Math 4527 (Number Theory 2)

Lecture #33 of 37 $\sim$ April 8, 2021

Sums of Three Squares + Ideal Class Groups

- Sums of Three Squares
- Ideal Class Groups of Quadratic Integer Rings

This material represents §9.1.3-9.2.1 from the course notes.

Minkowski's Convex-Body Theorem

Recall Minkowski's convex-body theorem from last time:

Theorem (Minkowski's Theorem for General Lattices)

Let Λ 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 Λ.

We used the theorem to analyze sums of 2 and 4 squares yesterday.

Sums of Three Squares, I

Today, we will use Minkowski's theorem to discuss sums of 3 squares, although this problem turns out to be much harder.

- By testing small examples, one is rapidly led to the conjecture that n may be written as the sum of three squares if and only if n is not a power of 4 times an integer that is 7 modulo 8 (i.e., when $n \neq 4^a(8b+7)$ for some a, b).
- As we showed back in chapter 6, if $n = 4^a(8b+7)$ then n is not the sum of three squares. For $a = 0$ this follows immediately by considering n modulo 8, and then we can induct on a.
- It remains to establish that integers not of this form can be written as the sum of three squares.

Sums of Three Squares, II

We will treat the case where $n \equiv 3 \pmod 8$, since the exposition is easiest to give there.

- Unlike in the case for sums of two squares and sums of four squares, the set of integers that are a sum of three squares is not closed under multiplication: both $3 = 1^2 + 1^2 + 1^2$ and $5 = 2^2 + 1^2 + 0^2$ are the sum of three squares, but $15 = 3 \cdot 5$ is not.
- Therefore, we cannot simply reduce to the case of considering representations of primes, as we did for the case of sums of two and four squares.
- Our approach will be to use Minkowski's theorem along with our characterization of integers that are expressible as the sum of two squares.

Sums of Three Squares, III

This particular argument was originally given by Ankeny in 1956.

Theorem (Sums of 3 Squares, 3 Mod 8 Case)

If n is a positive integer congruent to 3 modulo 8, then n is the sum of three squares.

The general idea of the proof is to use Minkowski's theorem to show that we can write $n = R^2 + N$ for integers R and N where N has a particular form that allows us to show it is a sum of two squares. Then n will be the sum of three squares, as desired.

Sums of Three Squares, IV

Proof:

- First, we observe that there exists a prime $q \equiv 1 \pmod 4$ such that $-2q$ is a quadratic residue modulo n.
- This follows from Dirichlet's theorem on primes in arithmetic progressions, since saying $-2q$ is a quadratic residue modulo n is simply a congruence condition modulo n.
- So, since $-2q$ is a quadratic residue modulo n, its reciprocal is also: say with $-1/(2q) \equiv t^2 \pmod n$.
- Next, we observe that $\left(\frac{-2q}{n}\right) = \left(\frac{-2}{n}\right)\left(\frac{q}{n}\right) = \left(\frac{-2}{n}\right)\left(\frac{n}{q}\right) = \left(\frac{-n}{q}\right)$ by quadratic reciprocity for Jacobi symbols and the fact that $n \equiv 3 \pmod 8$ so that $\left(\frac{-2}{n}\right) = +1$ and that $q \equiv 1 \pmod 4$ so that $\left(\frac{n}{q}\right) = \left(\frac{q}{n}\right)$ and $\left(\frac{-1}{q}\right) = +1$.

Sums of Three Squares, V

Proof (continued):

- Therefore, $-n$ is a quadratic residue modulo q, say with $-n \equiv b^2 \pmod{q}$ where we may assume that b is odd.
- This means $b^2 + n = qh'$ for some $h' \in \mathbb{Z}$, but now since $n \equiv 3 \pmod{8}$, reducing both sides modulo 4 yields $h' \equiv 0 \pmod{4}$, and so $h' = 4h$.
- So, to summarize, we have integers q, b, and h such that $-1/(2q) \equiv t^2 \pmod{n}$ and $b^2 + n = 4qh$.

