THE GEOMETRY OF NUMBERS

DR. BRIAN CONRAD
JULY 11, 2012

The goal of this lecture is to illustrate the idea that geometric arguments in Euclidean space can be used to prove number-theoretic statements about integers. The phrase "the geometry of numbers" is originally due to Minkowski. In this lecture, a geometric argument will be used to prove the Lagrange Four-Square Theorem.

Theorem (Lagrange Four-Square Theorem)**.** *Every natural number n is of the form $n = x^2 + y^2 + z^2 + w^2$, $x, y, z, w \in \mathbb{Z}$.*

Of course, some of these will have to be zero, as in $0 = 0^2+0^2+0^2+0^2$, $1 = 1^2+0^2+0^2+0^2$, and $2 = 1^2+1^2+0^2+0^2$. Additionally, four squares will be necessary, because any $n = 7 \bmod 8$ cannot be written as the sum of three squares (since the squares are $0, 1, 4 \bmod 8$).

The proof will be formulated geometrically in $\mathbb{R}^4$ and uses the rather unrelated fact that $\pi^2 > 8$. In this formulation, the theorem claims that every sphere $x^2 + y^2 + z^2 + w^2 = n$ with n a natural number intersects the lattice of integers $\mathbb{Z}^4 \subset \mathbb{R}^4$.

Proof. One can use Euler's identity, or

$$\left(\sum_{j=1}^{4} x_j^2\right)\left(\sum_{k=1}^{4} y_k^2\right) = \sum_{h=1}^{4} B_n(\mathbf{x}, \mathbf{y})^2$$

where B_n is some bilinear operation $B_n = \sum_{i,j=1}^{4} \pm x_i y_i$, to show that a number is a sum of four squares if its prime factors are.

(This is motivated by the fact that the norm on $\mathbb{C}$ is commutative, so that for real x, y, u, v,

$$(xu - yv)^2 + (xv - yu)^2 = (x^2 + y^2)(u^2 + v^2).$$

If this is generalized to the quarternions $\mathbb{H}$, then one obtains Euler's identity, which can be checked fairly straightforwardly by multiplying out. But it takes some insight to see beforehand — and Euler himself had no conception of the quarternions.)

With $0, 1$, and 2 shown above, then the only numbers for which the four-squares theorem needs to be checked are the odd primes. This doesn't seem particularly helpful, but it will be.

Lemma. Suppose p is an odd prime. Then, $x^2 + y^2 + 1 \equiv 0 \bmod p$ has a solution for $x, y \in \mathbb{Z}$.

Proof. Use a counting argument. Consider the p numbers $0, 1, \dots, p-1$. When squaring them, you get $\frac{p-1}{2}$ pairs of identical squares plus zero, since $p - 1 \equiv -1 \bmod p$, so $(p-1)^2 = (-1)^2 = 1^2$ and so on. (Specifically, since p is odd, then $u^2 \equiv v^2 \bmod p$ iff $u \equiv v \bmod p$.)

Including 0, there are therefore $\frac{p+1}{2}$ squares mod p, so there are $\frac{p+1}{2}$ possibilities for x^2 and also the same number of possibilities for $-1 - y^2$. Each is more than half of p, so there must be some x, y for which they coincide, and for which $x^2 \equiv -1 - y^2 \bmod p$, or $x^2 + y^2 + 1 \equiv 0 \bmod p$. ⊠

Given this lemma and some odd prime p, choose a, b such that $a^2 + b^2 + 1 \equiv 0\ (p)$.

Definition. A lattice $\Lambda \subset \mathbb{R}^n$ is the $\mathbb{Z}$-span of an $\mathbb{R}$-basis: if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of $\mathbb{R}^n$, then $\Lambda = \left\{\sum_{j=1}^{n} m_j \mathbf{v}_j, m_j \in \mathbb{Z}\right\}$.

For example, the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ corresponds to the lattice $\Lambda = \mathbb{Z}^n \subset \mathbb{R}^n$.

For the theorem, consider the lattice

$$\Lambda = \left\{ (u_1, u_2, u_3.u_4) \in \mathbb{Z}^4 \;\middle|\; \begin{array}{l} u_1 \equiv au_3 + bu_4\ (p) \\ u_2 \equiv bu_3 - au_4\ (p) \end{array} \right\}.$$

Though it is not directly obvious, this is in fact a lattice, a fact which depends on some higher algebra. However, it can be directly checked that

$$\Lambda = \mathbb{Z}\text{-span}\left\{ \begin{pmatrix} a \\ b \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} b \\ -a \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} p \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ p \\ 0 \\ 0 \end{pmatrix} \right\}.$$

This does require that these four vectors are linearly independent, but one can check this by showing their determinant is p^2 and thus nonzero.

