# Math 4527 (Number Theory 2)

## Lecture #32 of 37 $\sim$ April 7, 2021

---

Minkowski's Theorems + Sums of Two and Four Squares

- Minkowski's Convex-Body Theorem
- Sums of Two Squares
- Sums of Four Squares

This material represents §9.1.1-9.1.2 from the course notes.

# Overview

We now start the final chapter of the course, which is on the geometry of numbers.

- Broadly speaking, we will discuss applications of geometric ideas in number theory.
- These will primarily revolve around interpreting various quantities as giving lattices in Euclidean space, and then using various geometric methods to establish the existence of a "small" vector in the lattice.
- Although this might not seem to be a particularly useful thing, in fact we will be able to get quite a lot of mileage out of the results we discuss.

First, we will review some basic terminology[1] for sets in $\mathbb{R}^n$.

[1] I promise to do as little analysis and topology as possible, even though analysis and topology are wonderful!

# Some Terminology, I

We will denote the set of all points in $\mathbb{R}^n$ all of whose coordinates are integers by $\mathbb{Z}^n$.

**Definition**

*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 $B$.*

Examples:

1. The $n$-ball of radius $r$ centered at the origin in $\mathbb{R}^n$, given by the points $(x_1, x_2, \ldots, x_n)$ with $x_1^2 + x_2^2 + \cdots + x_n^2 \leq r^2$, is a convex set.
2. The unit cube, given by the points $(x_1, x_2, \ldots, x_n)$ with $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$, is a convex set.

# Some Terminology, II

We may distinguish three different classes of points in $\mathbb{R}^n$ relative to $B$, based on their behaviors when we draw balls around them.

1. If we can draw a ball around $P$ that is entirely contained in $B$, then $P$ is called an interior point of $B$.
2. If we can draw a ball around $P$ that is entirely contained in $B^c$, the complement of $B$, then we call $P$ an exterior point of $B$. (Equivalently, it is an interior point of $B^c$.)
3. Otherwise, no matter what size of ball we draw, it will always contain some points in $B$ and some points in $B^c$. Points with this property are called boundary points. Note that boundary points can be in $B$ or in $B^c$.

# Some Terminology, III

Definition: The interior of the set $B$, denoted $\mathrm{int}(B)$, is the set of its interior points. A set $B$ is open if $B = \mathrm{int}(B)$.

- Example: The open unit $n$-ball $B$ in $\mathbb{R}^n$, given by the points $(x_1, x_2, \ldots, x_n)$ with $x_1^2 + x_2^2 + \cdots + x_n^2 < 1$, is indeed an open set, since any point in this set is an interior point. (If a point is a distance $r = 1 - \epsilon$ from the origin, then the ball of radius $\epsilon/2$ is contained in $B$.)
- Non-Example: The closed unit $n$-ball in $\mathbb{R}^n$, given by the points $(x_1, x_2, \ldots, x_n)$ with $x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1$, is not an open set, since any point with $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$ is a boundary point, rather than an interior point.

It is moderately straightforward to see that if $B$ is an $n$-dimensional convex set in $\mathbb{R}^n$, then its interior is also convex.

# Some Terminology, IV

Definition: A set $B$ in $\mathbb{R}^n$ is symmetric about the origin if, for any $x$ in $B$, the point $-x$ is also in $B$.

- Example: The $n$-ball of radius $r$ centered at the origin in $\mathbb{R}^n$ is symmetric about the origin.
- Non-Example: The unit cube, given by the points $(x_1, x_2, \ldots, x_n)$ with $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$, is not symmetric about the origin.

We will also use the concept of (Lebesgue) measurability, which I don't want to define since this isn't an analysis class.

- Informally, a set $B$ is measurable if we can assign a sensible notion of $n$-dimensional volume to it.
- We can then compute the volume of $B$ by integrating the characteristic function (1 on $B$, 0 elsewhere) on $\mathbb{R}^n$.