Sums of Three Squares, VI

Proof (continued more):

- We have integers q, b, and h such that $-1/(2q) \equiv t^2 \pmod{n}$ and $b^2 + n = 4qh$.
- Let Λ be the lattice in $\mathbb{R}^3$ spanned by the vectors $\langle 2tq, \sqrt{2q}, 0\rangle$, $\langle tb, b/\sqrt{2q}, \sqrt{n/(2q)}\rangle$, and $\langle n, 0, 0\rangle$. The determinant of these three vectors is $n^{3/2}$, so the volume of the fundamental domain is $n^{3/2}$.
- Now let B be the convex, centrally-symmetric open set in $\mathbb{R}^3$ defined by $x_1^2 + x_2^2 + x_3^2 < 2n$, whose volume is $\frac{4}{3}\pi(2n)^{3/2}$ since it is merely a sphere of radius $\sqrt{2n}$.
- Since the volume of B is larger than 2^3 times the volume of the fundamental domain of Λ (since $\frac{4}{3}\pi \cdot 2^{3/2} > 8$), we conclude that there is a nonzero element of Λ in B.

Sums of Three Squares, VII

Proof (continued even more):

- We have integers q, b, and h such that $-1/(2q) \equiv t^2 \pmod{n}$ and $b^2 + n = 4qh$, and a nonzero element $\langle R, S, T\rangle = x\langle 2tq, \sqrt{2q}, 0\rangle + y\langle tb, b/\sqrt{2q}, \sqrt{n/(2q)}\rangle + z\langle n, 0, 0\rangle$ of Λ.
- Then $R = 2tqx + tby + nz$, $S = \sqrt{2q}x + \frac{b}{\sqrt{2q}}y$, $T = \sqrt{\frac{n}{2q}}y$ so

$$\begin{aligned} R^2 + S^2 + T^2 &= (2tqx + tby + nz)^2 + (\sqrt{2q}x + \tfrac{b}{\sqrt{2q}}y)^2 + (\sqrt{\tfrac{n}{2q}}y)^2 \\ &\equiv (t^2 + 1/(2q))(2qx + by)^2 \pmod{n} \\ &\equiv 0 \pmod{n}. \end{aligned}$$

- Notice also that $R^2 + S^2 + T^2$ is an integer, because it equals $R^2 + 2qx^2 + 2bxy + 2hy^2$, and these quantities are all integers.
- Since $R^2 + S^2 + T^2 < 2n$, we must have $R^2 + S^2 + T^2 = n$.

Sums of Three Squares, VIII

Proof (continued even yet more):

- Now we will show that the integer $N = S^2 + T^2 = 2qx^2 + 2bxy + 2hy^2$ is actually the sum of two integer squares, which will complete the proof because then $n = R^2 + N$ is then the sum of three squares.
- So suppose p is an odd prime dividing N to an odd power, meaning that p^{2a+1} divides N but p^{2a+2} does not: we wish to show that $p \equiv 1 \pmod 4$.
- We have two cases: either p divides n, or it does not.

Sums of Three Squares, VIII

Proof (continued even yet still more):

- First suppose p does not divide n: then $n \equiv R^2 \pmod{p}$ and so $\left(\frac{n}{p}\right) = +1$.
- Also, if $p = q$ then since $-2q$ is a quadratic residue modulo n we have $\left(\frac{-n}{p}\right) = +1$.
- Otherwise, if $p \neq q$ then $2qN = 4q^2x^2 + 4bqxy + 4qhy^2 = (2qx + by)^2 + ny^2$.
- But the only way that this quantity can be divisible by an odd power of p is if there is a nonzero solution to $e^2 + nf^2 \equiv 0 \pmod{p}$, which forces $-n$ to be a quadratic residue modulo p.
- In both cases we have $\left(\frac{n}{p}\right) = +1$ and $\left(\frac{-n}{p}\right) = +1$, so $\left(\frac{-1}{p}\right) = +1$ and so $p \equiv 1 \pmod{4}$.

