have

$$\operatorname{Vol} U = \sum_{\lambda \in \Lambda} \operatorname{Vol} \{(\lambda + \mathscr{P}_{\mathscr{B}}) \cap U\}. \tag{10.3}$$

Next, since volume in $\mathbb{R}^n$ is translation invariant, we have

$$\operatorname{Vol} \{(\lambda + \mathscr{P}_{\mathscr{B}}) \cap U\} = \operatorname{Vol} \{\mathscr{P}_{\mathscr{B}} \cap (-\lambda + U)\}.$$

Denote the set $\mathscr{P}_{\mathscr{B}} \cap (-\lambda + U)$ by $\mathscr{P}_\lambda$. Note that $\mathscr{P}_\lambda \subset \mathscr{P}_{\mathscr{B}}$. Since by assumption $\operatorname{Vol} U > \operatorname{Vol} \Lambda = \operatorname{Vol} \mathscr{P}_{\mathscr{B}}$, Equation (10.3) gives

$$\operatorname{Vol} \mathscr{P}_{\mathscr{B}} < \sum_{\lambda \in \Lambda} \operatorname{Vol} \mathscr{P}_\lambda.$$

This equation implies that there are distinct elements $\lambda_1, \lambda_2 \in \Lambda$ such that $\mathscr{P}_{\lambda_1} \cap \mathscr{P}_{\lambda_2} \neq \varnothing$. This means that there are $u_1, u_2 \in U$ such that $-\lambda_1 + u_1 = -\lambda_2 + u_2$ with. This last equation implies the statement of the proposition with $\lambda = \lambda_1 - \lambda_2$. □

Before we state the main theorem of this chapter we need a couple of definitions.

Definition 10.9. Let S be a set in $\mathbb{R}^n$.

- We call S *symmetric* if $x \in S$ implies $-x \in S$.
- We call S *convex* if for $x, y \in S$ and $0 < \alpha < 1$ we have

$$\alpha x + (1 - \alpha) y \in S,$$

  i.e., if $x, y \in S$, the line segment connecting x and y lies in S.

The quintessential example of a convex symmetric set in $\mathbb{R}^2$ is a filled ellipse of the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} < 1.$$

A particularly important example that makes an appearance later in this chapter is a disk

$$x^2 + y^2 < r^2.$$

The set

$$1 < x^2 + y^2 < 4$$

is symmetric but not convex, and the disk

$$(x-2)^2 + y^2 < 3$$

is convex but not symmetric.

We can now state and prove the following important theorem:

Theorem 10.10 (Minkowski). *Let Λ be a lattice in $\mathbb{R}^n$. Suppose S is a convex symmetric open set in $\mathbb{R}^n$ such that* $\operatorname{Vol} S > 2^n \operatorname{Vol} \Lambda$. *Then there is a non-zero vector in the intersection $\Lambda \cap S$.*

Proof. Let $S' = (1/2)S$ be the scaled down version of S. Then S' is open and $\operatorname{Vol} S' > \operatorname{Vol} \Lambda$. By Proposition 10.8, there are distinct elements $u_1, u_2 \in S$ such that

$$u := \frac{u_1}{2} - \frac{u_2}{2} \in \Lambda - \{0\}.$$

Since $u_2 \in S$ and S is symmetric, $-u_2 \in S$. Also $(u_1 - u_2)/2 \in S$ as S is convex and $(u_1 - u_2)/2$ is the middle point of the line segment connecting $u_1, -u_2 \in S$. The theorem is proved. □

Example 10.11. A lattice Λ is called *unimodular* if $\operatorname{Vol} \Lambda = 1$. Let Λ be a unimodular lattice in $\mathbb{R}^2$. Define the set $S_r \subset \mathbb{R}^2$ to be the disk

$$x^2 + y^2 < r^2.$$

For each $r > 0$, S_r is convex, symmetric, and open, and has area πr^2. If $r > 2/\sqrt{\pi} = 1.12837916709551...$, then $\operatorname{Vol} S_r > 4 = 2^2 \cdot \operatorname{Vol} \Lambda$. Theorem 10.10 implies that the set $A_r := S_r \cap \Lambda - \{(0,0)\}$ is a finite non-empty set. Basic properties of compact sets, e.g., [41, Theorem 2.36], imply that

$$A := \bigcap_{r > 2/\sqrt{\pi}} (S_r \cap \Lambda - \{(0,0)\}) = \overline{S_{2/\sqrt{\pi}}} \cap \Lambda - \{(0,0)\}$$

is non-empty. This means that every unimodular lattice $\Lambda \subset \mathbb{R}^2$ contains at least one non-zero vector v whose length is less than or equal to $2/\sqrt{\pi}$. This result can be generalized to every dimension; see Exercise 10.13.

Despite its innocent abstract appearance, Theorem 10.10 is a powerful result with many applications. In the remainder of this chapter we give three applications of this theorem to questions involving sums of squares.

