Algebraic Number Theory Lecture Notes

George Boxer

January 20, 2025

Contents

1 Lecture 1: Unique factorization and Fermat's Last theorem

One of the goals of number theory is to solve Diophantine equations, that is find all solutions to polynomial equations like

$$y^2 = x^3 + 5, \quad x^n + y^n = z^n$$

where the variables are either integers or rational numbers. This is in general very difficult! You are probably familiar with the following famous theorem, known as "Fermat's last theorem"

Theorem 1.1 (Wiles, 1994)**.** *If $x^n + y^n = z^n$ for $x, y, z \in \mathbf{Z}$ for $x, y, z \in \mathbf{Z}$ and $n \geq 3$ then $xyz = 0$.*

This is in stark contrast to the case $n = 2$, when the equation $x^2 + y^2 = z^2$ arises naturally from the Pythagorean theorem, and there are many solutions: $3^2+4^2 = 5^2, 5^2+12^2 = 13^2, \ldots$. We will review in a moment how to find all such solutions.

Pierre de Fermat was a French lawyer and bureaucrat, who, in his free time, was also one of the greatest mathematicians and number theorists of the 17th century. As the story goes, he stated Theorem 1.1 in the margin of his copy of Diophantus' Arithmetica and wrote that he had a proof that was too large to fit in the margin. This problem subsequently became widely known to mathematicians, and was a major source of motivation for the development of algebraic number theory over hundreds of years.

As a warmup, we will consider the case $n = 2$ and recall how unique factorization can be used to find all solutions to $x^2 + y^2 = z^2$ with $x, y, z \in \mathbf{Z}$.

First we make some reductions. We may assume that $x, y, z > 0$, as if (x, y, z) is a solution then so is $(\pm x, \pm y, \pm z)$, and moreover the solutions where one of x, y, z is 0 are just $(\pm a, 0, \pm a)$ and $(0, \pm a, \pm a)$. Next we may assume x, y, z are pairwise coprime. Indeed if some prime p divides two of x, y, z then it also divides the third, and $(x/p, y/p, z/p)$ is a solution. A solution where x, y, z are pairwise coprime is called a primitive pythagorean triple, and the imprimitive ones may be described as (dx, dy, dz) for $d > 0$ and (x, y, z) primitive.

Finally we note that if $n \in \mathbf{Z}$, then $n^2 \equiv 0, 1 \pmod 4$. Since we are assuming that not all of x, y, z are even, we see that z must be odd and exactly one of x or y must be odd. Thus upon possibly switching x and y we may assume that x is even and y, z are odd.

To summarize our reductions, we are looking for $x, y, z > 0$ pairwise coprime with x even and y, z odd satisfying $x^2 + y^2 = z^2$.

Now the idea is to rewrite the equation $x^2 + y^2 = z^2$ in the following manner:

$$y^2 = z^2 - x^2 = (z - x)(z + x).$$

We first claim that the two factors $z - x$ and $z + x$ on the right are coprime. Indeed if p is some prime and $p|z - x$ and $p|z + x$ then

$$p|(z + x) + (z - x) = 2z, \quad p|(z + x) - (z - x) = 2x.$$

As z and x have no common factors by assumption we see that the only possibility is $p = 2$. But we also assumed z is odd and x is even, hence $z + x, z - x$ are odd, so $p = 2$ is also not possible. This proves the claim.

Now we apply unique factorization. We have $z - x, z + x > 0$ and their product is a square. It follows that $z - x = m^2$, $z + x = n^2$ must both be squares. We now solve and find

$$x = \frac{n^2 - m^2}{2}, \quad y = nm, \quad z = \frac{n^2 + m^2}{2}.$$

We can check moreover that in order for $x, y, z > 0$ and to have no common factors, we should take $n > m > 0$ and n, m should both be odd and have no common factors. For example then $(n, m) = (3, 1)$ gives $4^2 + 3^2 = 5^2$, $(n, m) = (5, 1)$ gives $12^2 + 5^2 = 13^2$, $(n, m) = (5, 3)$ gives $8^2 + 15^2 = 17^2$, etc.

Returning to Fermat's last theorem, it is enough to treat the case $n = 4$, which is not too hard and was proved by Fermat, and the case that $n = p$ is an odd prime. Indeed, once FLT holds for some exponent n it holds for any multiple of n.

So for now on we consider the case of $n = p$ an odd prime. Now the basic idea is to mimic the strategy for finding pythagorean triples and consider a factorization

$$x^p + y^p = (x + y)(x + \zeta_p y)(x + \zeta_p^2 y) \cdots (x + \zeta_p^{p-1} y) = z^p$$

where $\zeta_p = e^{\frac{2\pi i}{p}}$ is a primitive pth root of unity. Now we are looking for $x, y, z \in \mathbf{Z}$, but the factors in the above formula won't be integers. They will lie in the ring

$$\mathbf{Z}[\zeta_p] = \left\{ \sum_{i=0}^{p-2} a_i \zeta_p^i \mid a_i \in \mathbf{Z} \right\}.$$

One might hope then that the ring $\mathbf{Z}[\zeta_p]$ has unique factorization. If it did, and the factors $x + y, x + \zeta_p y, \ldots, x + \zeta_p^{p-1} y$ had no common factors, then one might hope to obtain a formula like

$$x + \zeta_p y = u\alpha^p$$

for some $\alpha \in \mathbf{Z}[\zeta_p]$ and unit $u \in \mathbf{Z}[\zeta_p]^\times$, and one might then hope to use this to understand all solutions to the original equation $x^p + y^p = z^p$ as in the case of pythagorean triples.

Unfortunately there is no reason to expect $\mathbf{Z}[\zeta_p]$ to have unique factorization in general, and you may have seen examples of rings like $\mathbf{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbf{Z}\}$ not having unique factorization in your algebra class:

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

It was already discovered in the mid nineteenth century that $\mathbf{Z}[\zeta_{23}]$ does not have unique factorization, and unfortunately it turns out that $\mathbf{Z}[\zeta_p]$ has unique factorization if and only if $p \leq 19$.

Nonetheless, Kummer had the remarkable insight that one could partially recover from the failure of unique factorization in $\mathbf{Z}[\zeta_p]$ by working with ideals. More precisely what Kummer realized, and Dedekind further developed, was that in rings like $\mathbf{Z}[\zeta_p]$ that we shall study in this class, you always have unique factorization into *prime ideals*. In fact, this discovery predates the development of abstract ring theory and was the reason that ideals were defined in the first place. It also explains the name: ideals, originally called ideal elements by Kummer, behaved the way elements ideally would! In particular unique factorization is recovered.

With this idea in hand, Kummer was able to make progress on FLT for many more values of p, including $p = 23$. We will return to this much later in this course, but roughly Kummer's insights were:

- If the factors $x + y, x + \zeta_p y, \ldots, x + \zeta_p^{p-1} y$ are coprime, then we always have

  $$(x + \zeta_p y) = I^p$$

  for some ideal $I \subset \mathbf{Z}[\zeta_p]$. Here this is an equality of ideals, where $(x + \zeta_p y) = \mathbf{Z}[\zeta_p](x + \zeta_p y)$ is the principal ideal generated by $x + \zeta_p y$. (We will recall ideal theory, and in particular what I^p even means in this formula, later in the course.)

- There is an important invariant of the ring $\mathbf{Z}[\zeta_p]$ called the *ideal class group* $\mathrm{Cl}(\mathbf{Z}[\zeta_p])$ of $\mathbf{Z}[\zeta_p]$. To define it, we put an equivalence relation on nonzero ideals in $\mathbf{Z}[\zeta_p]$: we let $I \sim J$ if there exists $0 \neq \alpha, \beta \in \mathbf{Z}[\zeta_p]$ with $\alpha I = \beta J$. It is easy to check that this is an equivalence relation, and that furthermore writing $[I]$ for the equivalence class

of I, the formula $[I] \cdot [J] = [IJ]$ gives a well defined, commutative and associative binary operation on ideal classes. Moreover the principal ideals form an ideal class, which is an identity element for multiplication.

What is not at all obvious is that this actually forms an Abelian group (the existence of inverses is the hard part!) and that it is finite. The class group $\mathrm{Cl}(\mathbf{Z}[\zeta_p])$ is then in a sense a more refined way of measuring the failure of unique factorization in $\mathrm{Cl}(\mathbf{Z}[\zeta_p])$, in particular $\mathrm{Cl}(\mathbf{Z}[\zeta_p])$ is trivial if and only if $\mathbf{Z}[\zeta_p]$ has unique factorization.

Returning to Fermat, the equation

$$(x + \zeta_p y) = I^p$$

expresses that the ideal class $[I]$ has order 1 or p in $\mathrm{Cl}(\mathbf{Z}[\zeta_p])$. To conclude that $I = (\alpha)$ is a principal ideal, we don't really need unique factorization, it would suffice for $\mathrm{Cl}(\mathbf{Z}[\zeta_p])$ to have no elements of order p, or $p \nmid \#\mathrm{Cl}(\mathbf{Z}[\zeta_p])$, which will happen for many more primes p, including $p = 23$.

Using the above ideas, Kummer was able to show:

Theorem 1.2 (Kummer)**.** *If $p \nmid \#\mathrm{Cl}(\mathbf{Z}[\zeta_p])$ then FLT holds for exponent p.*

Primes p are called regular if $p \nmid \#\mathrm{Cl}(\mathbf{Z}[\zeta_p])$ and irregular otherwise. Unfortunately irregular primes do exist and there are 3 less than 100: $p = 37, 59, 67$. Remarkably, it is an open problem whether infinitely many regular primes exist, although it is expected. It is known that infinitely many primes are irregular.