# Minkowski's Convex-Body Theorem, I

Our goal now is to prove that if a convex set is sufficiently nice and has a sufficiently large $n$-measure (i.e., $n$-volume), it must contain a lattice point. First, a preliminary result:

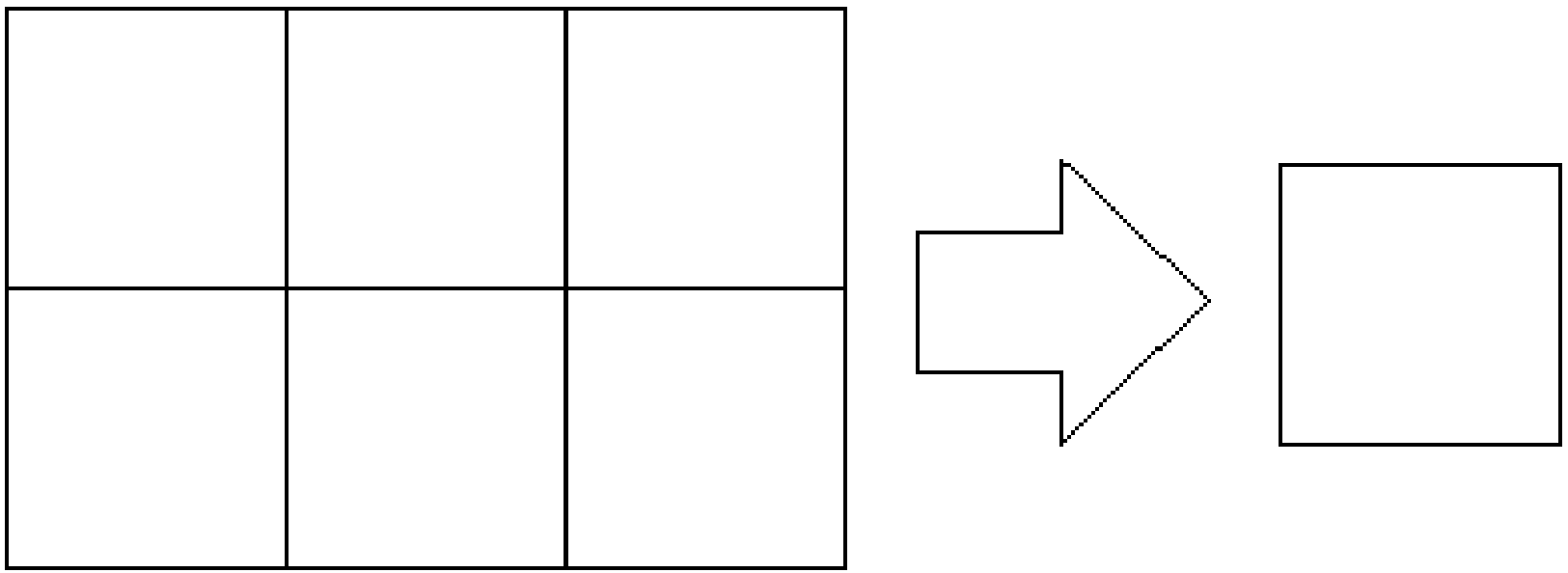
**Proposition (Blichfeldt's Principle)**

*If $S$ is a bounded measurable set in $\mathbb{R}^n$ whose $n$-measure is greater than 1, then there exist two points $x$ and $y$ in $S$ such that $x - y$ has integer coordinates.*

The idea of the proof is essentially to use the pigeonhole principle, but with measure.

# Minkowski's Convex-Body Theorem, II

Let me give the "proof by picture" first:



Take each piece of the region inside the grid of $1 \times 1$ squares on the left, and translate them into the grid on the right. If the area of the set on the left is $> 1$, then the area inside the box on the right is also $> 1$, so there is an overlap.

