Minkowski's Theorem

Preyarnsi Saikia

Indian Institute of Technology Delhi

1. Hermann Minkowski

Hermann Minkowski [7] was a mathematician born on June 22, 1864, in Aleksota (now in Kaunas, Lithuania). Minkowski completed his doctorate in 1885 under the guidance of Ferdinand von Lindemann at the Albertina University of Königsberg in East Prussia. He was granted the Mathematics Prize of the French Academy of Sciences in 1883 for his paper on the theory of quadratic structures.

He worked as a professor at Königsberg, Zürich, and Göttingen and taught great personalities like Albert Einstein. He created and developed the "Geometry of numbers" while exploring the arithmetic of quadratic forms, especially concerning n variables. This method has helped solve several problems in Number Theory and Mathematical Physics. Hermann Minkowski was the mathematician who introduced the four-dimensional space known as the "Minkowski spacetime". His other contributions to Mathematics are Minkowski Sausage and the Minkowski cover of a curve. Minkowski died unexpectedly of an infected appendix in Göttingen on January 12, 1909, after living a short yet accomplished life.

This article will explore in details a celebrated theorem of Minkowski.

2. Preliminaries

Before going into the Minkowski's Theorem, let us recall a few definitions and propositions.

Definition 2.1. A set is said to be a **bounded set** if it is contained in some ball of finite radius.

Definition 2.2. A set A is said to be a **convex set** if for any $x, y \in A$, all points of the line segment joining x and y, i.e., $(1-t)x + ty$ is also in A for $t \in [0, 1]$.

Definition 2.3. A set B is said to be **symmetric about the origin** if for any element $x \in B$,

the element $-x$ is also in B.

Definition 2.4. If $v_1, v_2, \ldots, v_n$ are linearly independent vectors in $\mathbb{R}^n$, the set $\Lambda = \{c_1v_1 + c_2v_2 + \ldots + c_nv_n : c_i \in \mathbb{Z}\}$ is called a **lattice** (n-dimensional) and $B = \{v_1, v_2, \ldots, v_n\}$ is called the **basis**.

Definition 2.5. [6] If Λ is a lattice, then we define the **volume of** Λ as

$$\text{Vol } \Lambda = \left| \text{Det} \begin{pmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \ldots & v_{nn} \end{pmatrix} \right|,$$

where $v_1 = (v_{11}, v_{12}, \ldots, v_{1n}), v_2 = (v_{21}, v_{22}, \ldots, v_{2n}), \ldots, v_n = (v_{n1}, v_{n2}, \ldots, v_{nn})$.

Definition 2.6. Let $a, b \in \mathbb{Z}$. We say that a is **congruent** to b modulo m (which we abbreviate as mod m) if m divides $a - b$. We write this as $a \equiv b \pmod{m}$.

Definition 2.7. A **class of residue** mod m is the class of all the numbers congruent to a given residue mod m, and every member of the class is called a representative of the class. There are in all m classes, represented by

$$0, 1, 2, \ldots, m - 1.$$

These m numbers form a complete system of incongruent residues to modulus m.

Definition 2.8. We define the group $U(n)$ to be the set of positive integers less than n and relatively prime to n. Then $U(n)$ is a group under multiplication modulo n.

Proposition 2.1. $U(n)$ is a multiplicative group and for each $a \in U(n)$, there exists unique $a^{-1} \in U(n)$ such that $a.a^{-1} \equiv 1 \pmod{n}$.

Proposition 2.2 (The Pigeonhole Principle)**.** Suppose n items are to be stored in m containers with $n > m$, then at least one container will contain more than one item.

3. Minkowski's Theorem

Theorem 3.1. *Let $K \subset \mathbb{R}^n$ be a bounded, convex, centrally symmetric set. If in addition, the volume of K satisfies Vol$(K) > 2^n$, then K contains at least one non-zero vector of $\mathbb{Z}^n$.*

Proof. [4] Take the cube $Q = [-1, 1]^n$. This cube is centered at the origin and it translates by even coordinate vectors to partition $\mathbb{R}^n$. We can thus say that

$$\mathbb{R}^n = \bigcup_{u \in 2\mathbb{Z}^n} (Q + u).$$

Let us denote $Q + u$ by Q_u.