As an aside, this is not the end of Kummer's work on Fermat's last theorem. He was also able to prove Kummer's criteria: p is regular if and only if p does not divide the numerators of the Bernoulli numbers $B_2, B_4, \ldots, B_{p-3}$. The Bernoulli numbers are defined as the coefficients of the Taylor expansion

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}$$

and the first several are $B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}, B_6 = \frac{1}{42}, B_8 = -\frac{1}{30}, B_{10} = \frac{5}{66}, B_{12} = -\frac{691}{2730}$ and so $p = 691$ is an irregular prime. As another example,

$$B_{32} = \frac{7709321041217}{510} = \frac{37 \cdot 683 \cdot 305065927}{510}$$

and hence $p = 37$ is irregular. It is no coincidence that the Bernoulli numbers are connected to values of the Riemann ζ function, by the following formula of Euler:

$$\zeta(2m) = \sum_{n=1}^{\infty} \frac{1}{n^{2m}} = (-1)^{m+1} \frac{(2\pi)^{2m}}{2(2m)!} B_{2m}$$

You are probably at familiar with at least the case $m = 1$ of this formula: $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

2 Lecture 2: Algebraic integers

Proposition 2.1. *Let $\alpha \in \mathbf{C}$ be a nonzero algebraic number with minimal monic polynomial $m_\alpha \in \mathbf{Q}[x]$. Then α is an algebraic integer if and only if $m_\alpha \in \mathbf{Z}[x]$.*

Proof. Since $m_\alpha(\alpha) = 0$, if $m_\alpha \in \mathbf{Z}[x]$ then α is an algebraic integer. For the converse, suppose α is an algebraic integer and let $f \in \mathbf{Z}[x]$ be a monic polynomial with $f(\alpha) = 0$. Then by the definition of the minimal monic polynomial we have a factorization

$$f = g \cdot m_\alpha$$

for some polynomial $g \in \mathbf{Q}[x]$. Furthermore g is monic as f and m_α are. We conclude by applying the following lemma. □

Lemma 2.2 (Gauss' Lemma)**.** *Let $g, h \in \mathbf{Q}[x]$ be monic polynomials and suppose that $g \cdot h \in \mathbf{Z}[x]$. Then $g, h \in \mathbf{Z}[x]$.*

Proof. Let $a \geq 1$ be the smallest integer such that $ag \in \mathbf{Z}[x]$. We claim that the greatest common divisor of the coefficients of ag is 1.[1] Indeed if some prime p divides all the coefficients of ag, then in particular it divides the leading coefficient a of ag (here we use g is monic!) and so $\frac{a}{p}$ is an integer smaller than a with $\frac{a}{p}g \in \mathbf{Z}[x]$. Similarly let $b \geq 1$ be the smallest integer such that $bh \in \mathbf{Z}[x]$. Again the greatest common divisor of the coefficients of bh is 1.

If $a = b = 1$ then in fact we have $g, h \in \mathbf{Z}[x]$ as claimed. If not, $ab > 1$ and so we can consider a prime factor p of ab. We can consider the equality

$$(ag)(bh) = abgh$$

in $\mathbf{Z}[x]$, and reduce the coefficients mod p to obtain an equality in $(\mathbf{Z}/p\mathbf{Z})[x]$. The right hand side is 0 because $gh \in \mathbf{Z}[x]$ and p divides ab, while we observed before that the two factors, ag and bh are nonzero mod p. This contradicts the fact that $(\mathbf{Z}/p\mathbf{Z})[x]$ is an integral domain.[2] □

The proposition gives as an effective way to check if an algebraic number is an algebraic integer. Rather than having to check whether for all polynomials $f \in \mathbf{Z}[x]$, $f(\alpha) = 0$, we can simply determine the minimal monic polynomial m_α and see if its coefficients are integers.

Example 2.3 (Rational algebraic integers)**.** When is a nonzero rational number $\alpha \in \mathbf{Q}$ an algebraic integer? Well, its minimal monic polynomial is $m_\alpha = x - \alpha$, and so by the proposition α is an algebraic integer if and only if α is an integer, as we might have expected![3]

[1] A polynomial $f \in \mathbf{Z}[x]$ is called primitive if the gcd of its coefficients is 1. Gauss' Lemma is commonly formulated as "If $f, g \in \mathbf{Z}[x]$ are primitive polynomials then fg is also primitive". In our version of the lemma, this is the situation we are in when we consider $(ag)(bh)$. You likely saw some version of Gauss' lemma in your algebra course. It is used to prove that if a polynomial $f \in \mathbf{Z}[x]$ factors in $\mathbf{Q}[x]$, then it also factors in $\mathbf{Z}[x]$, and to prove that $\mathbf{Z}[x]$ is a UFD.

[2] Recall from algebra that if R is an integral domain then $R[x]$ is also an integral domain. Indeed you see this by considering leading terms: $(ax^n + \cdots)(bx^m + \cdots) = abx^{n+m} + \cdots$

[3] The "rational root test" from algebra states that if $f = x^n + \cdots + d \in \mathbf{Z}[x]$ is a polynomial with $d \neq 0$, and if $\alpha \in \mathbf{Q}$ with $f(\alpha) = 0$, then in fact $\alpha \in \mathbf{Z}$ and $\alpha \mid d$. This gives us another way to conclude that rational numbers which are algebraic integers are in fact integers.

2.1 Quadratic integer rings

In this long example, we will work out the quadratic algebraic integers. This will be a fundamental example that we will return to repeatedly throughout this course. We could of course consider quadratic polynomial $x^2 + ax + b$ with $a, b \in \mathbf{Z}$ and find its roots using the quadratic formula, but it will be better to be a bit more organized.

By the quadratic formula, any quadratic algebraic number will lie in a quadratic field

$$\mathbf{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbf{Q}\}$$

where $d \in \mathbf{Q}^\times$.[4] If $a \in \mathbf{Q}^\times$, then $\mathbf{Q}(\sqrt{d}) = \mathbf{Q}(\sqrt{a^2 d})$. Thus it suffices to consider $\mathbf{Q}(\sqrt{d})$ for $d \in \mathbf{Z}$, $d \neq 0$, and d squarefree. Here squarefree means that for each prime p, $p^2 \nmid d$, or in other words in the factorization of d into primes, no prime factor occurs more than once. Moreover as $\mathbf{Q}(\sqrt{1}) = \mathbf{Q}$, we should also assume $d \neq 1$ in what follows.

Exercise 2.4. Verify that for $d \in \mathbf{Z}$, $d \neq 0, 1$, d squarefree, $\mathbf{Q}(\sqrt{d}) \neq \mathbf{Q}$, and moreover if $d_1, d_2 \in \mathbf{Z}$, $d_1, d_2 \neq 0, 1$, d_1, d_2 squarefree, then $\mathbf{Q}(\sqrt{d_1}) \neq \mathbf{Q}(\sqrt{d_2})$.

So now let us fix $d \neq 0, 1$ squarefree and ask, what are the algebraic integers in $\mathbf{Q}(\sqrt{d})$?

Before answering this we recall some notation.

Definition 2.5. Conjugation on $\mathbf{Q}(\sqrt{d})$ is defined by the formula

$$\overline{a + b\sqrt{d}} = a - b\sqrt{d}.$$

For $\alpha, \beta \in \mathbf{Q}(\sqrt{d})$ we have the following formulas:

- $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$.
- $\overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$.
- $\overline{\overline{\alpha}} = \alpha$.

The first two of these formulas say that conjugation, viewed as a function $\mathbf{Q}(\sqrt{d}) \to \mathbf{Q}(\sqrt{d})$, is a ring homomorphism. Furthermore it is clearly a bijection, and so it is an *automorphism* of $\mathbf{Q}(\sqrt{d})$. The third property says that conjugation has order two, so it is an *involution*.

We note that the elements $\alpha \in \mathbf{Q}(\sqrt{d})$ with $\overline{\alpha} = \alpha$ are exactly $\alpha \in \mathbf{Q}$. This may help motivate the following definitions:

Definition 2.6. The norm and trace of $\alpha \in \mathbf{Q}(\sqrt{d})$ are defined by the formulas

$$\mathrm{tr}(\alpha) = \alpha + \overline{\alpha}, \qquad N(\alpha) = \alpha\overline{\alpha}.$$

From the properties of conjugation recalled above, we immediately deduce that for $\alpha \in \mathbf{Q}(\sqrt{d})$, $N(\alpha), \mathrm{tr}(\alpha) \in \mathbf{Q}$ and for $\alpha, \beta \in \mathbf{Q}(\sqrt{d})$,

$$\mathrm{tr}(\alpha + \beta) = \mathrm{tr}(\alpha) + \mathrm{tr}(\beta), \qquad N(\alpha\beta) = N(\alpha)N(\beta).$$

[4] Here $\sqrt{d}$ is supposed to be an element of $\mathbf{C}$, and so we should really specify *which* square root of d we mean. We adopt the convention that for $d \in \mathbf{R}$, if $d \geq 0$ then $\sqrt{d} \geq 0$, while if $d < 0$ then $\sqrt{d}$ is defined to be $\sqrt{-d}i$. However this choice doesn't really matter, as should be clarified by the discussion of conjugation.

We also observe the explicit formulas

$$\mathrm{tr}(a+b\sqrt{d}) = 2a, \qquad N(a+b\sqrt{d}) = a^2 - db^2.$$

For $\alpha \in \mathbf{Q}(\sqrt{d})$ we compute

$$0 = \alpha^2 - (\alpha+\overline{\alpha})\alpha + \alpha\overline{\alpha} = \alpha^2 - \mathrm{tr}(\alpha)\alpha + N(\alpha).$$

In other words, α is a root of the polynomial $x^2 - \mathrm{tr}(\alpha)x + N(\alpha) \in \mathbf{Q}[x]$. If $\alpha \notin \mathbf{Q}$, then the minimal monic polynomial of α cannot have degree one, and so this degree two polynomial must be the minimal monic polynomial of α.

In other words, by the proposition, $\alpha = a + b\sqrt{d} \in \mathbf{Q}(\sqrt{d})$ is an algebraic integer if and only if $\mathrm{tr}(\alpha), N(\alpha) \in \mathbf{Z}$, or more explicitly if and only if $2a, a^2 - db^2 \in \mathbf{Z}$. Let us now find all $a, b \in \mathbf{Q}$ with this property.

First make the substitution $a' = 2a$ and $b' = 2b$. Our condition now becomes $a' \in \mathbf{Z}, (a')^2 - d(b')^2 \in 4\mathbf{Z}$. We first claim that these conditions imply $b' \in \mathbf{Z}$. Indeed they imply $d(b')^2 \in \mathbf{Z}$, and now the fact that d is square free implies $b' \in \mathbf{Z}$.

Now we must find all solutions to the congruence

$$(a')^2 - d(b')^2 \equiv 0 \pmod 4.$$

We note that if n is even then $n^2 \equiv 0 \pmod 4$ and if n is odd then $n \equiv 1 \pmod 4$. We also note that $d \not\equiv 0 \pmod 4$ because d is squarefree. By considering all the possibilities we find that the solutions are

- Either a', b' are even and hence the original $a, b \in \mathbf{Z}$, or
- a', b' are odd and $d \equiv 1 \pmod 4$.