Sums of Three Squares, IX

Proof (continued even yet still more also finally):

- Now suppose p divides n. Then $R^2 + N = n$, so since p divides N it must also divide R. Rewriting the equation as $R^2 + \frac{1}{2q}\left[(2qx + by)^2 + ny^2\right] = n$, we see that p must also divide $2qx + by$. Dividing through by p and then reducing modulo p yields $\frac{1}{2q} \cdot \frac{n}{p} y^2 \equiv \frac{n}{p} \pmod{p}$, so since n/p is nonzero modulo p as n is squarefree, we get $y^2 \equiv 2q \pmod{p}$ and thus $\left(\frac{2q}{p}\right) = +1$. Since we assumed at the very beginning that $\left(\frac{-2q}{p}\right) = +1$, this implies $\left(\frac{-1}{p}\right) = +1$ and so $p \equiv 1 \pmod{4}$ once again.
- This takes care of all cases, so we are finally done.

Sums of Three Squares, X

The proof for the case $n \equiv 3 \pmod 8$ can be adapted to establish the other cases $n \equiv 1, 2, 5, 6 \pmod 8$ as well, by suitable minor modifications on the conditions taken at the beginning.

- In these cases, we instead take q to be a prime with $q \equiv 1 \pmod 4$ such that $-q$ is a quadratic residue modulo n. If n is even we also take q such that $\left(\frac{-2}{q}\right) = (-1)^{(m-2/4)}$ (this imposes a condition on q mod 8).
- Also take t odd such that $t^2 \equiv -1/q \pmod n$ and set $b^2 - qh = -n$.
- Then we apply Minkowski's theorem to the lattice spanned by $\langle tq, \sqrt{q}, 0\rangle$, $\langle tb, b/\sqrt{q}, \sqrt{n/q}\rangle$, and $\langle m, 0, 0\rangle$.
- Following the same argument through shows that $R^2 + S^2 + T^2 = n$ and $S^2 + T^2 = N$ is a sum of two squares.

Ideal Class Groups, I

We will now discuss some additional properties of the ideals in quadratic integer rings.

- Our task today is to introduce the ideal class group, and then use some geometric methods to establish that the ideal class group of any quadratic integer ring is finite and provide methods for computing it.
- In particular we will then use our results to compute explicitly the ideal class groups of various quadratic integer rings.
- Next week, we will use Minkowski's theorem to write down stronger results.

Ideal Class Groups, II

As we have already discussed, a quadratic integer ring $\mathcal{O}_{\sqrt{D}}$ has unique factorization if and only if it is a principal ideal domain, and (thus) any examples of non-unique factorization necessarily arise from nonprincipal ideals.

- Our goal now is to quantify more precisely how "non-unique" the non-unique factorization in $\mathcal{O}_{\sqrt{D}}$ can be, which is (in a sense we will make precise) the same as asking about the various possible classes of nonprincipal ideals.

Let me give some motivation by working (once again) with our standard example of a quadratic integer ring with non-unique factorizations, namely $\mathbb{Z}[\sqrt{-5}]$.

Ideal Class Groups, III

We have shown that $\mathbb{Z}[\sqrt{-5}]$ is not a PID by constructing explicit nonprincipal ideals $I_2 = (2, 1+\sqrt{-5})$, $I_3 = (3, 1+\sqrt{-5})$, and $I_3' = (3, 1-\sqrt{-5})$.