Claim. If $\lambda \in \Lambda$, then the square of the norm of λ is an integer multiple of p (i.e. $\|\lambda\|^2 \in \mathbb{Z} \cdot p$).

Proof. Write $\lambda = (u_1, u_2, u_3, u_4)$, where $u_1 \equiv au_2 + bu_4$ (p) and $u_2 = bu_3 - au_4$ (p). Brute force could be used to solve the equation $\sum_{i=1}^4 u_i^2 = 0$, but it's a lot easier to work mod p:

$$\begin{aligned}(au - 3 + bu_4)^2 + (bu_3 - au_4)^2 + u_3^2 + u_4^2 &\equiv a^2(u_3^2 + u_4^2) = b^2(u_3^2 + u_4^2) + u_3^2 + u_4^2 \bmod p \\ &\equiv (a^2 + b^2 + 1)(u_3^2 + u_4^2) \bmod p \\ &\equiv 0 \bmod p\end{aligned}$$

by the way a and b were chosen. ⊠

Much of this is a generalization of something similar done in $\mathbb{R}^2$, so if it looks magical, try playing with the simpler case.

With this, the requirement to prove the theorem becomes finding a point in $(\Lambda - \{0\}) \cap \{\mathbf{v} : \|\mathbf{v}\|^2 < 2p\}$ (i.e. some nonzero lattice point with distance less than $2p$ from the origin), since if such a point exists, then its distance is necessarily p. This boils down into a further question: given a lattice $\Lambda \subset \mathbb{R}^n$ and a "nice" $B \subset \mathbb{R}^n$, how can one tell when B contains a nonzero lattice point of Λ? Specifically, B should be convex, so that if $x, y \in B$, then $[x, y] = \{tx + (1-t)y : 0 \le t \le 1\} \in B$ as well, and symmetric about the origin (so that $x \in B$ iff $-x \in B$). As an example, consider any open ball centered at 0.

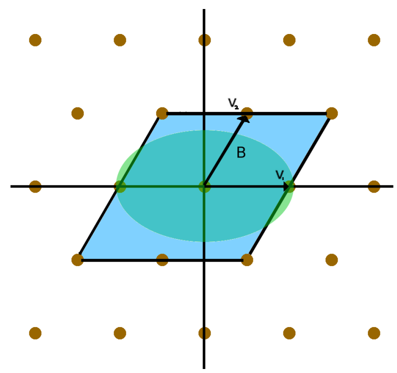

FIGURE 1. Example parallelograms, Λ, and B.

Looking at the plane (which is easier to visualize, as in Figure 1), it is possible to make a parallelogram that is just slightly smaller than 4 of the basic parallelograms tiled together and contains no nonzero lattice points. (The basic parallelogram is just the one bounded by the basis vectors.) In n dimensions, this is generalized to the parallelotope with volume $2^n|\det(\mathbf{v}_1, \dots, \mathbf{v}_n)|$.

However, strange things can happen to the fundamental parallelotope, since a lattice can have multiple $\mathbb{Z}$-bases. For example, $\left\{\binom{1}{0}, \binom{0}{1}\right\}$ and $\left\{\binom{1}{1}, \binom{1}{2}\right\}$ both represent the lattice $\mathbb{Z}^2$. A lattice is invariant under any change-of-basis matrix $T \in M_2(\mathbb{Z})$ provided that T^{-1} has integer entries. Thanks to some nice properties of $\mathbb{Z}$, this is equivalent to $\det T = \pm 1$, or that $T \in \mathrm{GL}_2(\mathbb{Z})$.

Definition. If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a $\mathbb{Z}$-basis of a lattice $\Lambda \subset \mathbb{R}^n$, then a fundamental parallelotope with respect to Λ is

$$P = P_{\{\mathbf{v}_1, \dots, \mathbf{v}_n\}} = \left\{ \sum_{i=1}^n t_i \mathbf{v}_i \mid 0 \le t_i \le 1 \right\}.$$

This parallelotope and its translates cover $\mathbb{R}^n$.

Claim. All fundamental parallelotopes of a given lattice have the same volume, called vol_Λ.[1]

Proof. Suppose P is a fundamental parallelotope corresponding to a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ for some lattice Λ and P' is another fundamental parallelotope corresponding to a basis $\{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ of Λ. Then, there is some change-of-basis matrix C such that $|\det C| = 1$. Then,

$$\mathrm{vol}\, P = |\det(\mathbf{v}_1, \dots, \mathbf{v}_n)| = |\det(C)\det(\mathbf{v}'_1, \dots, \mathbf{v}'_n)| = |\det C||\det(\mathbf{v}'_1, \dots, \mathbf{v}'_n)| = |\det C|\,\mathrm{vol}\, P' = \mathrm{vol}\, P'.$$ ⊠

[1]Notice this is not the volume *of* Λ, which is 0, because it is a discrete lattice.