10.3 Sums of two squares

In our first application we give a second proof of Theorem 5.7.

The second proof of Theorem 5.7. Recall that the non-trivial part of Theorem 5.7 is the statement that every prime p of the form $4k + 1$ is a sum of two squares. As we observed in our proof of Theorem 5.7, it suffices to find a pair of integers (u, v) with the following properties:

1. $u^2 + v^2 < 2p$;
2. $p|u^2 + v^2$;
3. $(u, v) \neq (0, 0)$.

Consider the set

$$S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 2p\}.$$

The set S is convex, symmetric, and open. Also, $\mathrm{Vol}\, S = 2\pi p$. In order to apply Theorem 10.10, we need to find a lattice Λ such that

(i) $4\mathrm{Vol}\, \Lambda < 2\pi p$;
(ii) for all $(a, b) \in \Lambda$, $p \mid a^2 + b^2$.

Note that since p is of the form $4k+1$, by (6.3), there is z such that $z^2+1 \equiv 0 \bmod p$. Clearly, the sensible thing to do is to use z to construct the lattice. Consider the vectors:

$$v_1 = (p, 0), \quad v_2 = (z, 1).$$

We have

$$\det \begin{pmatrix} p & 0 \\ z & 1 \end{pmatrix} = p \neq 0,$$

hence the vectors v_1, v_2 are linear independent. Let Λ be the lattice generated by v_1, v_2. By Proposition 10.3, $\mathrm{Vol}\, \Lambda = p$. Since $4p < 2\pi p$, condition (ii) is satisfied. Next, we verify (i). A typical vector in Λ can be written as $c_1 v_1 + c_2 v_2$ with $c_1, c_2 \in \mathbb{Z}$. We compute the coordinates of the vector as

$$c_1 v_1 + c_2 v_2 = (c_1 p + c_2 z, c_2).$$

We compute the sum of the squares of the coordinates to obtain

$$(c_1 p + c_2 z)^2 + c_2^2 \equiv c_2^2 (z^2 + 1) \equiv 0 \mod p.$$

This finishes the proof. □

10.4 Sums of four squares

In this section and the next we give a proof of Theorem 9.8. Our goal in this section is to show that every natural number is a sum of four squares following an idea of Davenport [71]. Davenport gives credit to Hermite (1830) for this proof, though Hermite did not have Theorem 10.10 at his disposal, so he had to use other methods. Davenport notes this is a very non-trivial result. According to [71], Euler tried to prove the result unsuccessfully many times between 1730 and 1750, see [90]. This is a testimony to the effectiveness of Minkowski's innocuous looking theorem. We will present another proof of the Four Squares Theorem in Chapter 11 where we will use quaternions.

We start with an identity discovered by Euler [90]. This identity is the analogue of Lemma 5.3 in this setting.

Lemma 10.12 (Euler's identity). *For all complex numbers* a, b, c, d, e, f, g, h,

$$(a^2+b^2+c^2+d^2)(e^2+f^2+g^2+h^2) =$$

$$(ae-bf-cg-dh)^2+(af+be+ch-dg)^2$$

$$+(ag-bh+ce+df)^2+(ah+bg-cf+de)^2$$

Proof. See Exercise 10.14 or Lemma 11.4 for a conceptual proof. □

Lemma 10.12 implies that in order to prove that every natural number is the sum of four squares, it suffices to prove that every prime number is a sum of four squares. Since

$$2 = 1^2+1^2+0^2+0^2,$$

we just need to prove the assertion for an odd prime p. As in the case of the Two Squares Theorem, we need to show that there are integers u, v, w, t such that

1. $u^2+v^2+w^2+t^2 < 2p$;
2. $p \mid u^2+w^2+v^2+t^2$;
3. $(u, v, w, t) \neq (0, 0, 0, 0)$.

By Exercise 9.4 the volume of the set S in $\mathbb{R}^4$ defined by

$$S = \{(u, v, w, t) \in \mathbb{R}^4 \mid u^2+v^2+w^2+t^2 < 2p\}$$

is

$$\frac{\pi^2}{2}(\sqrt{2p})^4 = 2\pi^2 p^2.$$

Also, we note that the set S is convex, symmetric, and open. In order to apply Theorem 10.10 we need to construct a lattice Λ such that

(i) $16\text{Vol}\,\Lambda < \text{Vol}\,S$;
(ii) for all $(a, b, c, d) \in \Lambda$, $p \mid a^2+b^2+c^2+d^2$.