Exercise 2.7. Check this!

To summarize, we have found that the set of algebraic integers in the field $\mathbf{Q}(\sqrt{d})$. They are:

$$\mathcal{O}_d = \begin{cases} \{a+b\sqrt{d} \mid a, b \in \mathbf{Z}\} & d \not\equiv 1 \pmod 4 \\ \left\{ \frac{a'+b'\sqrt{d}}{2} \mid a', b' \in \mathbf{Z}, a' \equiv b' \pmod 2 \right\} & d \equiv 1 \pmod 4 \end{cases}$$

In particular we observe that when $d \not\equiv 1 \pmod 4$, $\mathcal{O}_d$ is the ring $\mathbf{Z}[\sqrt{d}]$, while when $d \equiv 1 \pmod 4$ we note that

$$\left(\frac{1+\sqrt{d}}{2}\right)^2 = \frac{\frac{1+d}{2} + \sqrt{d}}{2} \in \mathcal{O}_d$$

and so

$$\mathcal{O}_d = \mathbf{Z}\left[\frac{1+\sqrt{d}}{2}\right] = \left\{ a + b\frac{1+\sqrt{d}}{2} \mid a, b \in \mathbf{Z} \right\}$$

is also a ring.

The rings $\mathcal{O}_d$ are the rings of quadratic algebraic integers and they will be a fundamental example throughout the course. We will revisit them shortly once we have developed a bit more theory.

Example 2.8. When $d = -1$, $\mathcal{O}_{-1} = \mathbf{Z}[i]$ is called the Gaussian integers. When $d = -3$, $\mathcal{O}_{-3} = \mathbf{Z}[\zeta_6] = \mathbf{Z}[\zeta_3]$ are called the Eisenstein integers. Note that $\zeta_6 = e^{2\pi i/6} = \frac{1+\sqrt{-3}}{2}$ and $\zeta_3 = e^{2\pi i/3} = \frac{-1+\sqrt{-3}}{2} = \zeta_6 - 1$.

2.2 Another characterization of algebraic integers

Our example of quadratic algebraic integers suggests that if α, β are algebraic integers then $\alpha + \beta$ and $\alpha\beta$ are algebraic integers, or in other words, algebraic integers form a ring. We would now like to prove this. It will be useful to give another characterization of algebraic integers.

Proposition 2.9. *Let $\alpha \in \mathbf{C}$. The following are equivalent:*

1. *α is an algebraic integer.*
2. *$\mathbf{Z}[\alpha]$ is finitely generated as an abelian group under addition.*
3. *There exists a nonzero, finitely generated subgroup $M \subset \mathbf{C}$ such that $\alpha \cdot M \subseteq M$.*[5]

We recall that $\mathbf{Z}[\alpha]$ means the subring of $\mathbf{C}$ generated by α, i.e. the intersection of all subrings of $\mathbf{C}$ containing α, or alternatively it is given explicitly by

$$\mathbf{Z}[\alpha] = \{f(\alpha) \mid f \in \mathbf{Z}[x]\}$$

Indeed any subring of $\mathbf{C}$ containing α must contain α^i for all $i \geq 0$, as well as any $\mathbf{Z}$-linear combination of these elements.

We also recall that to say that an abelian group M is finitely generated means that there are finitely many elements $m_1, \ldots, m_n \in M$ such that any $m \in M$ may be expressed as a $\mathbf{Z}$-linear combination of $m_1, \ldots, m_n$, or more explicitly, $m = a_1 m_1 + \cdots a_n m_n$ for $a_1, \ldots, a_n \in \mathbf{Z}$. We don't demand that this expression is unique.

We emphasize here that generators for $\mathbf{Z}[\alpha]$ as a ring and generators for $\mathbf{Z}[\alpha]$ as an abelian group mean very different things!

Exercise 2.10. Describe $\mathbf{Z}[1/2]$ explicitly. Explain directly why it isn't finitely generated as an abelian group.

Proof. We prove that 1 implies 2. We suppose that $\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0 = 0$ with $a_0, \ldots, a_{n-1} \in \mathbf{Z}$. We claim that $1, \ldots, \alpha^{n-1}$ generate $\mathbf{Z}[\alpha]$ as an abelian group under addition. From the explicit description of $\mathbf{Z}[\alpha]$ above we see that the infinite set $\{\alpha^k \mid k \geq 0\}$ generates $\mathbf{Z}[\alpha]$ under addition. Thus it suffices to show that each α^k for $k \geq n$ may be expressed as a $\mathbf{Z}$-linear combination of $1, \ldots, \alpha^{n-1}$.

For the base case of $k = n$ we have the formula

$$\alpha^n = -a_0 - a_1\alpha - \cdots - a_{n-1}\alpha^{n-1}.$$

Now for the inductive step, supposing we can write $\alpha^k = \sum_{i=0}^{n-1} c_i \alpha^i$, we multiply this expression by α and use the formula for α^n to obtain one for α^{k+1}.

We prove that 2 implies 3. As $\mathbf{Z}[\alpha]$ is a ring and $\alpha \in \mathbf{Z}[\alpha]$, we have $\alpha \cdot \mathbf{Z}[\alpha] \subseteq \mathbf{Z}[\alpha]$. Hence we may take $M = \mathbf{Z}[\alpha]$.

We prove that 3 implies 1. Let $m_1, \ldots, m_n$ generate M. Because $\alpha M \subseteq M$, we have expressions

$$\alpha m_j = \sum_{i=1}^{n} a_{ij}$$

[5] Here $\alpha \cdot M$ means $\{\alpha m \mid m \in M\}$.

for $j = 1, \ldots, n$, and coefficients $a_{ij} \in \mathbf{Z}$ for $1 \leq i, j \leq n$. We can package these coefficients into a matrix $A = (a_{ij}) \in M_n(\mathbf{Z})$. We can view A as giving the matrix for $\alpha \cdot : M \to M$ with respect to the coordinates on M given by the generating set $m_1, \ldots, m_n$. More formally, we have a surjective[6] homomorphism of abelian groups:

$$p : \mathbf{Z}^n \to M$$
$$(a_1, \ldots, a_n) \mapsto \sum_{i=1}^{n} a_i m_i$$

and for $v \in \mathbf{Z}^n$, we have that $\alpha \cdot p(v) = p(Av^T)$. It follows from this that for any polynomial $f \in \mathbf{Z}[x]$ and $v \in \mathbf{Z}^n$, $f(\alpha) \cdot p(v) = p(f(A)v^T)$.

We can consider the characteristic polynomial

$$\chi_A(x) = \det(x \cdot I - A) \in \mathbf{Z}[x]$$

which is furthermore monic (the leading term of the characteristic polynomial of an n by n matrix is always x^n). By the Cayley-Hamilton theorem, $\chi_A(A) = 0$. It follows that $\chi_A(\alpha)M = \{0\}$. In particular since M is nonzero, we can pick any nonzero $m \in M$ and conclude that $\chi_A(\alpha)m = 0$ and hence $\chi_A(\alpha) = 0$. Hence α is an algebraic integer. □

3 Lecture 3: Number fields, norms, traces

We now deduce some consequences of the characterization of algebraic integers from the end of last lecture.

Theorem 3.1. *1. If α, β are algebraic integers, then so are $\alpha + \beta$, $\alpha - \beta$, $\alpha\beta$.*

2. If $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbf{C}[x]$ is a monic polynomial with $a_0, \ldots, a_{n-1}$ algebraic integers, and $\alpha \in \mathbf{C}$ with $f(\alpha) = 0$ then α is an algebraic integer.

We write $\overline{\mathbf{Z}}$ for the set of algebraic integers in $\mathbf{C}$. The first part of the theorem may be rephrased as saying that $\overline{\mathbf{Z}}$ is a ring. The second part of the theorem tells us that we don't get anything new if we consider roots of monic polynomials whose coefficients are algebraic integers.

Proof. We prove part 1. We consider the ring

$$\mathbf{Z}[\alpha, \beta] = \{\sum_{i,j} a_{ij}\alpha^i\beta^j \mid a_{i,j} \in \mathbf{Z}\}$$

It is clearly generated by $\alpha^i\beta^j$ for $i, j \geq 0$. But if α satisfies a monic polynomial of degree n and β satisfies a monic polynomial of degree m, then in fact it is generated by the finitely many elements $\alpha^i\beta^j$ for $0 \leq i < n$ and $0 \leq j < m$, as in the proof of Proposition 2.9. Now for any element $\gamma \in \mathbf{Z}[\alpha, \beta]$, we have $\gamma \cdot \mathbf{Z}[\alpha, \beta] \subseteq \mathbf{Z}[\alpha, \beta]$ and so Proposition

[6]The structure theory of finitely generated abelian groups will tell us that as M is a finitely generated torsion free abelian group it is actually free, that is isomorphic to $\mathbf{Z}^m$ for some m. So we could have chosen the generating set so that it is actually a basis, and p is actually an isomorphism. However it is an interesting feature of this argument, and quite useful for generalizations, that we do not need the generating set to be a basis.

2.9 implies that γ is an algebraic integer. This applies in particular to $\alpha+\beta$, $\alpha-\beta$, and $\alpha\beta$.

We prove part 2. We consider the ring $\mathbf{Z}[a_0,\ldots,a_{n-1},\alpha]$. If a_i satisfies a monic polynomial of degree m_i, we see that the monomials

$$a_0^{i_0}\cdots a_{n-1}^{i_{n-1}}\alpha^j$$

for $0\le i_k<m_k$ for $0\le k<n$ and $0\le j<n$. This is a finite set, hence any element of this ring, and in particular α, is an algebraic integer. □

Now we can make some of the basic definitions for this course:

Definition 3.2. By a number field, we mean a subfield $K\subset\mathbf{C}$ which is finite dimensional as a $\mathbf{Q}$-vector space. The dimension of K as a $\mathbf{Q}$-vector space is called the degree of K and is denoted $[K:\mathbf{Q}]$.

