LAGRANGE'S FOUR SQUARE THEOREM VIA CONVEX GEOMETRY

STEVEN V SAM

This document was originally written February 10, 2009 and was based on some notes I took of a presentation given by Christian Haase in the summer of 2007. I've edited it and expanded it a bit to make it more self-contained. I don't know the original source but I'm sure it's something standard.

Theorem 1 (Lagrange)**.** *Every nonnegative integer can be written as a sum of four squares, i.e., the function* $\mathbf{Z}_{\geq 0}^4 \to \mathbf{Z}_{\geq 0}$ *given by* $(x, y, z, w) \mapsto x^2 + y^2 + z^2 + w^2$ *is surjective.*

First thing out of the way, it suffices to prove that every prime can be written as a sum of four squares:

Lemma 2. *If each of* m *and* n *can be written as a sum of four squares, then so can* mn*.*

Proof. The fancy way to say this is that $a_1^2 + a_2^2 + a_3^2 + a_4^2$ is the square of the norm of the quaternion $a_1 + a_2\mathbf{i} + a_3\mathbf{j} + a_4\mathbf{k}$, and the norm is multiplicative.

More concretely, this says that (this is Euler's four-square identity):

$$\begin{aligned} &(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) \\ &= ((a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 + (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 \\ &\quad + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2. \end{aligned}$$

□

Lemma 3. *Let* p *be a prime number. There exist integers* α, β *such that* $\alpha^2 + \beta^2 \equiv -1 \pmod{p}$*.*

Proof. If $p = 2$, take $\alpha = 1$, $\beta = 0$. So we may assume that p is odd. Define

$$S := \{\alpha^2 + p\mathbf{Z} \mid 0 \leq \alpha < p/2\} \subseteq \mathbf{Z}/p,$$

i.e., we're taking the residues of $0^2, 1^2, \ldots, ((p/2) - 1)^2$ modulo p.

I claim that $|S| = (p+1)/2$, i.e., all of the squares above are distinct modulo p. To prove this, choose $0 \leq \alpha, \alpha' < p/2$ such that $\alpha^2 \equiv \alpha'^2 \pmod{p}$. Then

$$(\alpha + \alpha')(\alpha - \alpha') = \alpha^2 - \alpha'^2 \equiv 0 \pmod{p},$$

and $\alpha + \alpha' \not\equiv 0 \pmod{p}$ since $\alpha + \alpha' < p$, which implies that $\alpha - \alpha' \equiv 0 \pmod{p}$ since $\mathbf{Z}/p$ has no nonzerodivisors. Hence $\alpha = \alpha'$, which proves the claim.

Similarly, define

$$S' := \{-1 - \beta^2 + p\mathbf{Z} \mid 0 \leq \beta < p/2\} \subseteq \mathbf{Z}/p.$$

Then $|S'| = (p+1)/2$ (either same argument or simply note that S' is naturally in bijection with S). Hence $S \cap S' \neq \varnothing$ by the pigeonhole principle, so we can find α and β such that $\alpha^2 \equiv -1 - \beta^2 \pmod{p}$. □

Date: June 21, 2023.

For the final step, we need to invoke some convex geometry. I'll start with some definitions.

A **lattice** Λ is a discrete subgroup of $\mathbf{R}^d$ which spans $\mathbf{R}^d$ in the sense of vector spaces. A more direct way to describe these is as follows: let $\{v_1, \dots, v_d\}$ be a collection of linearly independent vectors in $\mathbf{R}^d$. Then the subgroup spanned by them is a lattice, and they all arise in this way; we consider $\{v_1, \dots, v_d\}$ as a basis for Λ.

If $\{v_1, \dots, v_d\}$ is a basis for Λ, define the corresponding **fundamental parallelepiped** to be the set

$$\Pi = \{a_1 v_1 + \cdots + a_d v_d \mid 0 \le a_i < 1\}.$$

Then we have

$$\operatorname{vol} \Pi = |\det(v_1 \cdots v_d)|$$

where we're taking the determinant of the $d \times d$ matrix whose columns are $v_1, \dots, v_d$ and vol just means usual volume. Furthermore, this quantity does not depend on the choice of a basis: the group $\mathbf{GL}_n(\mathbf{Z})$ acts transitively on bases and the determinant of every matrix in $\mathbf{GL}_n(\mathbf{Z})$ is ± 1, so we can define $\operatorname{vol} \Lambda = \operatorname{vol} \Pi$.[1]