We need a lemma:

Lemma 10.13. *If p is an odd prime number, there are integers r, s such that*

$$r^2 + s^2 + 1 \equiv 0 \mod p.$$

Proof. We define functions f, g from $\mathbb{Z}$ to $\mathbb{Z}/p\mathbb{Z}$, the set of congruence classes modulo p, by setting $f(x) = x^2 \bmod p$ and $g(x) = -1 - x^2 \bmod p$. The assertion of the lemma is equivalent to the existence of integers r, s such that $f(r) = g(s)$. We claim

$$\#f(\mathbb{Z}) = \#g(\mathbb{Z}) = \frac{p+1}{2},$$

where here, for example, $f(\mathbb{Z})$ is the image of the function f and $\#f(\mathbb{Z})$ is the number of elements in the image. We will prove the statement involving f; the one for g follows similarly. The first point to note is that if $x \equiv y \mod p$, then $f(x) = f(y)$ as an element of $\mathbb{Z}/p\mathbb{Z}$. This means that we may in fact think of f as a function from $\mathbb{Z}/p\mathbb{Z}$ to $\mathbb{Z}/p\mathbb{Z}$. Next, $f(x) = f(y)$ if and only if $x \equiv \pm y \mod p$. Now, if $x \not\equiv 0 \mod p$, then $x \not\equiv -x \mod p$. This means that f is 2-to-1 for non-zero congruence classes. Since there are $p - 1$ non-zero congruence classes, there will be $(p-1)/2$ elements in the images of these classes. We also need to account for $f(0) = 0$. Consequently, the total number of elements in the image of f is $(p-1)/2 + 1 = (p+1)/2$. This finishes the proof of $\#f(\mathbb{Z}) = (p+1)/2$. As mentioned above, the proof of $\#g(\mathbb{Z}) = (p+1)/2$ is similar. Now, we notice

$$\#f(\mathbb{Z}) + \#g(\mathbb{Z}) = \frac{p+1}{2} + \frac{p+1}{2} = p + 1 > p,$$

hence by the Pigeon-Hole Principle, Theorem A.7, there has to be an overlap between the images of the functions f, g. □

Fix r, s as in the lemma, and consider the four vectors

$$v_1 = (p, 0, 0, 0), \quad v_2 = (0, p, 0, 0), \quad v_3 = (r, s, 1, 0), \quad v_4 = (s, -r, 0, 1).$$

We have

$$\det \begin{pmatrix} p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ r & s & 1 & 0 \\ s & -r & 0 & 1 \end{pmatrix} = p^2 \neq 0,$$

and consequently the vectors $\{v_1, v_2, v_3, v_4\}$ are linearly independent. If Λ is the lattice generated by these vectors, Proposition 10.3 implies that $\mathrm{Vol}\,\Lambda = p^2$. Note that since $2\pi^2 > 16$,

$$16\mathrm{Vol}\,\Lambda < \mathrm{Vol}\,S.$$

Now we can apply Theorem 10.10 to conclude that there is $(a, b, c, d) \in S \cap \Lambda$ such that $(a, b, c, d) \neq (0, 0, 0, 0)$. Next, we show that if $(a, b, c, d) \in \Lambda$, then

$$p \mid a^2 + b^2 + c^2 + d^2.$$

In order to see this, we write

$$(a, b, c, d) = c_1 v_1 + c_2 v_2 + c_3 v_3 + c_4 v_4$$

with $c_1, c_2, c_3, c_4 \in \mathbb{Z}$. Then

$$\begin{cases} a = c_1 p + c_3 r + c_4 s, \\ b = c_2 p + c_3 s - c_4 r, \\ c = c_3, \\ d = c_4. \end{cases}$$

Finally,

$$\begin{aligned} a^2 + b^2 + c^2 + d^2 &\equiv (c_1 p + c_3 r + c_4 s)^2 + (c_2 p + c_3 s - c_4 r)^2 + c_3^2 + c_4^2 \\ &\equiv (c_3 r + c_4 s)^2 + (c_3 s - c_4 r)^2 + c_3^2 + c_4^2 \\ &\equiv c_3^2 (r^2 + s^2 + 1) + c_4^2 (r^2 + s^2 + 1) \equiv 0 \mod p \end{aligned}$$

by the choices of r, s. This finishes the proof of the Four Square Theorem.

Remark 10.14. Davenport's original proof [71] differs slightly from the above argument. Davenport notes that if

$$m = x^2 + y^2 + z^2 + t^2,$$

then

$$2m = (x + y)^2 + (x - y)^2 + (z + t)^2 + (z - t)^2.$$

So it suffices to prove the theorem for odd m. So we assume that m is an odd natural number. There are integers r, s such that

$$r^2 + s^2 + 1 \equiv 0 \mod m.$$

Then consider the four vectors

$$v_1 = (m, 0, 0, 0), \quad v_2 = (0, m, 0, 0), \quad v_3 = (r, s, 1, 0), \quad v_4 = (s, -r, 0, 1).$$

and form the lattice Λ generated by them. The remainder of the argument is identical to what was presented above. Davenport's clever idea of reducing the general case to the odd m case should be compared to the division-by-$(1 + i)$ step in the proof of the Four Square Theorem presented in §11.3.