We know, K is bounded. Thus K intersects only a finite of these Q_u's. Let us call them χ. Now, let us look at the sets Q_u from χ and their translations back to Q. These translations will create an accumulation of parts of K inside Q.

However, we know that,

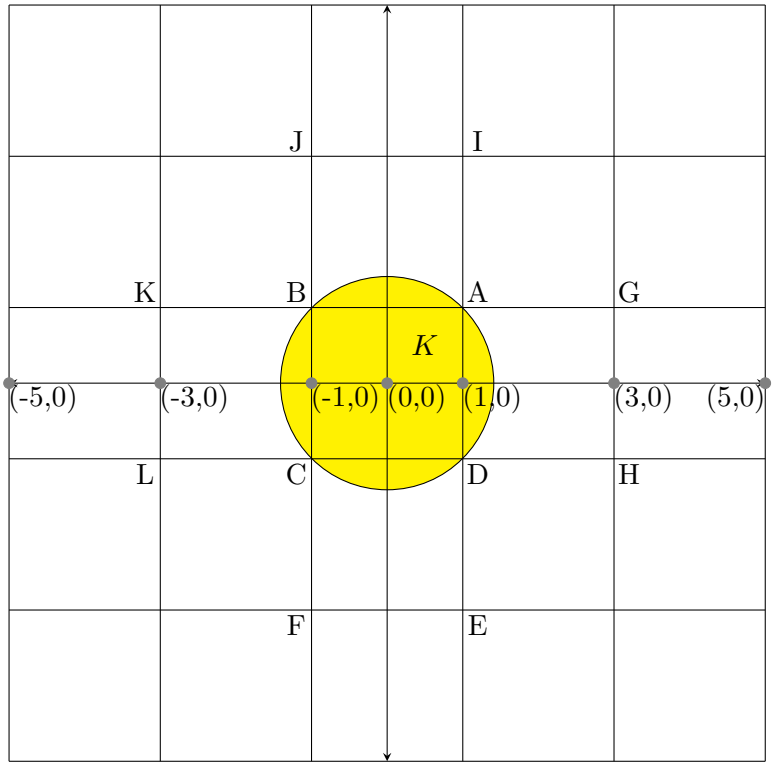
$$\text{Vol}\,(K) > 2^n \text{ whereas } \text{Vol}\,(Q) = 2^n.$$

Therefore there will be at least an overlap of two translated Q_u's. Consider a point x lying in that overlap. This point can be written as $x = v + y = w + z$ for some distinct points y, z in K and some distinct vectors v, w in $2\mathbb{Z}^n$. In particular, we get that the point $\frac{y-z}{2} = \frac{w-v}{2}$ is in $\mathbb{Z}^n$.