Theorem (Minkowski)**.** *Suppose $\Lambda \subset \mathbb{R}^n$ is a lattice and $B \subset \mathbb{R}^n$ is convex and symmetric around the origin. Then, if* $\operatorname{vol}(B) > 2^n \operatorname{vol}_\Lambda$ *(which is just* $\operatorname{vol}_{2\Lambda}$*), then* $B \cap (\Lambda - \{0\}) \neq \emptyset$*.*

Minkowski's Theorem is applicable to the four-square problem. Take $B_p = \{\|\mathbf{v}\|^2 < 2p\} \subset \mathbb{R}^4$ and Λ as given before, so that $\operatorname{vol}_\Lambda = 2p^2$. In order for the theorem to be satisfied, we want $\operatorname{vol}(B_p) > 16p^2$. Using the four-dimensional volume of a sphere,

$$\operatorname{vol} B_p = \frac{\pi^2(2p)^2}{2} = 2\pi^2 p^2 > 16p^2$$

because $\pi^2 > 8$. Step back and see how this number-theoretic property about squares of integers rests on this completly geometric property of π, which is totally unexpected.

Proof of Minkowski's Theorem. Consider the region $2P = \{\sum_{i=1}^n t_i \mathbf{v}_i : 0 \leq t \leq 2\}$, and for any lattice point $\mathbf{m} = \sum_{j=1}^n m_j \mathbf{v}_j$ define

$$D_{\mathbf{m}} = 2\mathbf{m} + 2P = \sum_{j=1}^n 2m_j \mathbf{v}_j + 2P$$

Thus, $D_{\mathbf{m}}$ is the parallelotope translated so that one of the corners is at $\mathbf{m}$. Thus, its volume is constant, and $\operatorname{vol}(D_{\mathbf{m}}) = 2^n \operatorname{vol}_\Lambda$. Additionally, they tesselate, since $\mathbf{m}$ is a lattice point: $\mathbb{R}^n = \bigcup_{\mathbf{m}\in\Lambda} D_{\mathbf{m}}$, and they basically don't intersect (the intersections are hyperplanes with measure 0). Thus, $B \cap D_{\mathbf{m}}$ are also essentially disjoint, so since $B = \cup(B \cap D_{\mathbf{m}})$, then

$$\operatorname{vol} B = \sum_{\mathbf{m}} \operatorname{vol}(B \cap D_{\mathbf{m}}) > \operatorname{vol} 2P$$

by the original assumption. Now, it is possible to translate each of these pieces back to the ~~future~~ origin, within the parallelotope $2P$:

$$\implies \bigcup_{\mathbf{m}\in\Lambda} (-2\mathbf{m} + (B \cap D_{\mathbf{m}})) \subseteq 2P.$$

But since these pieces have volume greater than $2P$, there must be distinct $\mathbf{m}$, $\mathbf{m}'$ with a nontrivial intersection:

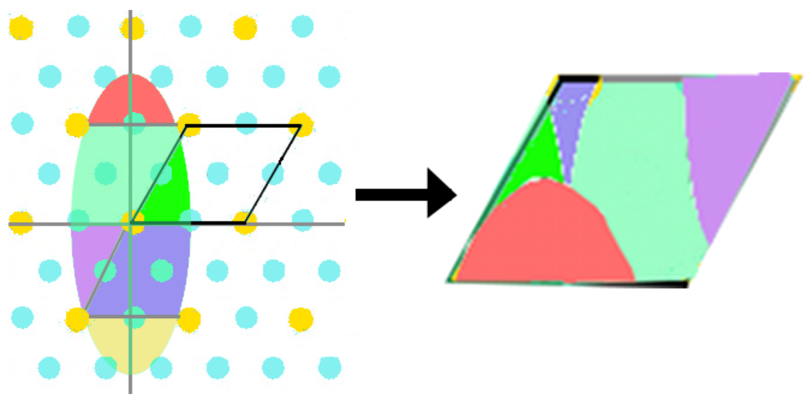

FIGURE 2. Translating B back to the origin to create an intersection.

$-2\mathbf{m} + x = -2\mathbf{m}' + x'$, with $\mathbf{m}, \mathbf{m}' \in \Lambda$ and for some $x, x' \in B$. Thus, $\frac{x'-x}{2} = \mathbf{m}' - \mathbf{m}$, which is also a nonzero lattice point (since $\mathbf{m}'$ and $\mathbf{m}$ are distinct) that is in B (by symmetry, since x is, then so is $-x$, and by convexity, their midpoint is as well). ⊠

The four-squares theorem follows as above. ⊠

A lot of problems in number theory, such as those relating to the theory of quadratic forms, can be solved in similar ways.