Basic examples are $\mathbf{Q}$, $\mathbf{Q}(\sqrt{d})$, $\mathbf{Q}(\sqrt[3]{2})$, etc. More generally:

Example 3.3. For any algebraic integer $\alpha\in\mathbf{C}$, $\mathbf{Q}(\alpha)$ is a number field. If m_α has degree n, then $1,\alpha,\ldots,\alpha^{n-1}$ form a basis for $\mathbf{Q}(\alpha)$ and hence $[\mathbf{Q}(\alpha):\mathbf{Q}]=n$.[7]

We note that if K is a number field and $\alpha\in K$ then α is an algebraic number. Indeed, there must be a non trivial $\mathbf{Q}$-linear dependence among $1,\alpha,\alpha^2,\ldots$ in the finite dimensional $\mathbf{Q}$-vector space K.

Definition 3.4. Given a number field K we denote the set of algebraic integers in K by $\mathcal{O}_K$. This is a ring by part 1 of the theorem. It is called the *integer ring* of K.

For example, we have seen that $\mathcal{O}_{\mathbf{Q}}=\mathbf{Z}$, $\mathcal{O}_{\mathbf{Q}(i)}=\mathbf{Z}[i]$, etc.

Example 3.5. Let p be prime. We can consider the Cyclotomic field $\mathbf{Q}(\zeta_p)$, where $\zeta_p=e^{2\pi i/p}$ is a primitive pth root of unity. We also considered the ring $\mathbf{Z}[\zeta_p]$. Since ζ_p is an algebraic integer, we have $\mathbf{Z}[\zeta_p]\subseteq\mathcal{O}_{\mathbf{Q}(\zeta_p)}$. Later in this course we shall prove that this is an equality, but this is not at all obvious. But do note that we have proved this for $p=3$ where $\mathbf{Q}(\zeta_3)=\mathbf{Q}(\sqrt{-3})$.

One of our basic goals in algebraic number theory is to understand these rings of integers better. For example we would like to ask questions like:

- Is $\mathcal{O}_K$ a unique factorization domain?
- Does the prime number p factor in $\mathcal{O}_K$? If so how?
- What is the structure of the group of units $\mathcal{O}_K^\times$?

When we studied quadratic integer rings the norm and trace maps turned out to be quite useful. We would like to study them for more general number fields.

Definition 3.6. Let K be a number field. We define maps $N_K,\mathrm{tr}_K:K\to\mathbf{Q}$ called the norm and the trace as follows: for $\alpha\in K$ we can view multiplication by α, $\alpha\cdot:K\to K$ as an endomorphism of the $\mathbf{Q}$-vector space K. Then $N_K(\alpha)$ is defined to be $\det(\alpha\cdot)$, the determinant of this endomorphism, and $\mathrm{tr}_K(\alpha)$ is defined to be $\mathrm{tr}(\alpha\cdot)$ the trace of this endomorphism.

[7] If you have taken Galois theory, you might have seen the primitive element theorem, which says that conversely for any number field K there is $\alpha\in K$ with $K=\mathbf{Q}(\alpha)$. We will not use this result in this course.

We deduce from the corresponding properties of traces and determinants of matrices that for $\alpha, \beta \in K$,

$$\mathrm{tr}_K(\alpha+\beta) = \mathrm{tr}_K(\alpha) + \mathrm{tr}_K(\beta), \quad N_K(\alpha\beta) = N_K(\alpha)N_K(\beta).$$

Remark 3.7. This definition looks different from the one that we gave for quadratic fields. You can check right now that this definition agrees with the other one for quadratic fields, and we will soon explain this more systematically.

Example 3.8. Let $K = \mathbf{Q}(\sqrt[3]{2})$. Let's compute $\mathrm{tr}_K(\sqrt[3]{2})$ and $N_K(\sqrt[3]{2})$. We should find a $\mathbf{Q}$-vector space basis for K and compute the matrix of $\sqrt[3]{2}\cdot$ with respect to this basis. For example we can take the basis $1, \sqrt[3]{2}, \sqrt[3]{4}$ and compute

$$\sqrt[3]{2}\cdot 1 = \sqrt[3]{2}, \quad \sqrt[3]{2}\cdot\sqrt[3]{2} = \sqrt[3]{4}, \quad \sqrt[3]{2}\cdot\sqrt[3]{4} = 2$$

and so the matrix of $\sqrt[3]{2}\cdot$ with respect to this basis is

$$\begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Now we just compute the trace and determinant of this matrix and conclude:

$$\mathrm{tr}_K(\sqrt[3]{2}) = 0, \quad N_K(\sqrt[3]{2}) = 2.$$

We can generalize the previous example with the following calculation of norms and traces.

Proposition 3.9. *Let $\alpha \in \mathbf{C}$ be a nonzero algebraic number with minimal monic polynomial $m_\alpha(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Then*

$$\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) = -a_{n-1}, \quad N_{\mathbf{Q}(\alpha)}(\alpha) = (-1)^n a_0.$$

In particular if α is actually an algebraic integer then $N_{\mathbf{Q}(\alpha)}(\alpha), \mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) \in \mathbf{Z}$.

Proof. We consider the matrix of multiplication by α with respect to the basis $1, \alpha, \ldots, \alpha^{n-1}$ of $\mathbf{Q}(\alpha)$. We compute that $\alpha\cdot\alpha^i = \alpha^{i+1}$ for $i = 0, \ldots, n-2$ and

$$\alpha\cdot\alpha^{n-1} = \alpha^n = -a_0 - a_1\alpha - \cdots - a_{n-1}\alpha^{n-1}.$$

In other words the matrix of multiplication by α with respect to this basis is what is in linear algebra called the companion matrix of the polynomial m_α:

$$C(m_\alpha) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

and we get the formulas of the proposition by computing the determinant and trace of this matrix.[8] □

[8] More generally, if f is a monic polynomial, f is the characteristic (and minimal) polynomial of the companion matrix $C(f)$.

We now give a reinterpretation of the formulas of Proposition 3.9. Let $\alpha \in \mathbf{C}$ be a nonzero algebraic number with minimal monic polynomial m_α. By the fundamental theorem of algebra[9] we can factor m_α into linear factors in $\mathbf{C}[x]$:

$$m_\alpha = (x-\alpha)(x-\alpha_2)\cdots(x-\alpha_n).$$

Here we are using that we know that $\alpha \in \mathbf{C}$ is one of the roots of m_α, and we are calling the other roots $\alpha_2, \ldots, \alpha_n$. The roots $\alpha, \alpha_2, \ldots, \alpha_n$ of m_α are called the *conjugates*[10] of α. Note that as they are also roots of the polynomial m_α, they are algebraic numbers, and they are also algebraic integers if α is. Using Proposition 3.9 we have the formulas:

$$\begin{aligned}\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) &= -a_{n-1} = \alpha + \alpha_2 + \cdots + \alpha_n \\ N_{\mathbf{Q}(\alpha)}(\alpha) &= (-1)^n a_0 = \alpha\alpha_2\cdots\alpha_n\end{aligned}$$

These formulas suddenly look very close to how we defined norms and traces for quadratic fields in the last lecture!

We would like to obtain formulas like these for $\mathrm{tr}_K(\alpha)$ and $N_K(\alpha)$ when $\mathbf{Q}(\alpha) \subsetneq K$. We will achieve this in the next lecture. In preparation, we first recall that m_α, because it is irreducible in $\mathbf{Q}[x]$, cannot have a repeated over $\mathbf{C}$.

If K is a field and $f \in K[x]$ is a polynomial, we can consider the derivative $f' \in K[x]$, defined formally by the familiar formula from calculus:

$$(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0)' = n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

We can check easily that this satisfies the usual identities from calculus: for $f, g \in K[x]$ and $a \in K$,

$$(af)' = af', \quad (f+g)' = f' + g', \quad (fg)' = f'g + fg'.$$

Indeed for the Leibniz rule, by the first two formulas it suffices to consider $f = x^n$, $g = x^m$.

Proposition 3.10. [11] *Let K be a number field and let $f \in K[x]$ be an irreducible polynomial of degree n. Then f has exactly n roots in $\mathbf{C}$.*

[9] Recall that the fundamental theorem of algebra says that any non constant polynomial $f \in \mathbf{C}[x]$ has a root in $\mathbf{C}$. We say that $\mathbf{C}$ is algebraically closed.

[10] One could also be more precise and call them the conjugates of α over $\mathbf{Q}$. More generally, given a number field K, we can consider $m_{\alpha,K} \in K[x]$, the minimal monic polynomial of α over K. That is, for any polynomial $f \in K[x]$ with $f(\alpha) = 0$, we have $m_{\alpha,K} \mid f$. Then in particular $m_{\alpha,K} | m_\alpha$, and they might or might not be equal. Then the complex roots of $m_{\alpha,K}$ are called the conjugates of α over K. They are a subset of conjugates of α over $\mathbf{Q}$.

Here is an example: take $\alpha = \sqrt[3]{2}$. Its minimal polynomial is $m_\alpha = x^3 - 2 = (x - \sqrt[3]{2})(x - \zeta_3\sqrt[3]{2})(x - \zeta_3^2\sqrt[3]{2})$. So over $\mathbf{Q}$, the three roots $\sqrt[3]{2}$, $\zeta_3\sqrt[3]{2}$, $\zeta_3^2\sqrt[3]{2}$ are conjugate. Over $K = \mathbf{Q}(\sqrt[3]{2})$, we have the factorization into irreducibles $x^3 - 2 = (x - \sqrt[3]{2})(x^2 + \sqrt[3]{2}x + \sqrt[3]{4})$. So over K, $\sqrt[3]{2}$ is only conjugate to itself, while $\zeta_3\sqrt[3]{2}$ and $\zeta_3^2\sqrt[3]{2}$ are still conjugate. Understanding what is going on here is the goal of Galois theory.

[11] If K is a field, a polynomial $f \in K[x]$ is called separable if f does not have a repeated root in any field extension $K \subseteq L$. The same proof as in this proposition shows that f is separable if and only if f and f' are coprime. If f is furthermore irreducible, the only way f and f' can fail to be coprime is if $f' = 0$. In particular if K has characteristic 0, this can never happen and so irreducible polynomials over fields of characteristic 0 are always separable.

However for fields of characteristic p this can fail. Here is the standard example you might have seen in one of your other courses: let p be prime, let $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ be the field with p elements, and let $K = \mathbf{F}_p(t) = \mathrm{Frac}(\mathbf{F}_p[t])$ be the field of rational functions in the variable t. Then $f = x^p - t \in K[x]$ is an irreducible polynomial with $f' = 0$, and so f is an example of an irreducible inseparable polynomial. Over the extension $\mathbf{F}_p(t) \subseteq \mathbf{F}_p(t^{1/p})$ the polynomial f factors as $f = (x - t^{1/p})^p$.