# Minkowski's Convex-Body Theorem, III

Proof:

- For each lattice point $\mathbf{a} = (a_1, \cdots, a_n)$, let $R(\mathbf{a})$ be the "box" consisting of the points $(x_1, \cdots, x_n)$ whose coordinates satisfy $a_i \leq x_i < a_{i+1}$, and also set $S(\mathbf{a}) = S \cap R(\mathbf{a})$.
- Then we have $\sum_{\mathbf{a} \in \mathbb{Z}^n} \text{vol}(S(\mathbf{a})) = \text{vol}(S)$, because each point of $S$ lies in exactly one of the boxes $R(\mathbf{a})$.
- Now imagine translating the set $S(\mathbf{a})$ by the vector $-\mathbf{a}$: this action will preserve measure, but it moves $S(\mathbf{a})$ to land inside $S(\mathbf{0})$. Denote this translated set by $S^*(\mathbf{a})$.
- Then $\sum_{\mathbf{a} \in \mathbb{Z}^n} \text{vol}(S^*(\mathbf{a})) = \text{vol}(S) > 1$.
- Now notice that each of the sets $S^*(\mathbf{a})$ lies inside $S(\mathbf{0})$, which has volume 1, so there must be some overlap.
- Hence, there exists some distinct $x, y \in S$ and $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{Z}^n$ such that $x - \mathbf{a}_1 = y - \mathbf{a}_2$. Then $x - y = \mathbf{a}_1 - \mathbf{a}_2$ is a nonzero lattice point, as required.

# Minkowski's Convex-Body Theorem, IV

Proof (once more, with feeling):

- Let $\chi_B(x)$ be the characteristic function of $B$ (1 if $x \in B$, 0 if $x \notin B$, which is integrable because $B$ is measurable.
- If we write $\psi(x) = \sum_{\mathbf{v} \in \mathbb{Z}^n} \chi_B(x + \mathbf{v})$, then $\psi$ is bounded because $B$ is bounded so there are only finitely many nonzero terms for any $\mathbf{v} \in \mathbb{Z}^n$.
- We may then integrate both sides and change the order of integration and summation (because the sum is a finite sum of nonnegative terms) and use the translation-invariance of the measure on $\mathbb{R}^n$ to see that $\int_{[0,1]^n} \psi(x)\, dx = \int_{[0,1]^n} \sum_{\mathbf{v} \in \mathbb{Z}^n} \chi_B(x + v)\, dx = \sum_{\mathbf{v} \in \mathbb{Z}^n} \int_{[0,1]^n} \chi_B(x + \mathbf{v})\, dx = \sum_{\mathbf{v} \in \mathbb{Z}^n} \int_{[0,1]^n + \mathbf{v}} \chi_B(x)\, dx = \int_{\mathbb{R}^n} \chi_B(x)\, dx > 1$, since this last integral is simply the measure of $B$.
- This means $\psi(x) \geq 2$ for some $x \in [0,1]^n$, which gives the desired points.

# Minkowski's Convex-Body Theorem, V

Now we may prove our first main result:

**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 n-measure is greater than $2^n$. Then $B$ contains a nonzero point all of whose coordinates are integers.*

We will remark that the bound here is sharp, in the sense that we cannot lower the bound to any number less than $2^n$. Also, if we change "open" to "closed", then we may weaken the condition to "measure at least $2^n$".

# Minkowski's Convex-Body Theorem, VI

Proof:

- 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 $\mathrm{vol}(B) > 2^n$, we have $\mathrm{vol}(\frac{1}{2}B) > 1$.
- Now apply Blichfeldt's principle to the set $\frac{1}{2}B$: we obtain distinct points $x, y \in \frac{1}{2}B$ such that $x - y$ has integer coordinates.
- Then $2x \in B$ and $2y \in B$. Furthermore, since $B$ is symmetric about the origin, $-2y \in B$.
- Then because $B$ 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 in $B$ all of whose coordinates are integers, as desired.

# Minkowski's Convex-Body Theorem, VII

The result of Minkowski's theorem does not apply merely to the lattice $\mathbb{Z}^n$ of points having integer coordinates.