10.5 Sums of three squares

We now give a proof of the only remaining statement of Theorem 9.8 that a positive integer m is expressible as a sum of three squares if and only if m is not of the form $4^a(8n+7)$. We will give one more proof of this fact using the theory of quadratic forms in §12.4.

The fact that numbers of the form $4^a(8n+7)$ are not expressible as a sum of three squares is not hard; see Exercise 10.17; however, the fact that a number m not of the form $4^a(8n+7)$ is expressible as a sum of three squares is a much harder theorem. There are several proofs of this result available in literature. We will present a beautiful proof due to Dirichlet in Chapter 12 following the exposition of the classical text by Landau [31]. The remarkable proof we give in this chapter is due to Ankeny [61].

It is clear that we may assume that m is square-free. Following [61], we present the detailed proof for the case where $m \equiv 3 \bmod 8$ to illustrate the method, and refer the reader to the exercises for the remaining cases.

Suppose $m = p_1 \cdots p_r$ is a square-free integer such that $m \equiv 3 \bmod 8$.

Step 1. There is a prime number q such that

- For each i, $1 \le i \le r$,
$$\left(\frac{-2q}{p_i}\right) = +1;$$
- $q \equiv 1 \bmod 4$.

To see this, we note that each condition $(-2q/p_i) = +1$ means that q belongs to some congruence classes modulo p_i. The Chinese Remainder Theorem 2.24 then implies that there is a congruence condition of the form $q \equiv a \bmod 4m$ such that all of these conditions are satisfied. Dirichlet's Arithmetic Progression Theorem, Theorem 5.11 in Notes to Chapter 5, implies the existence of infinitely many primes q with this property.

Step 2. There is an odd integer b and an integer h such that

$$b^2 - 4hq = -m.$$

To see this, we first have to show

$$\left(\frac{-m}{q}\right) = +1.$$

In fact,

$$+1 = \prod_{i=1}^{r}\left(\frac{-2q}{p_i}\right) = \prod_{i=1}^{r}\left(\frac{-2}{p_i}\right)\left(\frac{q}{p_i}\right)$$
$$= \left(\frac{-2}{m}\right)\prod_{i=1}^{r}\left(\frac{q}{p_i}\right).$$

Here $(-2/m)$ is the Jacobi symbol of §7.2. By Quadratic Reciprocity,

$$\left(\frac{q}{p_i}\right) = (-1)^{\frac{q-1}{2}\cdot\frac{p_i-1}{2}}\left(\frac{p_i}{q}\right) = \left(\frac{p_i}{q}\right)$$

as $q \equiv 1 \bmod 4$. Hence,

$$+1 = \left(\frac{-2}{m}\right)\prod_i \left(\frac{p_i}{q}\right) = \left(\frac{-2}{m}\right)\left(\frac{m}{q}\right).$$

This means

$$\left(\frac{-2}{m}\right) = \left(\frac{m}{q}\right).$$

Next, since $q \equiv 1 \bmod 4$, $(-1/q) = +1$, we have,

$$\left(\frac{-m}{q}\right) = \left(\frac{-1}{q}\right)\left(\frac{m}{q}\right) = \left(\frac{m}{q}\right).$$

Putting these identities together gives

$$\left(\frac{-m}{q}\right) = \left(\frac{-2}{m}\right).$$

Next, by Theorem 7.3,

$$\left(\frac{-2}{m}\right) = \left(\frac{-1}{m}\right)\left(\frac{2}{m}\right) = (-1)^{\frac{m-1}{2}}(-1)^{\frac{m^2-1}{8}} = (-1).(-1) = +1,$$

as $m \equiv 3 \bmod 8$. We finally obtain

$$\left(\frac{-m}{q}\right) = +1. \tag{10.4}$$

This means there is an integer b such that $b^2 \equiv -m \bmod q$. By adding q to b if necessary, we assume b is odd. Consequently, there is an integer h_1 such that

$$b^2 - qh_1 = -m.$$

Since b and q are odd, and $m \equiv 3 \bmod 8$, viewing this equation modulo 4 gives $1 - h_1 \equiv +1 \bmod 4$, or $h_1 \equiv 0 \bmod 4$. Write $h_1 = 4h$ to get

$$b^2 - 4qh = -m$$

as claimed.

Step 3. There is an integer t such that

$$t^2 \equiv -1/(2q) \mod m.$$

In fact, by the choice of q, for each i, there is an integer s_i such that

$$s_i^2 \equiv -2q \mod p_i.$$

If we set $t_i \equiv s_i^{-1} \mod p_i$, then $t_i^2 \equiv -1/(2q) \mod p_i$. By the Chinese Remainder Theorem, there is a t modulo m such that for each i, $t \equiv t_i \mod p_i$. This means $t^2 \equiv t_i^2 \equiv -1/(2q) \mod p_i$. Consequently, as $m = p_1 \dots p_r$, $t^2 \equiv -1/(2q) \mod m$.