- Notice, however, that the pairwise products of these nonprincipal ideals are all principal:
  $I_2^2 = (4, 2(1+\sqrt{-5}), -4+2\sqrt{-5}) = (2)$,
  $I_2 I_3 = (6, 2(1+\sqrt{-5}), 3(1+\sqrt{-5}), -4+2\sqrt{-5}) = (1+\sqrt{-5})$,
  $I_2 I_3' = (1-\sqrt{-5})$ since it is just the conjugate of $I_2 I_3$, and
  $I_3^2 = I_3 I' = (I_3')^2 = (3)$.
- This seems like it might just be a coincidence, so let's find some more nonprincipal ideals in $\mathbb{Z}[\sqrt{-5}]$.

Ideal Class Groups, IV

By factoring other integer primes, we can cook up some more nonprincipal ideals: for example $I_7 = (7, 3 + \sqrt{-5})$ and its conjugate $I_7' = (7, 3 - \sqrt{-5})$.

- But if we compute products, like $I_2 I_3$ or $I_3 I_7'$, we will discover that no matter which pair of ideals we multiply together, the result will always be principal. For example, $I_2 I_7 = (14, 6 + 2\sqrt{-5}, 7 + 7\sqrt{-5}, -2 + 4\sqrt{-5})$, so it contains $3 + \sqrt{-5} = 2(6 + 2\sqrt{-5}) - (7 + 7\sqrt{-5}) - (-2 + 4\sqrt{-5})$ and also each element in the ideal is divisible by $3 + \sqrt{-5}$, so in fact $I_2 I_7 = (3 + \sqrt{-5})$.
- Similarly, for $I_3 I_7' = (3, 1 + \sqrt{-5})(7, 3 - \sqrt{-5}) = (21, 9 - 3\sqrt{-5}, 7 + 7\sqrt{-5}, 8 + 2\sqrt{-5})$, we see this ideal contains $4 + \sqrt{-5} = 21 + (7 + 7\sqrt{-5}) - 3(8 + 2\sqrt{-5})$ and also each element in the ideal is divisible by $4 + \sqrt{-5}$, so in fact $I_3 I_7' = (4 + \sqrt{-5})$.

Ideal Class Groups, V

These calculations suggest that there might actually be only one type of nonprincipal ideal in $\mathbb{Z}[\sqrt{-5}]$, up to an appropriate notion of equivalence of ideals.

- If you're not convinced yet, you can try finding other primes that split in $\mathbb{Z}[\sqrt{-5}]$.
- You can use quadratic reciprocity to identify these, since $\left(\frac{-5}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{5}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{5}\right)$, so the primes that split are precisely those congruent to 1, 3, 7, or 9 modulo 20.
- For example, $(23) = (23, 8+\sqrt{-5})(23, 8-\sqrt{-5})$, and neither of these ideals I_{23}, I'_{23} is principal.
- However, in fact $I_{23}I_2 = (1+3\sqrt{-5})$ is principal.

There is also a natural composition operation on ideals, given by ideal multiplication, which seems to behave nicely with respect to this equivalence.

Ideal Class Groups, VI

We will now make all of this precise:

Definition

Let $R = \mathcal{O}_{\sqrt{D}}$ be a quadratic integer ring. We define a relation $\sim$ on the set of nonzero ideals of R by saying $I \sim J$ if $(a)I = (b)J$ for some nonzero principal ideals (a) and (b).

- Intuitively, we declare two ideals to be equivalent if they differ by a principal ideal factor.
- Example: Inside $\mathbb{Z}[i]$, since every nonzero ideal I is principal, we have $I \sim (1)$ for all such I.
- Example: Inside $\mathcal{O}_{\sqrt{-5}}$, with the notation as previously, we have $I_2 \sim I_3$: since $I_2^2 = (2)$ and $I_2 I_3 = (1 + \sqrt{-5})$, we see that $(1 + \sqrt{-5})I_2^2 = (2)I_2 I_3$ and thus cancelling I_2 gives $(1 + \sqrt{-5})I_2 = (2)I_3$.