- If $v_1, \ldots, v_n$ are ($\mathbb{R}$-)linearly independent vectors in $\mathbb{R}^n$, the set $\Lambda$ of vectors of the form $c_1 v_1 + \cdots + c_n v_n$, where each $c_i \in \mathbb{Z}$, is called a lattice.
- A fundamental region 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 parallelepiped (i.e., a "skew box").
- The points in this fundamental region give unique representatives for the quotient group $\mathbb{R}^n/\Lambda$, up to an appropriate choice of representatives on the boundary of the region.

# Minkowski's Convex-Body Theorem, VIII

Now we invoke a fact from linear algebra:

**Proposition (Volume of a Parallelepiped)**

*If $v_1, \ldots, v_n$ are arbitrary vectors in $\mathbb{R}^n$, then the signed volume of the parallelepiped they form is equal to the determinant of the matrix whose columns are the $v_i$.*

- One may prove this fact by a direct calculation and induction (namely, by projecting $v_n$ into the subspace spanned by the other vectors, and computing the resulting "height").
- Another quite efficient approach is to use wedge products.

# Minkowski's Convex-Body Theorem, IX

More structurally, the result follows by observing that the signed volume of the fundamental domain satisfies the same properties as the determinant:

1. Interchanging two vectors scales the signed volume by $-1$.
2. Scaling a vector scales the signed volume by the same amount.
3. Adding a multiple of one vector to another does not change the signed volume.
4. The signed volume for the standard basis is 1.

The determinant can be shown to be the only multilinear function satisfying these four properties, and so the signed volume is equal to the determinant. (The usual approach is to prove that the space of functions satisfying (1)-(3) is one-dimensional.)

# Minkowski's Convex-Body Theorem, X

By changing basis, we may give a version of Minkowski's theorem for general lattices:

**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 $> 2^n V$, then $B$ contains a nonzero point of $\Lambda$.*

The idea is just to change basis to the standard basis, which will turn $\Lambda$ into $\mathbb{Z}^n$ and rescale the volume of an arbitrary region by a factor of $1/V$.

# Minkowski's Convex-Body Theorem, XI

Proof:

- Apply the 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 the observation made about determinants above, so this 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. This immediately implies that $B$ contains a nonzero point of $\Lambda$, as required.

# Sums of Two Squares, I

As our first application of Minkowski's convex body theorem, we will prove that every prime $p$ congruent to 1 modulo 4 can be expressed as the sum of two squares.

- We have previously established this result as a consequence of studying factorizations in $\mathbb{Z}[i]$. The argument we will give using Minkowski's theorem is quite different.

**Theorem (Fermat's Two-Squares Theorem)**

*If $p$ is any prime congruent to 1 modulo 4, then there exist integers $a$ and $b$ such that $p = a^2 + b^2$.*

This result was first explicitly noted by Girard in 1625, about 15 years before Fermat observed it. Fermat also did not provide a proof; the first actual proof was given by Euler.

# Sums of Two Squares, II

Proof:

- We start with an observation that we have already made several times: if $p \equiv 1 \pmod 4$ then $-1$ is a square modulo $p$.
- This observation follows immediately from Euler's criterion $\left(\frac{-1}{p}\right) \equiv (-1)^{(p-1)/2} \equiv 1 \pmod p$, so since the Legendre symbol evaluates to $+1$, this means $-1$ is a quadratic residue.
- Alternatively, we could note that the group of nonzero residue classes modulo $p$ is cyclic and has order $p-1$, and so since 4 divides $p-1$, there exists an element $r$ of order 4. Then $r^2$ has order 2, but the only element of order 2 modulo $p$ is $-1$.