Step 4. Define

$$S = \{(u, v, w) \in \mathbb{R}^3 \mid u^2 + v^2 + w^2 < 2m\}.$$

Then S is an open ball in the three-dimensional space, and as such it is convex, symmetric, and open. The volume of S is

$$\frac{4}{3}\pi(2m)^{\frac{3}{2}}.$$

Step 5. We now define a lattice. Set

$$v_1 = (2tq, (2q)^{1/2}, 0), \quad v_2 = (tb, b/(2q)^{1/2}, m^{1/2}/(2q)^{1/2}), \quad v_3 = (m, 0, 0).$$

Since

$$\det \begin{pmatrix} tb & b/(2q)^{1/2} & m^{1/2}/(2q)^{1/2} \\ 2tq & (2q)^{1/2} & 0 \\ m & 0 & 0 \end{pmatrix} = -m^{3/2} \neq 0,$$

$\{v_1, v_2, v_3\}$ is a linearly independent set in $\mathbb{R}^3$. Let Λ be the lattice generated by these vectors. Then $\operatorname{Vol} \Lambda = m^{3/2}$.

Step 6. If $(u, v, w) \in \Lambda$, then v, w are not integers. However, we show that $u^2+v^2+w^2$ is an integer, and that

$$u^2 + v^2 + w^2 \equiv 0 \mod m.$$

We have for three integers x, y, z,

$$(u, v, w) = xv_1 + yv_2 + zv_3$$

$$= (2tqx + tby + mz, (2q)^{1/2}x + b/(2q)^{1/2}y, m^{1/2}/(2q)^{1/2}y).$$

Consequently,

$$u^2+v^2+w^2 = (2tqx+tby+mz)^2+((2q)^{1/2}x+b/(2q)^{1/2}y)^2+(m^{1/2}/(2q)^{1/2}y)^2$$

$$= (2tqx + tby + mz)^2 + \frac{1}{2q}(2qx + by)^2 + \frac{my^2}{2q} \tag{10.5}$$

$$= (2tqx + tby + mz)^2 + 2(qx^2 + bxy + hy^2). \tag{10.6}$$

This shows that $u^2 + v^2 + w^2$ is an integer. We now that it is divisible by m. From (10.5) we have

$$u^2 + v^2 + w^2 \equiv t^2(2qx + by)^2 + (2qx + by)^2/2q \equiv 0 \mod m$$

by the choice of t.

Step 7. Recall $\text{Vol}\,\Lambda = m^{2/3}$ and $\text{Vol}\,S = \frac{4}{3}\pi(2m)^{\frac{3}{2}}$. Since

$$\frac{4}{3}\pi(2m)^{\frac{3}{2}} > 8m^{2/3},$$

we see that $\text{Vol}\,S > 2^3\text{Vol}\,\Lambda$. Theorem 10.10 implies that there is a non-zero triple of integers (x_1, y_1, z_1) such that

$$(u_1, v_1, w_1) := x_1v_1 + y_1v_2 + z_1v_3 \in S.$$

Since $(u_1, v_1, w_1) \in S$, we have $u_1^2 + v_1^2 + w_1^2 < 2m$. Step 6 says $u_1^2 + v_1^2 + w_1^2$ is a non-zero integer, and that $m \mid u_1^2 + v_1^2 + w_1^2$. This means

$$u_1^2 + v_1^2 + w_1^2 = m. \tag{10.7}$$

Now let

$$R_1 = 2tqx + tby + mz, \quad v = qx^2 + bxy + hy^2. \tag{10.8}$$

The identity (10.6) combined with (10.7) shows

$$m = R_1^2 + 2v. \tag{10.9}$$

Step 8. Finally, we show that $2v$ is a sum of two squares, and this fact combined with Equation (10.9) finishes the proof of the theorem for $m \equiv 3 \bmod 8$.

It suffices to show that v is a sum of two squares. Indeed, $2 = 1^2 + 1^2$, and by Lemma 5.5 if v is a sum of two squares, $2v$ will be a sum of two squares.

To show that v is a sum of two squares, by Theorem 5.2, we need to show that if $p^{2k+1} \mid v$ but $p^{2k+2} \nmid v$, then $p \equiv 1 \bmod 4$.

There are two cases to consider: $p \nmid m$, and $p|m$. We treat each case separately.

If $p \nmid m$, then reducing Equation (10.9) modulo p implies $\left(\frac{m}{p}\right) = +1$. If $p = q$, then by Equation (10.4), $(-1/p) = +1$, and Lemma 6.7 implies $p \equiv 1 \bmod 4$.