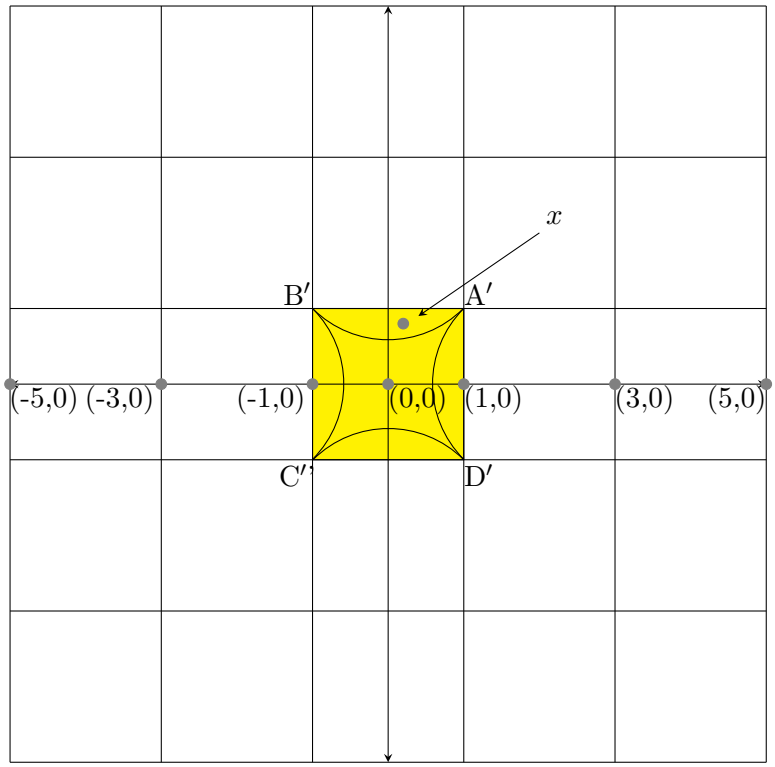
But $y \in K$ and $-z \in K$ as z is in K and K is symmetric with respect to the origin. Thus, the convexity of K yields that $\frac{y-z}{2}$ is also in K, which means that $\frac{y-z}{2}$ is a lattice point that lies in K.

This proves the theorem. □

Graphical representation of the proof for $n = 2$.



In the figure, $K = x^2 + y^2 \leq 2$ such that K is bounded, convex and centrally symmetric. We take $Q = [-1, 1]^2$, i.e., the square ABCD. Vol ABCD = Area ABCD = 4 sq. units. Clearly Vol K = Area $K > 4$ sq. units (from figure). We can translate ABCD by even coordinates to convert the entire $\mathbb{R}^2$ plane to a lattice as shown in the figure. As defined in the above proof, here χ = {ABCD, CDEF, AGHD, AIJB, BKLC}.



All the squares CDEF, AGHD, AIJB, BKLC of χ have been translated back to ABCD to form the new A′B′C′D′. x lies in the overlap of ABCD and CDEF. Thus $x = v + y = w + z$ for some distinct points y, z in K and some distinct vectors v, w in $2\mathbb{Z}^2$. Thus $\frac{y-z}{2} = \frac{w-v}{2}$ is in $\mathbb{Z}^2$. But $\frac{y-z}{2} \in K$ as proved in the theorem above. Thus we conclude that $\frac{y-z}{2}$ (a point in K) is also a point in $\mathbb{Z}^2$.

Minkowski's Theorem for general lattices

A more general version of the theorem is stated below.

Theorem 3.2. *Let Λ be a lattice in $\mathbb{R}^n$. Suppose S is a bounded, convex, centrally symmetric set in $\mathbb{R}^n$ such that Vol$(K) > 2^n$ Vol Λ. Then there is a non-zero vector in the set $\Lambda \cap S$.*

The proof can be found in [5].

4. Applications

4.1. Sum of two squares

Lemma 4.1. [1] *For $a \in U(p)$, where p is an odd prime, $a^2 \equiv 1 \pmod{p}$ if and only if $a \equiv 1 \pmod{p}$ or $a \equiv p-1 \pmod{p}$.*

Proof. Let us first assume that $a^2 \equiv 1 \pmod{p}$. Then p divides $a^2 - 1$. Since p is prime, either

$p \mid a+1$ or $p \mid a-1$.
If $p \mid a+1$, then, $a = pk-1, k \in \mathbb{Z}$. As $a \in U(p), a \equiv p-1 \pmod{p}$.
If $p \mid a-1$, then, $a = pm+1, m \in \mathbb{Z}$. As $a \in U(p), a \equiv 1 \pmod{p}$.

Conversely, if $a \equiv p-1 \pmod{p}$, then $a^2 = p^2 - 2p + 1 \equiv 1 \pmod{p}$.
If $a \equiv 1 \pmod{p}$, then $a^2 \equiv 1 \pmod{p}$.