Proof. The derivative f' of f has degree $n-1$ (in particular it is nonzero). As $K[x]$ is a PID, we can consider the greatest common divisor of f and f'. As f is irreducible it must be either f or 1. But f cannot divide f' for degree reasons, and hence it is 1. Hence there exist polynomials $g, h \in K[x]$ with

$$1 = gf + hf'.$$

Now suppose that $\alpha \in \mathbf{C}$ is a repeated root of f, or in other words

$$f = (x-\alpha)^2 p$$

for some $p \in \mathbf{C}[x]$. Then we have

$$f' = (x-\alpha)(2p + (x-\alpha)p')$$

and so in particular $f'(\alpha) = 0$. But then plugging α in to the formula $1 = gf + hf'$ gives a contradiction.

To conclude we have shown that $f = (x-\alpha_1)\cdots(x-\alpha_n) \in \mathbf{C}[x]$ where $\alpha_i \neq \alpha_j$ for $i \neq j$, and hence f has exactly n roots $\alpha_1, \ldots, \alpha_n$ in $\mathbf{C}$. □

4 Lecture 4: Embeddings

Last time we explained how to compute $\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha)$, $N_{\mathbf{Q}(\alpha)}(\alpha)$ in terms of the conjugates of α. More precisely if m_α is the minimal monic polynomial of α, with factorization over $\mathbf{C}$,

$$m_\alpha = (x-\alpha)(x-\alpha_2)\cdots(x-\alpha_n) \in \mathbf{C}[x]$$

then

$$\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) = \alpha + \alpha_2 + \cdots + \alpha_n, \quad N_{\mathbf{Q}(\alpha)}(\alpha) = \alpha\alpha_2\cdots\alpha_n.$$

We would like to generalize this to a calculation of $\mathrm{tr}_K(\alpha)$, $\det_K(\alpha)$ for any number field K and $\alpha \in K$. For this (and later purposes) it is helpful to make the following definition.

Definition 4.1. Let K be a number field. By an embedding of K we mean a ring homomorphism[12] $\tau : K \to \mathbf{C}$. We write Σ_K for the set of embeddings of K.[13]

Example 4.2.

1. With our definition of number fields as a subfield $K \subseteq \mathbf{C}$, the identity map $\tau(\alpha) = \alpha$ defines an embedding.
2. If $K = \mathbf{Q}(\sqrt{d})$ is quadratic, there are two embeddings: the identity map $\tau_1(\alpha) = \alpha$ and conjugation $\tau_2(\alpha) = \overline{\alpha}$.

[12] In this course, ring homomorphisms send 1 to 1. Then as K is a field τ is automatically injective: if $x \in K$ is nonzero then there is an inverse x^{-1} of x in K and $1 = \tau(1) = \tau(xx^{-1}) = \tau(x)\tau(x^{-1})$ in $\mathbf{C}$ and hence $\tau(x) \neq 0$. The word "embedding" can mean lots of different things in different areas of mathematics, but in algebra contexts, embedding is usually a synonym for "injective homomorphism".

[13] This notation is not standard. Many other notations for the set of embeddings of a number field K exist.

3. We consider the cubic field $\mathbf{Q}(\sqrt[3]{2})$.[14] The minimal monic polynomial of $\sqrt[3]{2}$ is $x^3 - 2$, which factors over $\mathbf{C}$ as

$$x^3 - 2 = (x - \sqrt[3]{2})(x - \zeta_3\sqrt[3]{2})(x - \zeta_3^2\sqrt[3]{2}) \in \mathbf{C}[x]$$

and hence the conjugates of $\sqrt[3]{2}$ are $\zeta_3\sqrt[3]{2}$ and $\zeta_3^2\sqrt[3]{2}$. We note that unlike the quadratic case, they are not elements of $\mathbf{Q}(\sqrt[3]{2})$. Indeed every element of $\mathbf{Q}(\sqrt[3]{2})$ is real while the conjugates are not.

We shall see below that there are three embeddings of $\mathbf{Q}(\sqrt[3]{2})$, characterized by:

$$\tau_1(\sqrt[3]{2}) = \sqrt[3]{2}, \quad \tau_2(\sqrt[3]{2}) = \zeta_3\sqrt[3]{2}, \quad \tau_3(\sqrt[3]{2}) = \zeta_3^2\sqrt[3]{2}$$

In other words there are three embeddings corresponding to the three conjugates of $\sqrt[3]{2}$.

We generalize the last example:

Proposition 4.3. *Let α be a nonzero algebraic number with minimal monic polynomial m_α of degree n. Then $\#\Sigma_{\mathbf{Q}(\alpha)} = n = [\mathbf{Q}(\alpha) : \mathbf{Q}]$. Moreover for each of the n distinct[15] roots $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$ of m_α there is an embedding $\tau_i : \mathbf{Q}(\alpha) \to \mathbf{C}$ with $\tau_i(\alpha) = \alpha_i$.*

Proof. Recall that we have an isomorphism

$$\begin{aligned} \mathbf{Q}[x]/(m_\alpha) &\to \mathbf{Q}(\alpha) \\ f + (m_\alpha) &\mapsto f(\alpha) \end{aligned}$$

Thus giving a ring homomorphism $\tau : \mathbf{Q}(\alpha) \to \mathbf{C}$ is the same as giving a ring homomorphism $\phi : \mathbf{Q}[x] \to \mathbf{C}$ such that $(m_\alpha) \subseteq \ker(\phi)$, with the relation being that for $f \in \mathbf{Q}[x]$, $\tau(f(\alpha)) = \phi(f)$.

But giving a ring homomorphism $\phi : \mathbf{Q}[x] \to \mathbf{C}$ is the same as giving the element $\beta = \phi(x) \in \mathbf{C}$, as indeed given β we can define $\phi(f) = f(\beta)$. Finally saying that $(m_\alpha) \subseteq \ker(\phi)$ is the same as saying that $0 = \phi(m_\alpha) = m_\alpha(\beta) = 0$.

To summarize, for each root β of m_α in $\mathbf{C}$ there is a unique $\tau : \mathbf{Q}(\alpha) \to \mathbf{C}$ with $\tau(\alpha) = \beta$. □

Now we can rewrite our formula from before using embeddings instead of conjugates:

$$\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) = \sum_{\tau \in \Sigma_{\mathbf{Q}(\alpha)}} \tau(\alpha), \quad N_{\mathbf{Q}(\alpha)}(\alpha) = \prod_{\tau \in \Sigma_{\mathbf{Q}(\alpha)}} \tau(\alpha)$$

This suggests a formula for $\mathrm{tr}_K(\alpha)$, $\det_K(\alpha)$ when $\mathbf{Q}(\alpha) \subsetneq K$ in terms of embeddings.

We finally state the main theorem on embeddings and traces/norms:

Theorem 4.4. *Let K be a number field.*

1. *We have $\#\Sigma_K = [K : \mathbf{Q}]$.*[16]

[14] We should say what we really mean by $\sqrt[3]{2}$. We mean the real cube root of 2.

[15] By proposition 3.10.

[16] This is a special case of a more general result you might have encountered in Galois theory, with essentially the same proof: if $K \subseteq L$ is a finite separable extension and $L \subseteq M$ is a "big enough" extension, then there are $[L : K]$ embeddings $L \to M$ which restrict to the identity on K. Here "big enough" could mean that M is algebraically closed, and in our situation we are taking $M = \mathbf{C}$. However if you look carefully at the proof, really all we really need is that any irreducible polynomial in $K[x]$ which has a root in L has all its roots in M. In field theory we say "M contains a normal closure of the extension $K \subseteq L$". In the very special case that we can take $M = L$, we say M is a normal extension of L. Then the "embeddings" of L into L are actually automorphism, and they form a group called the Galois group of the extension $K \subseteq L$.

2. For $\alpha \in K$ we have[17]

$$\mathrm{tr}_K(\alpha) = \sum_{\tau \in \Sigma_K} \tau(\alpha), \quad N_K(\alpha) = \prod_{\tau \in \Sigma_K} \tau(\alpha).$$

To summarize what we have seen so far, we have proved 1 for $K = \mathbf{Q}(\alpha)$[18] and we have proved 2 in the special case of $\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha)$ and $N_{\mathbf{Q}(\alpha)}(\alpha)$. It turns out that this is the essential case and the essential thing to understand, and what remains to do are some reductions to this case.

We begin with some preparations. First we recall that degrees are multiplicative in towers of fields.

Lemma 4.5. *Let $K \subseteq L \subseteq M$ be inclusions of fields, and suppose that $[L : K], [M : L] < \infty$. Then $[M : K] = [M : L][L : K]$. More precisely if $\alpha_1, \ldots, \alpha_n$ is a basis for L as a K-vector space and $\beta_1, \ldots, \beta_m$ is a basis for M as an L-vector space, then $\alpha_i \beta_j$ for $1 \leq i \leq n$, $1 \leq j \leq m$ form a basis for M as a K-vector space.*

Now we give a generalization of proposition 4.3. The proof is essentially the same, you just have to mention τ_0 in the right places. (I didn't give it in class, but it is here in the notes for completeness!)

Lemma 4.6. *Let K be a number field, $\tau_0 : K \to \mathbf{C}$ an embedding, and let α be an algebraic number. Then there are $[K(\alpha) : K]$ embeddings $\tau : K(\alpha) \to \mathbf{C}$ which restrict to τ_0 on K.*

Proof. We write $m_{\alpha,K} \in K[x]$ for the minimal monic polynomial of α over K, which is an irreducible polynomial of degree $[K(\alpha) : K]$ and hence has $[K(\alpha) : K]$ distinct roots in $\mathbf{C}$ by Proposition 3.10. Hence the polynomial $\tau_0(m_{\alpha,K}) \in \mathbf{C}[x]$ also has $[K(\alpha) : K]$ distinct roots in $\mathbf{C}$.