Now suppose $p \neq q$. The definition of v from (10.8) shows

$$4qv = (2qx_1 + by_1)^2 + my_1^2.$$

This equation implies that p^{2k+1} divides an expression of the form $e^2 + mf^2$, but p^{2k+2} does not. Consequently, again,

$$\left(\frac{-m}{p}\right) = +1.$$

Again, as before, $(-1/p) = 1$, and $p \equiv 1 \bmod 4$. This settles the case where $p \nmid m$.

Now we consider the case where $p|m$. Recall that we have

$$R_1^2 + 2v = m.$$

This identity implies that $p|R_1$. We can now rewrite this equation in the following form

$$R_1^2 + \frac{1}{2q}((2qx_1 + by_1)^2 + my_1^2) = m,$$

and this identity implies $p|(2qx_1 + by_1)$. Since by assumption m is square-free, these statements show

$$\frac{1}{2q}\frac{m}{p}y_1^2 \equiv \frac{m}{p} \mod p,$$

or what is the same

$$y_1^2 \equiv 2q \mod p.$$

Consequently, $(2q/p) = +1$. Recall from Step 1 that since $p \mid m$, we have $(-2q/p) = +1$. Hence, $(-1/p) = +1$, and again we arrive at the conclusion that $p \equiv 1 \bmod 4$.

For the cases where $m \equiv 1, 2, 5, 6 \bmod 8$, see Exercise 10.19.

Exercises

10.1 Find bases $\mathscr{B}$ and $\mathscr{B}'$ of $\mathbb{R}^2$ which generate the same lattice but $\mathscr{P}_{\mathscr{B}} \neq \mathscr{P}_{\mathscr{B}'}$.

10.2 Prove Proposition 10.3.

10.3 Show that the volume of $\mathscr{P}_{\mathscr{B}}$ is independent of the choice of the basis of $\mathscr{B}$ that generates the lattice Λ.

10.4 Show that for $n > 3$ we have

$$\det((ij-1)^2)_{1\le i,j\le n} = 0.$$

10.5 Show that for all $n > 4$

$$\det((ij-1)^3)_{1\le i,j\le n} = 0.$$

10.6 Generalize the previous two problems by showing that for natural numbers n, k, if $n > k+1$, we have

$$\det((ij-1)^k)_{1\le i,j\le n} = 0.$$

10.7 Define a matrix $D = (a_{ij})_{1\le i,j\le n}$ by setting

$$a_{ij} = \begin{cases} 0 & i = j; \\ 1 & i < j; \\ -1 & i > j. \end{cases}$$

Show that

$$\det D = \begin{cases} 0 & n \text{ odd}; \\ 1 & \text{otherwise}. \end{cases}$$

10.8 Define a matrix $E = (b_{ij})_{1\le i,j\le n}$ by setting

$$b_{ij} = \begin{cases} 1 + x^2 & i = j; \\ x & |i - j| = 1; \\ 0 & \text{otherwise.} \end{cases}$$

Compute $\det E$. Hint: Let $D_n = \det E$. Show that for $n \geq 3$ we have $D_n - D_{n-1} = x^2(D_{n-1} - D_{n-2})$.

10.9 For three complex numbers α, β, γ, and $r \in \mathbb{N}$, set $\sigma_r = \alpha^r + \beta^r + \gamma^r$. Show that for $n \in \mathbb{N}$,

$$\det(\sigma_{n+i+j-2})_{1 \leq i,j \leq 3} = (\alpha\beta\gamma)^n \{(\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)\}^2.$$

10.10 Define a matrix $F_n = (c_{ij})_{1 \leq i,j \leq n}$ by setting $c_{ij} = 1 + \delta_{ij}x$, where δ_{ij} is Kronecker's delta. Compute $f_n(x) := \det F_n$ by showing that $f_n'(x) = nf_{n-1}(x)$.

10.11 Define a subset $\Lambda \subset \mathbb{Z}^2$ by setting

$$\Lambda = \{(x, y) \in \mathbb{Z}^2 \mid x + y \equiv 0 \bmod 3\}.$$

Show that Λ is a lattice by finding a basis $\mathscr{B}$ of $\mathbb{R}^2$ such that $\Lambda = \Lambda_{\mathscr{B}}$. Compute $\operatorname{Vol} \Lambda$.

10.12 Fix a prime p. Define a subset $\Lambda_{p,n} \subset \mathbb{Z}^n$ by setting

$$\Lambda_{p,n} = \left\{ (x_1, \ldots, x_n) \in \mathbb{Z}^n \mid \sum_{i=1}^{n} x_i \equiv 0 \bmod p \right\}.$$

Prove that $\Lambda_{p,n}$ is a lattice. Compute $\operatorname{Vol} \Lambda_{p,n}$.