# Sums of Two Squares, III

Proof (continued):

- Now suppose $-1 \equiv m^2 \pmod{p}$, and consider the lattice $\Lambda$ be the lattice in $\mathbb{R}^2$ spanned by the two vectors $\langle 1, m \rangle$ and $\langle 0, p \rangle$.
- The determinant of these two vectors is $p$, so the volume of the fundamental domain is $p$.
- Let $B$ be the interior of the disc $x_1^2 + x_2^2 < 2p$ in $\mathbb{R}^2$, and observe that $B$ is open, convex and centrally-symmetric. From elementary geometry, the area of this disc is $2\pi p$.
- Since $2\pi > 4$, the volume of $B$ is larger than $2^2$ times the volume of the fundamental domain of $\Lambda$, and so by Minkowski's theorem, we conclude that there is a nonzero element $\langle x_1, x_2 \rangle = a \langle 1, m \rangle + b \langle 0, p \rangle$ of $\Lambda$ in $B$.

# Sums of Two Squares, IV

Proof (continued):

- We have $-1 \equiv m^2 \pmod{p}$ and we have just shown that there is a nonzero element $\langle x_1, x_2 \rangle = a \langle 1, m \rangle + b \langle 0, p \rangle$ of $\Lambda$ in $B$.
- But then

$$\begin{aligned} x_1^2 + x_2^2 &= a^2 + (ma + bp)^2 \\ &\equiv a^2(1 + m^2) \pmod{p} \\ &\equiv 0 \pmod{p} \end{aligned}$$

  and since $\langle x_1, x_2 \rangle$ is a nonzero integer and $x_1^2 + x_2^2 < 2p$, the only possibility is that $x_1^2 + x_2^2 = p$.
- Thus, $p$ is the sum of two squares, and we are done.

# Sums of Two Squares, V

With the characterization of primes congruent to 1 modulo 4 in hand, we could then give a classification of the integers that are the sum of two squares.

- We have already done this, however, so we will not bother to do it again.
- Instead, we will note that the key idea is to use the fact that the product of the sum of two squares is also the sum of two squares, which follows from the identity $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$.

# Sums of Four Squares, I

We can give a similar kind of argument to establish that every positive integer $n$ can be expressed as the sum of four squares, which is a result first proven by Lagrange.

- This result was known, in most respects, to the ancient Greeks, and was stated explicitly by Bachet in 1621 in his translation notes of the works of Diophantus.
- The first actual proof was given by Lagrange in 1770, and in 1834 Jacobi extended the result to give a formula for the number of representations of $n$ as a sum of four squares.
- Jacobi's result is as follows: if $\sigma(n)$ represents the sum of the divisors of $n$ and $r_4(n)$ is the number of ways of writing $n$ as the sum of four squares, then $r_4(n) = 8\sigma(n)$ if $n$ is odd and $r_4(n) = 24\sigma(d)$ if $n = 2^k d$ ($d$ odd) is even.

# Sums of Four Squares, II

We first show that if $a, b$ are the sum of four squares, then so is $ab$:

**Lemma (Products of Sums of Four Squares)**

*If $a$ and $b$ are the sum of four squares, then so is $ab$.*

Proof:

- This follows from the following identity:

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

which can be verified simply by multiplying out and verifying that all of the cross-terms cancel.

# Sums of Four Squares, III

Like the corresponding identity for sums of two squares, which arises from the fact that the norm map on $\mathbb{Z}[i]$ is multiplicative, the four-squares identity also arises from a norm map on a ring: here, it is the noncommutative ring $\mathbb{H}$ of quaternions.

- 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 on $\mathbb{H}$ is $\overline{a + bi + cj + dk} = a - bi - cj - dk$, and the norm map is $N(q) = q\overline{q}$. One may compute explicitly that $N(a + bi + cj + dk) = a^2 + b^2 + c^2 + d^2$, and the fact that the norm map is multiplicative (which is not obvious from its definition because the multiplication of quaternions is not commutative) amounts to the four-squares identity.

# Sums of Four Squares, IV

In fact, since the norm of a nonzero quaternion is nonzero, the multiplicativity of the norm map implies that every nonzero quaternion has a multiplicative inverse, which is to say, the quaternions form a division ring (which is the noncommutative equivalent of a field).

