1.1 The Geometry of Numbers and Lagrange’s Four-Squares Theorem

In this section, we discuss several results from the geometry of numbers. The underlying idea is (obviously enough) to use geometric ideas to study number theory.

- We have already encountered some of these ideas when discussing the Gaussian integers: they form a lattice inside the complex plane.

First, we recall some terminology:

- 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 $B$.
- A set $B$ in $\mathbb{R}^n$ is symmetric about the origin if, for any $x$ in $B$, the point $-x$ is also in $B$.
- We denote the set of all points in $\mathbb{R}^n$ all of whose coordinates are integers by $\mathbb{Z}^n$.

Our starting point is the following theorem, which says that if a convex set is sufficiently nice and large enough, it must contain an integer point.

**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 volume is $> 2^n$. Then $B$ contains a nonzero point all of whose coordinates are integers.

- We first show the following (sometimes called Blichfeldt’s principle): if $S$ is a bounded set in $\mathbb{R}^n$ whose volume is greater than 1, then there exist two points $x$ and $y$ in $S$ such that $x - y$ has integer coordinates.

  * **Proof**: The idea is essentially to use the pigeonhole principle.
  * For each lattice point $a = (a_1, \cdots, a_n)$, let $R(a)$ be the “box” consisting of the points $(x_1, \cdots, x_n)$ whose coordinates satisfy $a_i \leq x_i < a_{i+1}$.
  * If we then set $S(a) = S \cap R(a)$, we have $\sum_{a \in \mathbb{Z}^n} \text{vol}(S(a)) = \text{vol}(S)$, because each point of $S$ lies in exactly one of the boxes $R(a)$.
  * Now imagine translating the set $S(a)$ by the vector $-a$: it will preserve volume, but move $S(a)$ to land inside $S(0)$. Denote this translated set by $S^*(a)$.
  * Then $\sum_{a \in \mathbb{Z}^n} \text{vol}(S^*(a)) = \text{vol}(S) > 1$.
  * Now notice that each of the sets $S^*(a)$ lies inside $S(0)$, which has volume 1, so there must be some overlap.
  * Hence, there exists some distinct $x, y \in S$ and $a_1, a_2 \in \mathbb{Z}^n$ such that $x - a_1 = y - a_2$: but then $x - y = a_1 - a_2$ is a nonzero lattice point.

- **Proof** (of Minkowski’s Theorem): Suppose $B$ is a convex open set symmetric about 0 whose volume is $> 2^n$, and let $\frac{1}{2}B = \left\{ \frac{1}{2}x : x \in B \right\}$.

  * Notice that since $\text{vol}(B) > 2^n$, we have $\text{vol}(\frac{1}{2}B) > 1$. Apply Blichfeldt’s principle to $\frac{1}{2}B$: we obtain distinct points $x, y \in \frac{1}{2}B$ such that $x - y$ has integer coordinates.
  * Then $2x \in B$ and $2y \in B$.
  * Since $B$ is symmetric about the origin, $-2y \in B$. 

---

Number Theory (part n): The Geometry of Numbers (by Evan Dummit, 2014, v. 1.00)

**Contents**

1.1 The Geometry of Numbers and Lagrange’s Four-Squares Theorem ...................................... 1
* Since $B$ is convex, the midpoint of the line segment joining $2x$ and $-2y$ lies in $B$.
* But this point is simply $x - y$, which is a nonzero point all of whose coordinates are integers, as desired.

- The result of Minkowski’s theorem does not apply merely to the lattice of points having integer coordinates.
  - If $v_1, \ldots, v_n$ are ($\mathbb{R}$-)linearly independent vectors in $\mathbb{R}^n$, the set $A$ 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 domain 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”.
  - A basic fact from linear algebra says: the volume of the fundamental domain is equal to the determinant of the matrix whose columns are the vectors $v_1, \ldots, v_n$ (expressed in terms of the standard basis of $\mathbb{R}^n$).

- **Theorem** (Minkowski’s Theorem for general lattices): Let $\Lambda$ 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 $> 2^n \cdot V$, then $B$ contains a nonzero point of $\Lambda$.
  - **Proof**: Apply a linear transformation $T$ sending the basis vectors of $\Lambda$ 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 standard linear algebra), 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: then $B$ contains a nonzero point of $\Lambda$.

- There are many applications of Minkowski’s theorem in number theory.
  - One very important one, which we do not quite possess the tools to discuss at this stage, is to provide an effective way to determine whether the ring $\mathbb{Z}[\sqrt{D}]$ possesses unique factorization.
  - Instead, we will discuss a simpler application: proving that every positive integer can be written as the sum of four squares.

- **Theorem** (Lagrange): Every positive integer $n$ can be expressed as the sum of four squares of integers.
  - We first show that if $a, b$ are the sum of four squares, then so is $ab$. We then show every prime is the sum of four squares. By applying these results, we immediately see that every positive integer is the sum of four squares.
  - **Lemma 1**: If $a$ and $b$ are the sum of four squares, then so is $ab$.
    * **Proof**: This follows from the following magical-seeming identity:
      \[
      (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 + (x_1y_3 - x_2y_4 - x_3y_1 + x_4y_2)^2 + (x_1y_4 + x_2y_3 - x_3y_2 - x_4y_1)^2
      \]
      which can be verified (if not understood) simply by multiplying out and verifying that all of the cross-terms cancel.
    * Like the corresponding identity for sums of two squares
      \[
      (a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2
      \]
      which arises from the fact that the norm map on $\mathbb{Z}[i]$ is multiplicative, the four-squares identity does not “come from nowhere”: it in fact arises from a norm map on the ring $\mathbb{H}$ of 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 is $a + bi + cj + dk = a - bi - cj - dk$, and the norm map is $N(q) = q\overline{q}$. Explicitly, one can compute that $N(a + bi + cj + dk) = a^2 + b^2 + c^2 + d^2$, and the fact that this map is multiplicative amounts to the four-squares identity.

* In fact, since the norm of a nonzero quaternion is nonzero, we in fact see that every nonzero quaternion has a multiplicative inverse.

* 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 $i = (1,0,0)$, $j = (0,1,0)$, $k = (0,0,1)$ is used for the basis vectors, and then for example one has $i \times j = k$.

* 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.

○ **Lemma 2**: Every prime $p$ can be written as the sum of four squares.

* Proof: As shown on a homework assignment, $-1$ is the sum of two squares modulo $p$ for any prime $p$; say, $-1 \equiv r^2 + s^2 \pmod{p}$.

  * The argument was: if $p$ is odd, the set $S$ of squares modulo $p$ contains $(p+1)/2$ elements, as does the set $T$ of elements of the form $-1 - s^2$, so they must have a nontrivial intersection.

* Now let $\Lambda$ be the lattice spanned by the four vectors $(p,0,0,0)$, $(0,p,0,0)$, $(r,s,1,0)$, and $(s,-r,0,1)$. 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 with cleverness or a brute-force quadruple integration) 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 $\Lambda$ (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 $\Lambda$ in $B$.

* But then

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

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.