10.13 Verify the details of the argument in Example 10.11. Generalize to all $\mathbb{R}^n$.

10.14 Prove Lemma 10.12 by direct computation.

10.15 Supply the details of Davenport's proof of the Four Square Theorem.

10.16 (✠) Write 4594043492117928 as a sum of four squares. In how many ways is it possible to do this?

10.17 Show that a number of the form $4^a(8n + 7)$ is not expressible as a sum of three squares.

10.18 (✠) Can you write 4594043492117928 as a sum of three squares? In how many ways?

10.19 Complete Ankeny's proof of the remaining cases of the Three Square Theorem, i.e., for the cases where $m \equiv 1, 2, 5, 6 \bmod 8$. Let q be prime, $(-q/p_j) = +1$ for all odd prime divisors of m, and $q \equiv 1 \bmod 4$, and if m is even, $m = 2m_1$, $(-2/q) = (-1)^{(m_1-1)/2}$, $t^2 \equiv -1/q \bmod p_j$, t odd, $b^2 - qh = -m$, and $v_1 = (tq, q^{1/2}, 0)$, $v_2 = (tb, b/q^{1/2}, m^{1/2}/q^{1/2})$, $v_3 = (m, 0, 0)$.

10.20 Show that every integer can be represented as a sum of five cubes of integers in infinitely many ways. Show that 3 can be written as a sum of four cubes not equal to 0 or 1 in infinitely many ways.

Notes

Waring's Problem

In 1770 Edward Waring asked whether for a natural number k, there is an integer s, depending on k, such that every natural number could be written as the sum of at most s natural numbers everyone of which is a k-th power. If the answer is yes, then the smallest possible value of s is denoted by $g(k)$. For example, in this chapter we showed that every natural number is the sum of at most four perfect squares. We also saw that there are many integers that are not sums of three squares. This means that $g(2) = 4$. David Hilbert showed in 1909 that the answer to Waring's question was yes. The first few values of $g(k)$ are as follows: $g(2) = 4$, $g(3) = 9$, $g(4) = 19$, $g(5) = 37$, $g(6) = 73$, etc. The sequence of integers $g(k)$ appears as sequence A002804 in the *Online Encyclopedia of Integer Sequences* available at

https://oeis.org/A002804

The following conjecture dates back to the 19th century:

Conjecture 10.15 (The Ideal Waring's Theorem). For all k we have

$$g(k) = 2^k + [(3/2)^k] - 2,$$

where in this formula $[x]$ denotes the integer part of a real number x.

It is a theorem of L. E. Dickson and S. S. Pillai from 1936 that this formula for $g(k)$ holds if

$$2^k\{(3/2)^k\} + [(3/2)^k] \leq 2^k. \tag{10.10}$$

This last inequality is known to be true for $k \leq 471,600,000$, and for k large enough by a result of K. Mahler from 1957. Equation (10.10) is expected to hold for all k; see [106]. In fact, it is a result of David, Waldschmidt, and Laishram [82, 107] that an explicit form of the *abc* Conjecture implies the Ideal Waring's Theorem. For the statement of the *abc* Conjecture and its explicit form see Notes to Chapter 3.

A related sequence which is considerably more difficult to study is the sequence $G(k)$ defined as the smallest positive integer s such that every *sufficiently large* positive integer can be written as the sum of s k-th powers. It is clear that $G(k) \leq g(k)$. One could, however, imagine that there may exist some rogue integers early on that require a lot of k-powers, but past a certain point the situation would stabilize. The only values of $G(k)$ that are currently known are $G(2) = 4$ and $G(4) = 16$ obtained in 1939 by Davenport. It appears that the best available upper bound for $G(k)$ is provided by Trevor Wooley in 1995:

$$G(k) \leq k(\log k + \log\log k + 2 + O(\log\log k/\log k)).$$

This should be compared with the conjectured formula for $g(k)$ mentioned earlier. The conjectured value of $g(k)$ grows exponentially with k, whereas the inequality proved by Wooley shows that $G(k)$ grows essentially in a linear fashion. (Professor Ram Murty often jokes that analytic number theorists say "log, log, log" when they drown.)

A powerful method that has been employed to prove many of the results related to Waring's Problem, and other additive questions in number theory, is the *Circle Method* originally invented by Hardy and Ramanujan around 1916. The idea is to define a function

$$f_k(x) = \sum_{n=0}^{\infty} e^{2\pi i n^k x}$$

on the interval $[0, 1]$. Fix a natural number s, and suppose we wish to show that every natural number is the sum of s k-th powers. We have

$$f_k(x)^s = \sum_{n_1} \sum_{n_2} \cdots \sum_{n_s} e^{2\pi i x \sum_{j=1}^{s} n_j^k} = \sum_{n=0}^{\infty} R_s(n) e^{2\pi i n x}$$

with $R_s(n)$ being the number of representations of n as a sum of s k-th powers. Theorem A.3 now implies that for each l,