Then we have an isomorphism

$$\begin{aligned} K[x]/(m_{\alpha,K}) &\to K(\alpha) \\ f + (m_{\alpha,K}) &\mapsto f(\alpha) \end{aligned}$$

Giving a homomorphism $\tau : K(\alpha) \to \mathbf{C}$ which restricts to τ_0 on K is the same as giving a homomorphism $\phi : K[x] \to \mathbf{C}$ which restricts to τ_0 on K and for which $(m_{\alpha,K}) \subset \ker(\phi)$, where the relation between τ and ϕ is given by $\tau(f(\alpha)) = \phi(f)$ for $f \in K[x]$. Giving a homomorphism $\phi : K[x] \to \mathbf{C}$ which restricts to τ_0 on K is the same as giving the element $\beta = \phi(x) \in \mathbf{C}$ as then $\phi(f) = \tau_0(f)(\beta)$. Finally we have $(m_{\alpha,K}) \subseteq \ker(\phi)$ if and only $0 = \phi(m_{\alpha,K}) = \tau_0(m_{\alpha,K})(\beta) = 0$.

To sum up, giving an embedding $\tau : K(\alpha) \to \mathbf{C}$ restricting to τ_0 on K is the same as giving a root $\beta \in \mathbf{C}$ of $\tau_0(m_{\alpha,K}) \in \mathbf{C}[x]$, and there are $[K(\alpha) : K]$ such roots. □

Corollary 4.7. *Let $K \subseteq L$ be number fields and let $\tau_0 : K \to \mathbf{C}$ be an embedding. Then there are $[L : K]$ embeddings $\tau : L \to \mathbf{C}$ which restrict to τ_0 on K.*

[17] In some texts, these formulas are taken as the definition of the trace and the norm. The disadvantage of using this as a defintion is that you have to prove something about the set of embeddings Σ_K before you can really do anything, and it is not obvious that the trace and norm are actually in $\mathbf{Q}$.

[18] So if we used the primitive element theorem, we would have proved 1 in general, but we won't!

Proof. We prove this by induction on the degree $[L : K]$ of the extension. If $[L : K] = 1$ then $L = K$ and the result is trivial. Otherwise, pick $\alpha \in L$, $\alpha \notin K$ and consider the tower of fields $K \subseteq K(\alpha) \subseteq L$. Then $[L : K(\alpha)] < [L : K]$ and so by induction we may use the conclusion of the corollary for the extension $K(\alpha) \subseteq L$. Now by Lemma 4.6, τ_0 extends to $[K(\alpha) : K]$ embeddings of $K(\alpha)$, and by the induction hypothesis they each extend to $[L : K(\alpha)]$ embeddings of L, giving $[L : K(\alpha)][K(\alpha) : K] = [L : K]$ embeddings in total of L extending τ_0. □

As a special case of the corollary, we can take $K = \mathbf{Q}$ and we have proven part 1 of the theorem.

Proof of Theorem 4.4 2. As explained before, we have already proven these formulas in the special case of $\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha)$ and $N_{\mathbf{Q}(\alpha)}(\alpha)$ by combining Propositions 3.9 and 4.3. For the general case, first we will prove the formulas

$$\mathrm{tr}_K(\alpha) = [K : \mathbf{Q}(\alpha)] \, \mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha), \quad N_K(\alpha) = N_{\mathbf{Q}(\alpha)}(\alpha)^{[K:\mathbf{Q}(\alpha)]}.$$

by generalizing the proof of proposition 3.9. We start by finding a basis for K as a $\mathbf{Q}$-vector space.

Let $n = \deg(\alpha)$, $m = [K : \mathbf{Q}(\alpha)]$. Then $1, \alpha, \ldots, \alpha^{n-1}$ is a basis for $\mathbf{Q}(\alpha)$ as a $\mathbf{Q}$ vector space, and we can pick any basis $\beta_1, \ldots, \beta_m$ for K as a $\mathbf{Q}(\alpha)$-vector space. Then by Lemma 4.5

$$\beta_1, \alpha\beta_1, \alpha^2\beta_1, \ldots, \alpha^{n-1}\beta_1, \beta_2, \alpha\beta_2, \ldots, \beta_m, \alpha\beta_m, \ldots, \alpha^{n-1}\beta_m$$

is a basis for K as a $\mathbf{Q}$-vector space. We note that for $i = 1, \ldots, m$ the subspace spanned by $\beta_i, \alpha\beta_i, \ldots, \alpha^{n-1}\beta_i$ is invariant by $\alpha\cdot$, and moreover the matrix of $\alpha\cdot$ on these vectors is the companion matrix $C(m_\alpha)$ exactly as in the proof of Proposition 3.9. It follows that the matrix of $\alpha\cdot$ on K is a block diagonal matrix, with $m = [K : \mathbf{Q}(\alpha)]$ blocks which are $C(m_\alpha)$. Upon taking traces and determinants we obtain the formulas.

Now to conclude the proof we calculate

$$\begin{aligned}
\sum_{\tau \in \Sigma_K} \tau(\alpha) &= \sum_{\tau_0 \in \Sigma_{\mathbf{Q}(\alpha)}} \sum_{\tau \in \Sigma_K, \tau|_{\mathbf{Q}(\alpha)} = \tau_0} \tau(\alpha) \\
&= [K : \mathbf{Q}(\alpha)] \sum_{\tau_0 \in \Sigma_{\mathbf{Q}(\alpha)}} \tau_0(\alpha) \\
&= [K : \mathbf{Q}(\alpha)] \, \mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) \\
&= \mathrm{tr}_K(\alpha).
\end{aligned}$$

Here in the first equality, we break up Σ_K according to their restrictions τ_0 to $\mathbf{Q}(\alpha)$. As $\alpha \in \mathbf{Q}(\alpha)$, $\tau(\alpha) = \tau_0(\alpha)$ only depends on this restriction. By Corollary 4.7 the inner sum is over $[K : \mathbf{Q}(\alpha)]$ elements, giving the second equality. The third equality is the special case we had already proved before, and the fourth equality is what we just proved.

The proof for norms is exactly the same calculation, replacing sums by products. □

By now we should really prove the following important result:

Proposition 4.8. *Let K be a number field and let $\alpha \in \mathcal{O}_K$. Then $\mathrm{tr}_K(\alpha), N_K(\alpha) \in \mathbf{Z}$.*

We observed that this was true in the special case that $K = \mathbf{Q}(\alpha)$ in Proposition 3.9.

Proof. This follows immediately from the special case of 3.9 and the formulas

$$\mathrm{tr}_K(\alpha) = [K : \mathbf{Q}(\alpha)]\,\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha), \quad N_K(\alpha) = N_{\mathbf{Q}(\alpha)}(\alpha)^{[K:\mathbf{Q}(\alpha)]}.$$

obtained in the proof of Theorem 4.4.[19] □

We have now spent quite a bit of time talking about traces and norms! We will make a crucial use of the trace in the next lecture, while you have probably already seen applications of the norm for rings like $\mathbf{Z}[i]$ for studying units and factorizations. We will see many such uses soon. For example, an important application of the norm is that we can use it to characterize units in integer rings:

Proposition 4.9. *Let K be a number field and let $\alpha \in \mathcal{O}_K$. Then $\alpha \in \mathcal{O}_K^\times$ if and only if $N_K(\alpha) \in \mathbf{Z}^\times = \{1, -1\}$.*[20]

Proof. If there exists $\beta \in \mathcal{O}_K^\times$ with $\alpha\beta = 1$ then $1 = N_K(1) = N_K(\alpha\beta) = N_K(\alpha)N_K(\beta)$. Hence $N_K(\alpha) \in \mathbf{Z}^\times = \{1, -1\}$.

Conversely suppose $N_K(\alpha) = \pm 1$. Then by Theorem 4.4,

$$\pm 1 = N_K(\alpha) = \alpha \prod_{\tau \in \Sigma_K, \tau \neq \tau_{id}} \tau(\alpha)$$

where $\tau_{id}(x) = x$ for $x \in K$ is the "identity" embedding. Let

$$\beta = \pm \prod_{\tau \in \Sigma_K, \tau \neq \tau_{id}} \tau(\alpha)$$

so that $\alpha\beta = 1$. Then $\beta = \alpha^{-1} \in K$. On the other hand each $\tau(\alpha)$ is an algebraic integer, so β is a product of algebraic integers and hence itself an algebraic integer. Hence $\beta \in \mathcal{O}_K$. □

5 Lecture 5: The additive structure of integer rings

Our goal for the day is to prove the following important theorem:

Theorem 5.1. *Let K be a number field. Then as a group under addition, $\mathcal{O}_K$ is a free abelian group of rank $[K : \mathbf{Q}]$.*

We recall that an abelian group M is said to be free of rank n if either of the following two equivalent conditions holds:

1. M is isomorphic to $\mathbf{Z}^n$
2. M has a basis of size n: there exists $m_1, \ldots m_n$ such that any $m \in M$ has a unique[21] expression of the form $m = \sum c_i m_i$ with $c_i \in \mathbf{Z}$.

[19] I should have explained these formulas and then proved this proposition immediately after proposition 3.9 in Lecture 3. Sorry!

[20] The norm of a unit can really be -1. For example if $[K : \mathbf{Q}]$ is odd, then $N_K(-1) = -1$.

[21] We previously discussed finitely generated abelian groups. There we don't require uniqueness.

The proof that these are equivalent is familiar from linear algebra: given a basis $m_1, \ldots, m_n$ of M you can write down an isomorphism

$$\mathbf{Z}^n \to M$$
$$(c_1, \ldots, c_n) \to \sum c_i m_i.$$

Conversely $\mathbf{Z}^n$ has a basis of size n given by the standard basis vectors, and hence so does any isomorphic abelian group. We also note that the rank n of a free abelian group is well defined. For instance $|\mathbf{Z}^n/2\mathbf{Z}^n| = 2^n$.

Definition 5.2. We call a basis $\alpha_1, \ldots, \alpha_n$ for $\mathcal{O}_K$ an integral basis.

Example 5.3. We have already proved the theorem for the quadratic integer rings $\mathcal{O}_d$ in lecture 2. We found that $1, \sqrt{d}$ is an integral basis when $d \not\equiv 1 \pmod 4$ and $1, \frac{1+\sqrt{d}}{2}$ is an integral basis when $d \equiv 1 \pmod 4$.