- 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 $\mathbf{i} = \langle 1, 0, 0 \rangle$, $\mathbf{j} = \langle 0, 1, 0 \rangle$, $\mathbf{k} = \langle 0, 0, 1 \rangle$ is used for the basis vectors, and then for example one has $\mathbf{i} \times \mathbf{j} = \mathbf{k}$.
- You may also have encountered the quaternion group $Q_8$, which is simply the multiplicative subset $\{\pm 1, \pm i, \pm j, \pm k\}$.

# Sums of Four Squares, V

The quaternions were originally described[2] by Hamilton in 1843, which is why the ring of quaternions is denoted by $\mathbb{H}$.

- As a historical note, the development of quaternions actually predates the modern language of vectors by about 40 years, and so many of the classical results in physics (e.g., Maxwell's equations) predating the 20th century were originally written in terms of quaternions rather than vectors.
- 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.

[2] Someone should make a musical about this.

# Sums of Four Squares, VI

We need one additional lemma:

**Lemma**

*For any prime $p$, there exist integers $r$ and $s$ such that $r^2 + s^2 \equiv -1 \pmod{p}$. In other words, $-1$ is the sum of two squares modulo $p$.*

Proof:

- If $p = 2$ the result is obvious, so suppose $p$ is odd.
- From our results on quadratic residues, the set $S$ of squares $r^2$ modulo $p$ contains $(p+1)/2$ elements. Thus, the set $T$ of elements of the form $-1 - s^2$ also has $(p+1)/2$ elements.
- Since there are only $p$ residue classes modulo $p$, the sets $S$ and $T$ must intersect nontrivially: then we have $r^2 \equiv -1 - s^2 \pmod{p}$ and so $r^2 + s^2 \equiv -1 \pmod{p}$, as required.

# Sums of Four Squares, VII

We can now establish our main result:

**Theorem (Lagrange's Four-Square Theorem)**

*If $n$ is any positive integer, then $n$ can be written as the sum of four squares. In other words, there exist integers $a, b, c, d$ such that $n = a^2 + b^2 + c^2 + d^2$.*

Since products of sums of four squares are also sums of four squares, we only have to prove that every prime $p$ can be written as the sum of four squares.

# Sums of Four Squares, VIII

Proof:

- Let $p$ be a prime. By the Lemma, there exist integers $r$ and $s$ such that $r^2 + s^2 \equiv -1 \pmod{p}$.
- Now let $\Lambda$ be the lattice in $\mathbb{R}^4$ spanned by the four vectors $\langle p, 0, 0, 0\rangle$, $\langle 0, p, 0, 0\rangle$, $\langle r, s, 1, 0\rangle$, and $\langle s, -r, 0, 1\rangle$. 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 to be $2\pi^2 p^2$ (there are fairly efficient approaches via cylindrical or spherical coordinates).
- Since the volume of $B$ is larger than $2^4$ times the volume of the fundamental domain of $\Lambda$ (since $2\pi^2 p^2 > 16p^2$), Minkowski's theorem implies that there is a nonzero element of $\Lambda$ in $B$.

# Sums of Four Squares, IX

Proof (continued):

- We have $r^2 + s^2 \equiv -1$ and we showed that there is a nonzero element of $\Lambda$ in $B$. Suppose it is $\langle x_1, x_2, x_3, x_4 \rangle = a\langle p, 0, 0, 0\rangle + b\langle 0, p, 0, 0\rangle + c\langle r, s, 1, 0\rangle + d\langle s, -r, 0, 1\rangle$.
- Then

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

  and since $\langle x_1, x_2, x_3, x_4 \rangle$ is a nonzero integer with $x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2p$, 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.

# Summary

We proved Minkowski's convex-body theorem for general lattices.

We used Minkowski's theorem to prove every prime congruent to 1 modulo 4 is the sum of two squares.

We used Minkowski's theorem to prove every positive integer is the sum of four squares.

Next lecture: Sums of three squares, the ideal class group.