This proves our lemma. □

Lemma 4.2 (Wilson's Theorem)**.** [1] *Let p be a prime, then $(p-1)! \equiv -1 \pmod{p}$.*

Proof. From Proposition 2.1, we know that $U(n)$ is a group under multiplication modulo n. Every element of $U(n)$ has a unique multiplicative inverse.
For prime p,

$$U(p) = \{1, 2, \ldots, p-1\}.$$

For $p = 2$, $(p-1)! = 1 \equiv -1 \pmod{p}$.

For $p > 2$, from Lemma 4.1, the self inverse elements of $U(p)$ are 1 and $p-1$. $a = a^{-1}$ if and only if $a = 1$ or $a = p-1$.

We can arrange the product $1.2 \ldots (p-1)$ so that each $a \in \{2, 3, \ldots, (p-2)\}$ is adjacent to its multiplicative inverse. We can then simplify to get

$$(p-1)! = 1.1 \ldots 1.(p-1) = (p-1) \equiv -1 \,(\text{mod } p).$$

Thus, our result has been proved. □

Lemma 4.3. [1] *Suppose $p \equiv 1 \,(\text{mod } 4)$, there exists $z \in \mathbb{Z}$ such that $z^2 + 1 \equiv 0 \pmod{p}$.*

Proof. Expanding $(p-1)!$, we get

$$(p-1)! = 1 \times 2 \times 3 \ldots \left(\frac{p-1}{2}\right) \times \left(\frac{p+1}{2}\right) \ldots (p-2) \times (p-1).$$

We have,

$$\begin{aligned} p-1 &\equiv -1 \pmod{p}, \\ p-2 &\equiv -2 \pmod{p}, \\ &\vdots \\ \frac{p+1}{2} &\equiv -\frac{p-1}{2} \pmod{p}. \end{aligned}$$

The last $\frac{p-1}{2}$ factors in the product can be paired with the negatives of the first $\frac{p-1}{2}$ factors so that the factorial becomes

$$\begin{aligned}(p-1)! &\equiv 1\times 2\times 3\ldots\left(\frac{p-1}{2}\right)\times\left(-\left(\frac{p-1}{2}\right)\right)\ldots -3\times -2\times -1 \pmod p\\ &\equiv (-1)^{\frac{(p-1)}{2}}\left(1\times 2\times 3\ldots\left(\frac{p-1}{2}\right)\right)^2 \pmod p.\end{aligned}$$

Wilson's Theorem tells us that $(p-1)!\equiv -1 \pmod p$, so we can write

$$\begin{aligned}(-1)(-1) &\equiv (-1)(-1)^{\frac{(p-1)}{2}}\left(1\times 2\times 3\ldots\left(\frac{p-1}{2}\right)\right)^2 \pmod p\\ 1 &\equiv (-1)^{\frac{(p+1)}{2}}\left(1\times 2\times 3\ldots\left(\frac{p-1}{2}\right)\right)^2 \pmod p.\end{aligned}$$

Now, we know that $p\equiv 1 \pmod 4$, so $p=4k+1$ for some $k\in\mathbb{Z}$. Then $\frac{p+1}{2}=\frac{4k+1+1}{2}=2k+1$, which is odd. Thus, we have, $(-1)^{\frac{(p+1)}{2}}=-1$ and

$$1+\left(1\times 2\times 3\ldots\left(\frac{p-1}{2}\right)\right)^2\equiv 0 \pmod p.$$

Let $z=1\times 2\times 3\ldots\frac{p-1}{2}$, so that we have,

$$z^2+1\equiv 0 \pmod p.$$

This proves the theorem. □

Theorem 4.4. [2] *For prime p, if $p\equiv 1 \pmod 4$, we can always find some $a,b\in\mathbb{Z}$ such that $p=a^2+b^2$.*

Proof. Consider the set

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

The set S is convex, symmetric, and open as it is a circle. Also, Vol $S=2\pi p$. To apply Minkowski's Theorem for generalised lattice, we need to find a lattice Λ such that

(1) 4Vol $\Lambda<2\pi p$;
(2) for all $(a,b)\in\Lambda$, $p\mid a^2+b^2$.