Today we will use the following fact about about free abelian groups which is a consequence of the structure theory of finitely generated abelian groups from algebra:

Proposition 5.4. *Any subgroup of a free abelian group of rank n is a free abelian group of rank $m \leq n$.*[22]

Proof. We prove this by induction on n. Suppose we have a subgroup $M \subseteq \mathbf{Z}^n$. We consider the projection onto the first coordinate

$$p : \mathbf{Z}^n \to \mathbf{Z}$$
$$(c_1, \ldots, c_n) \mapsto c_1$$

We have $\ker(p) \simeq \mathbf{Z}^{n-1}$.

We consider $p(M) \subseteq \mathbf{Z}$. We know from algebra that the subgroups of $\mathbf{Z}$ are $\{0\}$ and $r\mathbf{Z}$ for $r \geq 1$. If $p(M) = \{0\}$ then $M \subseteq \ker(p)$ which is a free abelian group of rank $n-1$, so we are done by induction.

Otherwise suppose $p(M) = r\mathbf{Z}$ for some $r > 0$. Pick an element $m \in M$ with $p(m) = r$. We will show that

$$M = (\mathbf{Z} \cdot m) \times (M \cap \ker(p))$$

i.e. M is the direct product of its subgroups $\mathbf{Z} \cdot m$ and $M \cap \ker(p)$. This is enough to conclude because by induction $M \cap \ker(p)$ is free of rank $\leq n-1$.

Now recall that to check that an abelian group is the direct product of two of its subgroups we need to check:

- The intersection of the two subgroups is $\{0\}$. To see this, note that for $k \in \mathbf{Z}$, $p(km) = kp(m) = kr$, so $km \in \ker(p)$ only for $k = 0$.
- The two subgroups span M. Indeed given any $v \in M$, $p(v) = kr$ for some $k \in \mathbf{Z}$. Then $p(v - km) = 0$, so $v - km \in \ker(p)$, and we can write $v = km + (v - km)$.

[22] One formulation of the main result of the structure theory of finitely generated abelian groups is: if $M \subseteq \mathbf{Z}^n$ is a subgroup, then there is a basis $v_1, \ldots, v_n$ of $\mathbf{Z}^n$, an integer $m \leq n$, and integers $d_1 | d_2 | \ldots | d_m$ such that $d_1 v_1, d_2 v_2, \ldots, d_m v_m$ is a basis for M. In particular M is free of rank m.

□

Our strategy for proving Theorem 5.1 will be as follows. Let $\alpha_1, \ldots, \alpha_n \in \mathcal{O}_K$ be any $n = [K : \mathbf{Q}]$ elements which form a basis for K as a $\mathbf{Q}$ vector space. Recall that if α is an algebraic number then there exists $n > 0$ such that $n\alpha$ is an algebraic integer, so we may begin with any $\mathbf{Q}$ basis $\alpha_1, \ldots, \alpha_n$ for K and replace α_i with $n_i\alpha_i \in \mathcal{O}_K$.

Then to prove 5.1 it suffices to show that there is some integer $d > 0$ such that

$$\mathcal{O}_K \subseteq \mathbf{Z} \cdot \frac{\alpha_1}{d} \oplus \mathbf{Z} \cdot \frac{\alpha_2}{d} \oplus \cdots \oplus \mathbf{Z} \cdot \frac{\alpha_n}{d} \simeq \mathbf{Z}^n$$

Indeed this implies that $\mathcal{O}_K$ is free of rank at most n, but we also have the inclusion

$$\mathbf{Z}^n \simeq \mathbf{Z} \cdot \alpha_1 \oplus \mathbf{Z} \cdot \alpha_2 \oplus \cdots \oplus \mathbf{Z} \cdot \alpha_n \subseteq \mathcal{O}_K$$

and hence the rank of $\mathcal{O}_K$ must be at least n, and hence must be exactly n.

So let $\alpha \in \mathcal{O}_K$ and write $\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n$ for $c_i \in \mathbf{Q}$. Our goal is to bound the denominators of the c_i.

In the quadratic case $K = \mathbf{Q}(\sqrt{d})$ we considered $\alpha = c_1 \cdot 1 + c_2\sqrt{d}$ and considered $N_K(\alpha) = c_1^2 - c_2^2 d$. You might imagine we should do the same thing in general. However in general $N_K(c_1\alpha_1 + \cdots + c_n\alpha_n)$ will be a polynomial of degree n in the n variables $c_1, \ldots, c_n$! Clearly it will be very difficult to analyze under what conditions on the $c_i \in \mathbf{Q}$, $N_K(c_1\alpha_1 + \cdots + c_n\alpha_n) \in \mathbf{Z}$ beyond the quadratic case. Then we would still have to consider the other coefficients of the minimal polynomial of α. So this is probably not such a good way to try to proceed.

Instead we observe that $\alpha \in \mathcal{O}_K$ implies that $\alpha\alpha_1, \ldots, \alpha\alpha_n \in \mathcal{O}_K$ which implies that $\mathrm{tr}_K(\alpha\alpha_1), \mathrm{tr}_K(\alpha\alpha_2), \ldots, \mathrm{tr}(\alpha\alpha_n) \in \mathbf{Z}$. None of these implications are necessarily reversible but we will see that the integrality of these traces is enough to bound the denominators of the coefficients c_i, and rather advantageously, $\mathrm{tr}_K(\alpha\alpha_i)$ is a linear expression in the c_i. More explicitly:

$$\mathrm{tr}_K(\alpha\alpha_i) = \mathrm{tr}_K(c_1\alpha_1\alpha_i + \cdots + c_n\alpha_n\alpha_i) = c_1 \,\mathrm{tr}_K(\alpha_1\alpha_i) + \cdots + c_n \,\mathrm{tr}_K(\alpha_n\alpha_i).$$

Thus if we let $m_i = \mathrm{tr}(\alpha\alpha_i) \in \mathbf{Z}$, we in fact have a system of n linear equations in n unknowns:

$$\begin{aligned}
c_1 \,\mathrm{tr}_K(\alpha_1\alpha_1) + c_2 \,\mathrm{tr}_K(\alpha_2\alpha_1) + \cdots + c_n \,\mathrm{tr}_K(\alpha_n\alpha_1) &= m_1 \\
c_1 \,\mathrm{tr}_K(\alpha_1\alpha_2) + c_2 \,\mathrm{tr}_K(\alpha_2\alpha_2) + \cdots + c_n \,\mathrm{tr}_K(\alpha_n\alpha_2) &= m_2 \\
&\vdots \\
c_1 \,\mathrm{tr}_K(\alpha_1\alpha_n) + c_2 \,\mathrm{tr}_K(\alpha_2\alpha_n) + \cdots + c_n \,\mathrm{tr}_K(\alpha_n\alpha_n) &= m_n
\end{aligned}$$

or in matrix form

$$A \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix}.$$

where $A = (\mathrm{tr}_K(\alpha_i\alpha_j))_{1 \le i,j \le n} \in M_n(\mathbf{Z})$.

Our first order of business is verifying that this system of equations actually has a unique solution:

Lemma 5.5. *Let $\alpha_1, \ldots, \alpha_n \in K$ be any basis for K as a $\mathbf{Q}$-vector space. Form the matrix $A = (\mathrm{tr}_K(\alpha_i \alpha_j))_{1 \leq i,j \leq n}$. Then $\det(A) \neq 0$.*

Proof. We will show that the linear transformation $\mathbf{Q}^n \to \mathbf{Q}^n$ defined by A is injective. So suppose $v = (c_1, \ldots, c_n)^T \in \ker(A)$. Reversing the calculation above, saying that $Av = 0$ is exactly the same as saying that if $\alpha = c_1 v_1 + \cdots + c_n v_n$ then $\mathrm{tr}(\alpha \alpha_i) = 0$ for $i = 1, \ldots, n$. But now any $\beta \in K$ may be written as $\beta = d_1 \alpha_1 + \cdots d_n \alpha_n$ and we compute

$$\mathrm{tr}_K(\alpha\beta) = \sum d_i \, \mathrm{tr}_K(\alpha \alpha_i) = 0.$$

But now if $\alpha \neq 0$, we can take $\beta = \alpha^{-1}$ and we have

$$0 = \mathrm{tr}_K(\alpha\beta) = \mathrm{tr}_K(1) = n.$$

Thus $\alpha = 0$ and hence $v = (c_1, \ldots, c_n)^T = 0$ and hence $\det(A) \neq 0$. □

Returning to our system of equations, we can now solve it for the c_i:

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = A^{-1} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix}$$

We have $A^{-1} \in M_n(\mathbf{Q})$.[23] We can pick an integer $d > 0$ such that $dA^{-1} \in M_n(\mathbf{Z})$ (e.g. take d to be the least common multiple of the denominators of the n^2 entries of A.) Then we have

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = d^{-1}(dA^{-1}) \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix} \in d^{-1}\mathbf{Z}^n.$$

So we have proved Theorem 5.1!

It is natural to wonder what we can actually take d to be in the above proof. The following lemma gives a bound:

Lemma 5.6. *Let $A \in M_n(\mathbf{Z})$ with $\det(A) \neq 0$. Then $\det(A)A^{-1} \in M_n(\mathbf{Z})$.*

In class I mentioned that one way to prove this is using the adjugate matrix from linear algebra.

Definition 5.7. Let R be an arbitrary commutative ring and let $A \in M_n(R)$.

1. For $1 \leq i, j \leq n$, the (i,j) *minor*[24] $A_{ij} \in R$ of A is defined to be the determinant of the $n-1$ by $n-1$ matrix obtained by deleting the ith row and jth column of A.

2. The *adjugate* of A is $\mathrm{adj}(A) = ((-1)^{i+j} A_{ji})$, that is the matrix whose entry in the ith row and jth column is $(-1)^{i+j} A_{ji}$.

[23] Even though $A \in M_n(\mathbf{Z})$ it is certainly not necessarily the case that $A^{-1} \in M_n(\mathbf{Z})$. We will return to this point shortly.

[24] In class I mistakenly called it the cofactor but then wikipedia corrected me: the cofactor is $(-1)^{i+j} A_{ij}$. The formula I wrote for the adjugate was still correct. Mea culpa!

The point of the adjugate is the following formula:

Proposition 5.8. *For $A \in M_n(R)$,* $\mathrm{adj}(A) \cdot A = A \cdot \mathrm{adj}(A) = \det(A) I$.