Ideal Class Groups, VII

The relation $\sim$ is, perhaps unsurprisingly, an equivalence relation, and we can also use it to detect whether $\mathcal{O}_{\sqrt{D}}$ is a PID:

Proposition (Properties of Ideal Classes)

Suppose $R = \mathcal{O}_{\sqrt{D}}$ is a quadratic integer ring. Then the following properties hold for the relation $I \sim J$ if $(a)I = (b)J$ for some nonzero $a, b \in R$:

1. *The relation $\sim$ is an equivalence relation on nonzero ideals. The equivalence classes of this relation are called ideal classes.*
2. *We have $I \sim (1)$ if and only if I is principal. Thus, $\mathcal{O}_{\sqrt{D}}$ is a principal ideal domain if and only if $I \sim (1)$ for all nonzero ideals I of $\mathcal{O}_{\sqrt{D}}$.*
3. *Multiplication of ideals respects ideal classes: if $I \sim I'$ and $J \sim J'$, then $IJ \sim I'J'$.*

Ideal Class Groups, VIII

1. The relation is an equivalence relation on nonzero ideals.

Proof:

- Clearly $I \sim I$ since $(1)I = (1)I$.
- Also, if $I \sim J$ then $(a)I = (b)J$, and then by interpreting this as $(b)J = (a)I$ we see $J \sim I$.
- Finally, if $I \sim J$ and $J \sim K$ then $(a)I = (b)J$ and $(c)J = (d)K$, and so $(ac)I = (bc)J = (bd)K$ meaning $I \sim K$.

Ideal Class Groups, IX

2. We have $I \sim (1)$ if and only if I is principal. Thus, $\mathcal{O}_{\sqrt{D}}$ is a principal ideal domain if and only if $I \sim (1)$ for all nonzero ideals I of $\mathcal{O}_{\sqrt{D}}$.

Proof:

- If $I \sim (1)$ then $(a)I = (b)$ for some nonzero a and b. This equality requires that a divides b, say with $b = ka$. Then cancelling (a) yields $I = (k)$, so I is principal.
- The second statement follows immediately from the first.

---

3. Multiplication of ideals respects ideal classes: if $I \sim I'$ and $J \sim J'$, then $IJ \sim I'J'$.

Proof:

- Suppose $(a)I = (b)I'$ and $(c)J = (d)J'$. Multiplying these relations yields $(ac)IJ = (bd)I'J'$, so $IJ \sim I'J'$.

Ideal Class Groups, X

We have a natural multiplication operation on ideals, which makes the set of nonzero ideals into a semigroup.

- Because the multiplication of ideals respects ideal classes, the set of ideal classes inherits this multiplication operation.
- But things become even better with ideal classes, because the operation makes the set of ideal classes into an actual group, rather than just a semigroup!

Proposition (The Ideal Class Group)

Let $R = \mathcal{O}_{\sqrt{D}}$ be a quadratic integer ring and let $[I]$ represents the ideal class of an ideal I of R. Then the operation $[I] \cdot [J] = [IJ]$ makes the set of ideal classes into an abelian group. This group is called the ideal class group of R (often, just the class group of R).

Ideal Class Groups, XI

Proof:

- First, multiplication of ideal classes is well-defined by (3) from the proposition earlier.
- The operation is associative and commutative because multiplication of ideals is associative and commutative: $([I][J])[K] = [IJ][K] = [IJK] = [I][JK] = [I]([J][K])$ and $[I][J] = [IJ] = [JI] = [J][I]$.
- The ideal class of (1) is a multiplicative identity, since $(1)I = I$ and so $[(1)][I] = [I]$ for all I.
- Finally, every ideal class has an inverse: as we proved, for any ideal I the product $I \cdot \bar{I}$ is a principal ideal (a), and so $[I][\bar{I}] = [(a)] = [(1)]$.

Ideal Class Groups, XII

We see that the ideal classes have the structure of an abelian group under multiplication.