We have, $p\equiv 1 \pmod 4$. Therefore by Lemma 4.3, there exists $z\in\mathbb{Z}$ such that $z^2+1\equiv 0 \pmod p$. We will use this z to construct our required lattice. Consider the vectors:

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

We have,

$$\text{Det}\begin{pmatrix}p & 0\\ z & 1\end{pmatrix}=p.$$

Hence the vectors v_1 and v_2 are linearly dependent (as Det $= p \neq 0$). Let Λ be the lattice generated by v_1 and v_2. By Definition 2.5, Vol $\Lambda = p$. Since, $4p < 2\pi p \implies 4\text{Vol}\,\Lambda < 2\pi p$. So, condition (1) is satisfied.

A typical vector in Λ can be written as $c_1v_1 + c_2v_2$ with $c_1, c_2 \in \mathbb{Z}$. We compute the coordinates of the vector as

$$\begin{aligned} c_1v_1 + c_2v_2 &= (c_1p, 0) + (c_2z, c_2) \\ &= (c_1p + c_2z, c_2). \end{aligned}$$

We again compute the sum of the coordinates to obtain

$$\begin{aligned} (c_1p + c_2z)^2 + c_2^2 &= (c_1p)^2 + (c_2z)^2 + 2c_1c_2pz + c_2^2 \\ &\equiv (c_1z)^2 + c_2^2 \pmod{p} \\ &\equiv c_2^2(z^2 + 1) \pmod{p} \\ &\equiv 0 \pmod{p}. \qquad\qquad (as\ z^2 + 1 \equiv 0 \pmod{p}) \end{aligned}$$

Thus, condition (2) has also been verified.

By Minkowski's Theorem for general lattices, there exists a non-zero vector in the set $\Lambda \cap S$. Thus, there exists $(m, n) \in \Lambda \cap S$ such that

(i) $m^2 + n^2 < 2p$;
(ii) $p \mid m^2 + n^2$;
(iii) $(m, n) \neq (0, 0)$.

The only possible condition is $p = m^2 + n^2$. This proves our theorem. □

4.2. Sum of four squares

Lemma 4.5. [3] *If p is an odd prime, then there exists $x, y \in \mathbb{Z}$ such that*

$$x^2 + y^2 \equiv -1 \pmod{p}.$$

Proof. For $x = 0, 1, \ldots, \dfrac{p-1}{2}$, all of the numbers x^2 have different congruent modulo p. This is because if $x_1^2 \equiv x_2^2 \pmod{p}$, then

$$p \mid (x_1 - x_2)(x_1 + x_2) \Rightarrow x_1 \equiv \pm x_2 \pmod{p},$$

which is a contradiction. So, we have $\dfrac{p+1}{2}$ numbers which are incongruent modulo p.

For $y = 0, 1, \ldots, \dfrac{p-1}{2}$, the numbers $-1 - y^2$ are all incongruent modulo p. Again this is because if $-1 - y_1^2 \equiv -1 - y_2^2 \pmod{p}$, then

$$p \mid -1 - y_1^2 + 1 + y_2^2 = (y_2 - y_1)(y_2 + y_1) \Rightarrow y_1 \equiv \pm y_2 \pmod{p},$$

a contradiction. So, we have another set of $\dfrac{p+1}{2}$ numbers which are incongruent modulo p.