Lemma 4 (Blichfeldt)**.** *Let $\Lambda \subset \mathbf{R}^d$ be a lattice and $X \subseteq \mathbf{R}^d$ be a measurable set. If $\operatorname{vol} X > \operatorname{vol} \Lambda$, then there exist distinct $x, y \in X$ such that $x - y \in \Lambda$.*

Proof. Let Π be a fundamental parallelepiped of Λ. Then $\mathbf{R}^d = \coprod_{u \in \Lambda} \Pi + u$ (disjoint union), and hence $X = \coprod_{u \in \Lambda} X \cap (\Pi + u)$. Define $X_u := (X - u) \cap \Pi$. Then

$$\operatorname{vol} \Pi = \operatorname{vol} \Lambda < \operatorname{vol} X = \sum_{u \in \Lambda} \operatorname{vol}(X_u + u) = \sum_{u \in \Lambda} \operatorname{vol} X_u.$$

Since each $X_u \subseteq \Pi$, there must exist distinct $u, u' \in \Lambda$ such that $X_u \cap X_{u'} \neq \varnothing$. Take $v \in X_u \cap X_{u'}$ and set $x = v + u$ and $y = v + u'$. ☐

Theorem 5 (Minkowski)**.** *Let $\Lambda \subset \mathbf{R}^d$ be a lattice and $K \subseteq \mathbf{R}^d$ be a centrally symmetric (i.e., $x \in K$ implies $-x \in K$) convex measurable set such that $\operatorname{vol} K > 2^d \operatorname{vol} \Lambda$. Then K contains a nonzero element of Λ.*

Proof. Set $K' := \frac{1}{2}K = \{\frac{1}{2}x \mid x \in K\}$, so $\operatorname{vol} K' = \frac{1}{2^d} \operatorname{vol} K > \operatorname{vol} \Lambda$. Using Blichfeldt's lemma, there exist distinct $x, y \in K'$ such that $x - y \in \Lambda$. Since K' is centrally symmetric, we also have $-y \in K'$, and so $2x, -2y \in K$. Finally, by convexity, this means that $x - y \in K$. ☐

Now we can finish the proof.

Proof of Lagrange's theorem. By Lemma 2, it suffices to prove that every prime p can be written as a sum of four squares. Next, by Lemma 3, there exist integers α, β such that $\alpha^2 + \beta^2 \equiv -1 \pmod{p}$. Define

$$\Lambda := \{\mathbf{a} \in \mathbf{Z}^4 \mid a_1 \equiv \alpha a_3 + \beta a_4 \pmod{p}, \quad a_2 \equiv \beta a_3 - \alpha a_4 \pmod{p}\}.$$

Being a subgroup of $\mathbf{Z}^4$, it is clear that Λ is discrete. Also, the set $\{0, \dots, p-1\}^2 \times \{(0,0)\}$ surjects onto $\mathbf{Z}^4/\Lambda$ under the projection, so Λ is a finite index subgroup of $\mathbf{Z}^4$, and hence is lattice with $\operatorname{vol} \Lambda = |\mathbf{Z}^4/\Lambda| \le p^2$. Next, define the ball

$$B := \{\mathbf{a} \in \mathbf{R}^4 \mid \|\mathbf{a}\| < \sqrt{2p}\}.$$

This is convex, measurable, and centrally symmetric. Since

$$\operatorname{vol} B = \frac{\pi^2}{2}(\sqrt{2p})^4 = 2\pi^2 p^2 > 16p^2 \ge 2^4 \operatorname{vol} \Lambda,$$

[1] I don't think this is standard notation but it makes it easier for me to remember what it means.

we can apply Minkowski's theorem to find $\mathbf{a} \in \Lambda$ such that $0 < \|\mathbf{a}\|^2 < 2p$. Since $\|\mathbf{a}\|^2 = a_1^2 + a_2^2 + a_3^2 + a_4^2$, we conclude that (working modulo p):

$$\begin{aligned}\|\mathbf{a}\|^2 &\equiv (\alpha a_3 + \beta a_4)^2 + (\beta a_3 - \alpha a_4)^2 + a_3^2 + a_4^2 \\ &\equiv (\alpha^2 + \beta^2 - 1)a_3^2 + (\alpha^2 + \beta^2 - 1)a_4^2 \\ &\equiv 0 \pmod{p},\end{aligned}$$

and hence $a_1^2 + a_2^2 + a_3^2 + a_4^2$ is a positive integer multiple of p. Since $0 < \|\mathbf{a}\|^2 < 2p$, this multiple must be 1. □