$$\int_0^1 f_k(x)^s e^{-2\pi i l x}\, dx = \sum_{n=0}^{\infty} R_s(n) \int_0^{\infty} e^{2\pi i (k-l)x}\, dx = \sum_{n=0}^{\infty} R_s(n)\delta_{nl} = R_s(l),$$

with δ_{nl} being Kronecker's delta. So in order to show that $R_s(l) \neq 0$ to gain information about $g(k)$, or $R_s(l) \neq 0$ for l large enough to gain information about $G(k)$ one needs to show that

$$\int_0^1 f_k(x)^s e^{-2\pi i l x}\, dx \neq 0.$$

In order to see that this integral is non-zero the idea is to concentrate on those x's for which the value of $f(x)$ is large. Note that if x is a rational number, then $e^{2\pi i k^2 x} = 1$ for infinitely many k. So for such x, the function $f(x)$ *blows up*. The *art* of the Circle Method is to partition $[0, 1]$ to two pieces: $\mathfrak{M}$, called the *major arcs*, consisting of those x's which are *close* to a rational number with small denominator, and $\mathfrak{m}$, the *minor arcs*, the complement of $\mathfrak{M}$. Then we have

$$\int_0^1 f_k(x)^s e^{-2\pi i l x}\, dx = \int_{\mathfrak{M}} f_k(x)^s e^{-2\pi i l x}\, dx + \int_{\mathfrak{m}} f_k(x)^s e^{-2\pi i l x}\, dx.$$

In most applications, including Waring's Problem, the major arcs integral is not too hard to analyze to obtain a fairly explicit asymptotic formula. The real problem is to show that the contribution of the major arcs is not canceled out by the integral over the minor arcs. To see how this is done in a series of instructive examples, see Vaughan [54]. To see the analysis of the major arcs in a situation where we do not know how to handle the minor arcs see [33, Ch. 14]. A major new development in

applications of the Circle Method is Harald Helfgott's recent proof of the Ternary Goldbach Conjecture which asserts that every odd integer larger than 5 is the sum of three prime numbers, available at

https://arxiv.org/abs/1312.7748

Geometry of numbers

Minkowski's Theorem 10.10 and Gauss' Circle Theorem 9.4 belong to an area of mathematics called *geometry of numbers*. Minkowski proved Theorem 10.10 in the course of his work on quadratic forms in relation to Diophantine approximation. To get a feel for the sort of problem Minkowski was interested in, suppose we want to study minimal values of positive definite quadratic forms on integral points. Note if $f(x_1, \ldots, x_n)$ is a positive definite quadratic form with real coefficients, then the set of real points $(x_1, \ldots, x_n)$ such that $f(x_1, \ldots, x_n) < \lambda$ is a bounded convex symmetric domain of the sort considered in this chapter. The volume of this set is $\text{Vol}\{f < 1\} \cdot \lambda^{\frac{n}{2}}$. Since $\text{Vol}\,\mathbb{Z}^n = 1$, Theorem 10.10 implies that if

$$\text{Vol}\{f < 1\} \cdot \lambda^{\frac{n}{2}} > 2^n, \text{ i.e., } \lambda > \frac{4}{\text{Vol}\{f < 1\}^{2/n}},$$

then there is at least one non-zero integral point $\underline{x} \in \mathbb{Z}^n$ such that $f(\underline{x}) < \lambda$. This means that the minimal value of f on non-zero points in $\mathbb{Z}^n$ is at most

$$\frac{4}{\text{Vol}\{f < 1\}^{2/n}}.$$

Suppose, for example, that $n = 2$, and $f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$. The positive-definiteness of f means $a, c > 0$ and $b^2 < 4ac$. In this case, it is a nice exercise to show that

$$\text{Vol}\{f < 1\} = \frac{2\pi}{\sqrt{4ac - b^2}}.$$

Putting everything together, we see that the minimum value of a positive definite quadratic form $f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ on non-zero integral points (x_1, x_2) is at most

$$\frac{2}{\pi}\sqrt{4ac - b^2}.$$

Analogues of the Gauss's Circle Theorem 9.4 appear often in contemporary research papers. In many number theoretic problems one needs to count integral points in a certain domain. Ideally one should be able to replace the number of integral points by the area of the region, which is often not too hard to compute, plus an error term contributed by the points along, or close to, the boundary. However, in order to bound the error one needs to show that there are not too many points on the boundary of the regions, or *tucked away in corners*. This can be a real challenge, as anyone studying the works of Manjul Bhargava (Fields medal, 2014) might notice. Geometry of

numbers methods have featured prominently in Bhargava's groundbreaking works. To see an expository article on the work of Bhargava and the role played by the geometry of numbers, see Gross' article [78].

Davenport's little article [71] is a nice entry way to the subject of geometry of numbers. The classical texts by Cassels [11] and Siegel [45] are wonderful introductions to this exciting area.