But there are $p + 1$ numbers altogether in these two sets and only p possible residues modulo p. Then, by the "Pigeonhole Principle", at least one number x^2 in the first set must be congruent to a number $-1 - y^2$ in the second set. Hence

$$\begin{aligned} x^2 &\equiv -1 - y^2 \pmod{p} \\ \Rightarrow x^2 + y^2 &\equiv -1 \pmod{p}. \end{aligned}$$

This proves our theorem. □

Theorem 4.6 (Lagranges four square theorem)**.** *Every positive integer n can be expressed as the sum of four squares of integers.*

Proof. For $n = 1$: $1 = 1^2 + 0^2 + 0^2 + 0^2$.

For $n \geq 2$: The identity

$$\begin{aligned} (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = &(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 \\ &+ (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 + (x_1y_3 - x_3y_1 + \\ &x_4y_2 - x_2y_4)^2 + (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2 \end{aligned} \tag{1}$$

can be verified very easily by basic algebra. Through identity (1) we can conclude that if a and b are sum of four squares, then so is ab.

Now we show that every prime is the sum of four squares. From these two conclusions we can show that every positive integer is the sum of four squares.

Let Λ' be the lattice spanned by the four vectors $(p, 0, 0, 0)$, $(0, p, 0, 0)$, $(r, s, 1, 0)$, and $(s, -r, 0, 1)$.

$$\text{Vol}\,\Lambda' = \text{Det}\begin{pmatrix} p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ r & s & 1 & 0 \\ s & -r & 0 & 1 \end{pmatrix} = p^2.$$

Let $B = \{(x_1, x_2, x_3, x_4) : x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2p\}$. B is a convex, centrally symmetric and open set in $\mathbb{R}^4$.

$$\text{Vol } B = 2\pi^2 p^2.$$

Since $2\pi^2 p^2 > 2^4 p^2 \Rightarrow \text{Vol } B > 2^4 \text{ Vol } \Lambda'$, we can use Minkowski's Theorem for general lattices to conclude that there exists a non-zero vector (x_1, x_2, x_3, x_4) in the set $\Lambda' \cap B$.
Thus $(x_1, x_2, x_3, x_4) = a(p, 0, 0, 0) + b(0, p, 0, 0) + c(r, s, 1, 0) + d(s, -r, 0, 1)$.
But then

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 + x_4^2 &= (ap + cr + ds)^2 + (bp + cs - dr)^2 + c^2 + d^2 \\ &\equiv (c^2 + d^2)(1 + r^2 + s^2) \pmod{p} \\ &\equiv 0 \pmod{p}. \qquad \text{(From Lemma 4.5)} \end{aligned}$$

And since (x_1, x_2, x_3, x_4) is non zero 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. This proves our theorem. □

References

[1] Jahnavi Bhaskar. Sum of two squares. *University of Chicago (REU 2008)*, 2008.

[2] Zichao Dong. Minkowski's theorem and its applications. *Semester Notes, Carneige-Mellon University*, 2019.

[3] Matilde N Lalın. Every positive integer is the sum of four squares!(and other exciting problems). *CiteseerX*, 2002.

[4] Cosmin Pohoata. How to look at minkowski's theorem. *AwesomeMath*, 2013.

[5] Noah Stephens-Davidowitz. Introduction to lattices and minkowski's theorem. *Lecture notes NYU*, 2016.

[6] Akshay Venkatesh. Geometry of numbers. *Lecture notes MIT*, 2016.

[7] Wikipedia contributors. Hermann minkowski — Wikipedia, the free encyclopedia, 2021. URL: `https://en.wikipedia.org/w/index.php?title=Hermann_Minkowski&oldid=1032167976`.

"Everyone believes that mathematics is a dry, boring science, relying only on the ability to count. This is ridiculous. Numbers in mathematics play the most insignificant, the last role. Maths is the purest philosophical area, the science of the greatest poets... I don't know how to count at all; you know yourself that many of my students count better than me. I often get confused when solving some numerical problem, and if I were taking an exam, let's say, with Pauker, he would give me zero. And yet I am still a mathematician, and Pauker is not"

– Mikhail Ostrogradsky (1801 – 1862)