- By itself, this fact does not yield very much useful information about the ideal classes.
- What we really want to do is compute the structure of the ideal class group.
- In particular, it would be quite nice if this group were finitely generated, or (even better) finite, since we can say lots of things about the structure of finite(ly generated) abelian groups.

Ideal Class Groups, XIII

In fact, the ideal class group of $\mathcal{O}_{\sqrt{D}}$ *is* always finite:

Theorem (Properties of the Class Group)

Suppose $R = \mathcal{O}_{\sqrt{D}}$ is a quadratic integer ring and let $[I]$ denote the ideal class of an ideal I of R. Then the following are true:

1. *If I is a nonzero ideal of R, then I contains a nonzero element α such that $N(\alpha) \leq (|D| + 1)N(I)$.*
2. *Every ideal class of R contains an ideal J such that $N(J) \leq |D| + 1$.*
3. *The ideal class group of $\mathcal{O}_{\sqrt{D}}$ is finite.*

The headline result is (3), but in fact (2) is even better than (3), because it actually allows us to compute the class group. (I will do examples next time.)

Ideal Class Groups, XIV

1. If I is a nonzero ideal of R, then I contains a nonzero element α such that $N(\alpha) \leq (|D|+1)N(I)$.

Proof:

- Let $m = \lfloor \sqrt{N(I)} \rfloor$ so that $m^2 \leq N(I) < (m+1)^2$.
- Then since the cardinality of R/I is $N(I) < (m+1)^2$, by the pigeonhole principle at least two of the $(m+1)^2$ elements $\{a + b\sqrt{D} : 0 \leq a, b \leq m\}$ in R must be congruent modulo I, so their difference is in I.
- Thus, there exists a nonzero element $\gamma \in I$ of the form $a + b\omega$ where $-m \leq a, b \leq m$.
- Then $N(\gamma) = (a + b\sqrt{D})(a - b\sqrt{D}) = a^2 - Db^2 \leq |a^2| + |Db^2| = m^2(|D|+1) \leq (|D|+1)N(I)$, as claimed.

When $D \equiv 1 \pmod 4$ this bound can be improved by working instead with the elements of the form $a + b\omega$. But we will do better next week anyway.

Ideal Class Groups, XV

2. Every ideal class of R contains an ideal J such that $N(J) \leq |D| + 1$.

Proof:

- Let $\mathcal{C}$ be an ideal class and let I be any ideal in the inverse class $\mathcal{C}^{-1}$.
- By (1), there exists a nonzero element $\alpha \in I$ such that $N(\alpha) \leq (|D| + 1)N(I)$. Because $\alpha \in I$, by the equivalence of divisibility and containment we see that I divides (α) and so $(\alpha) = IJ$ for some ideal J.
- Taking norms yields $N(\alpha) = N(I)N(J)$, so $N(J) = \dfrac{N(\alpha)}{N(I)} \leq |D| + 1$. Finally, taking ideal classes gives $[1] = [(\alpha)] = [I][J]$ so $J \in [I]^{-1} = (\mathcal{C}^{-1})^{-1} = C$, as required.

Ideal Class Groups, XVI

3. The ideal class group of $\mathcal{O}_{\sqrt{D}}$ is finite.

Proof:

- By (2), every ideal class contains some ideal J with $N(J) \leq |D| + 1$.
- But there are only finitely many possible ideals J with $N(J) \leq |D| + 1$: there are only finitely many possible prime ideals that could occur in the prime factorization of J (namely, the primes of norm at most $|D| + 1$) and the power to which each such ideal can occur is bounded (since a prime power P^a has norm $N(P)^a$, we must have $a \leq \log_{N(P)}(|D| + 1)$ for all such P).
- Thus, the ideal classes are all represented by a finite list of ideals, so there are finitely many ideal classes.

Summary

We used Minkowski's theorem to classify the integers that are the sum of three squares.

We introduced the ideal class group of a quadratic integer ring and proved that it is always a finite abelian group.

Next lecture: Computing ideal class groups.