This formula is proved by interpreting the entries of the product as Laplace expansions of a determinant. This formula immediately implies lemma 5.6 as well as the following important corollary

Corollary 5.9. *For $A \in M_n(R)$, we have $A \in M_n(R)^\times$ if and only if* $\det(A) \in R^\times$

Remark 5.10. *I didn't say this during the lecture, but actually we can improve lemma 5.6 slightly: if d' denotes the greatest common divisor of the entries of $\mathrm{adj}(A)$ then $d'|d$ and $(d/d')A^{-1} \in M_n(\mathbf{Z})$, and moreover this is the best possible as this matrix will have coprime entries.

Lemma 5.6 and this improvement can also be understood using the structure theory of finitely generated abelian groups. Applying the statement recalled in the footnote of proposition 5.4 we deduce that there are integers $d_1|d_2|\cdots|d_n$ and matrices $S_1, S_2 \in M_n(\mathbf{Z})^\times$ so that

$$A = S_1 \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} S_2$$

This is called the smith normal form of A. Computing A^{-1} using this formula we see that $d_n A^{-1} \in M_n(\mathbf{Z})$. In fact $d_1 \cdots d_n = \det(A)$ while $d_1 \cdots d_{n-1}$ is the greatest common divisor of the entries of $\mathrm{adj}(A)$.

6 Lecture 6: Discriminants

Motivated by the last lecture we make the following definition.

Definition 6.1. Let K be a number field of degree $n = [K : \mathbf{Q}]$. We define the *discriminant* of $\alpha_1, \ldots, \alpha_n \in K$ as

$$\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = \det(\mathrm{tr}_K(\alpha_i \alpha_j)_{1 \le i,j \le n}) \in \mathbf{Q}.$$

The matrix of traces we take the determinant of arose in the last lecture when we studied the system of equations $\mathrm{tr}(\alpha\alpha_i) = m_i$. You might be familiar with the discriminant of a polynomial. We will soon see that there is a relation.

Example 6.2. Let $K = \mathbf{Q}(\sqrt{d})$. We compute

$$\mathrm{disc}(1, \sqrt{d}) = \det \begin{pmatrix} \mathrm{tr}_K(1) & \mathrm{tr}_K(\sqrt{d}) \\ \mathrm{tr}_K(\sqrt{d}) & \mathrm{tr}_K(d) \end{pmatrix} = \det \begin{pmatrix} 2 & 0 \\ 0 & 2d \end{pmatrix} = 4d,$$

$$\mathrm{disc}\left(1, \frac{1+\sqrt{d}}{2}\right) = \det \begin{pmatrix} 2 & 1 \\ 1 & \frac{d+1}{2} \end{pmatrix} = d.$$

Here are two basic facts about the discriminant:

1. If $\alpha_1, \ldots, \alpha_n$ is a basis for K as a $\mathbf{Q}$-vector space, then $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) \neq 0$. This was Lemma 5.5 of the last lecture.

2. If $\alpha_1, \ldots, \alpha_n \in \mathcal{O}_K$ then $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}$.

Exercise 6.3. Prove the converse to Lemma 5.5: if $\alpha_1, \ldots, \alpha_n$ are linearly dependent then $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = 0$.

We first study how the discriminant changes when we change the basis:

Lemma 6.4. *Suppose that $\alpha_1, \ldots, \alpha_n \in K$ and $\beta_1, \ldots, \beta_n \in K$ are related by*

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = S \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

for some matrix $S = (c_{ij}) \in M_n(\mathbf{Q})$. Then writing $A = (\mathrm{tr}_K(\alpha_i\alpha_j))$ and $B = (\mathrm{tr}_K(\beta_i\beta_j))$ we have

$$B = SAS^t,$$

and hence

$$\mathrm{disc}(\beta_1, \ldots, \beta_n) = \det(S)^2 \mathrm{disc}(\alpha_1, \ldots, \alpha_n).$$

Proof. This is an exercise in matrix multiplication and indices. We have:

$$\begin{aligned} \mathrm{tr}_K(\beta_i\beta_j) &= \sum_{k,l=1}^{n} \mathrm{tr}_K(c_{i,k}\alpha_k c_{j,l}\alpha_l) \\ &= \sum_{k,l=1}^{n} c_{i,k}\,\mathrm{tr}_K(\alpha_k\alpha_l)c_{j,l} \end{aligned}$$

and this is the i,jth entry of SAS^t.[25] □

Corollary 6.5. *Let $M \subset K$ be a free abelian subgroup of rank n. Then if $\alpha_1, \ldots, \alpha_n \in M$ and $\beta_1, \ldots, \beta_n \in M$ are two $\mathbf{Z}$-bases, then* $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = \mathrm{disc}(\beta_1, \ldots, \beta_n)$.

Proof. In this case the change of basis matrix S is in $M_n(\mathbf{Z})^\times$, and so $\det(S)^2 = 1$. □

Definition 6.6.

1. If $M \subset K$ is a free abelian subgroup of rank n we define $\mathrm{disc}(M) = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)$ for any $\mathbf{Z}$-basis of M. In view of the corollary this is independent of the choice of basis.

2. We define the *discriminant* of K to be $D_K = \mathrm{disc}(\mathcal{O}_K) \in \mathbf{Z}$. In other words, $D_K = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)$ for any integral basis $\alpha_1, \ldots, \alpha_n$.

Example 6.7. If $d \neq 0, 1$ is a squarefree integer, then by 6.2

$$D_{\mathbf{Q}(\sqrt{d})} = \begin{cases} 4d & d \not\equiv 1 \pmod 4 \\ d & d \equiv 1 \pmod 4 \end{cases}$$

The discriminant of a number field is a fundamental invariant that shows up in many different contexts throughout algebraic number theory. For now we see that it is closely connected with computing the integer ring $\mathcal{O}_K$.

We can rephrase the change of basis formula for discriminants in the following way:

[25] This is really just the "change of basis formula" for bilinear forms.

Proposition 6.8. *Suppose $M \subseteq N \subseteq K$ are free abelian subgroups of rank n inside K. Then*

$$\mathrm{disc}(M) = [N : M]^2 \cdot \mathrm{disc}(N).$$

Recall that the index $[N : M]$ is defined to be the cardinality of the quotient N/M.

Proof. By the structure theory of finitely generated abelian groups, there is a basis $\alpha_1, \ldots, \alpha_n$ of N and integers $d_1, \ldots, d_n$ such that $d_1\alpha_1, \ldots, d_n\alpha_n$ is a basis for M. Then by Lemma 6.4 we have

$$\mathrm{disc}(M) = \mathrm{disc}(d_1\alpha_1, \ldots, d_n\alpha_n) = (d_1 \cdots d_n)^2 \mathrm{disc}(\alpha_1, \ldots, \alpha_n) = (d_1 \cdots d_n)^2 \mathrm{disc}(N)$$

On the other hand we have

$$N/M \simeq \frac{\mathbf{Z} \oplus \mathbf{Z} \oplus \cdots \oplus \mathbf{Z}}{d_1\mathbf{Z} \oplus d_2\mathbf{Z} \oplus \cdots \oplus d_n\mathbf{Z}} \simeq (\mathbf{Z}/d_1\mathbf{Z}) \oplus \cdots \oplus (\mathbf{Z}/d_n\mathbf{Z})$$

and so $[N : M] = d_1 \cdots d_n$. □

Typically we apply proposition 6.8 with $N = \mathcal{O}_K$ and with $M \subseteq \mathcal{O}_K$ the subring of algebraic integers we have found so far (and may be hoping is all of $\mathcal{O}_K$!) We restate it in this case, as well as some important consequences.

Proposition 6.9. *Suppose $M \subseteq \mathcal{O}_K$ is a free abelian subgroup of rank n. Then we have*

$$\mathrm{disc}(M) = [\mathcal{O}_K : M]^2 \cdot D_K.$$

In particular:

1. *$D_K \mid \mathrm{disc}(M)$*
2. *If $p \mid [\mathcal{O}_K : M]$ for some prime p then $p^2 | \mathrm{disc}(M)$.*
3. *If $\mathrm{disc}(M)$ is squarefree then $\mathcal{O}_K = M$ and $D_K = \mathrm{disc}(M)$.*

Example 6.10. We revisit the computation of integer rings of quadratic fields using what we have learned. So let $K = \mathbf{Q}(\sqrt{d})$ with $d \neq 1$ a squarefree integer.

We start with the "obvious" subring $\mathbf{Z}[\sqrt{d}] \subset \mathcal{O}_K$. We compute $\mathrm{disc}(\mathbf{Z}[\sqrt{d}]) = 4d$. The only prime p for which $p^2 | \mathrm{disc}(\mathbf{Z}[\sqrt{d}])$ is $p = 2$. Then proposition 6.9 tells us there are two possibilities:

1. $[\mathcal{O}_K : \mathbf{Z}[\sqrt{d}]] = 1$, i.e. $\mathcal{O}_K = \mathbf{Z}[\sqrt{d}]$ and $D_K = 4d$.
2. $[\mathcal{O}_K : \mathbf{Z}[\sqrt{d}]] = 2$ and $D_K = d$.

If we were in the second case, we would have $\mathcal{O}_K \subseteq \frac{1}{2}\mathbf{Z}[\sqrt{d}]$[26], and so one of $\frac{1}{2}, \frac{\sqrt{d}}{2}, \frac{1+\sqrt{d}}{2}$ would be an algebraic integer. $\frac{1}{2}$ and $\frac{\sqrt{d}}{2}$ clearly aren't, so we just compute the minimal polynomial of $\frac{1+\sqrt{d}}{2}$ and see that it is $x^2 - x + \frac{1-d}{4}$. So we see that the second case happens exactly when $d \equiv 1 \pmod 4$.

[26] I got asked about this in class and didn't give a very good answer: if $M \subseteq N$ are abelian groups and $[N : M]$ is finite, then $[N : M](N/M) = 0$ by Lagrange's theorem, or in other words $[N : M]N \subseteq M$. In particular applied to the case at hand, if $[\mathcal{O}_K : \mathbf{Z}[\sqrt{d}]] = 2$, then $2\mathcal{O}_K \subseteq \mathbf{Z}[\sqrt{d}]$, or $\mathcal{O}_K \subseteq \frac{1}{2}\mathbf{Z}[\sqrt{d}]$. Alternatively as I suggested in class you can come to the same conclusion using the kind of calculation of the proof of Proposition 6.8: there is a basis m_1, m_2 of $\mathcal{O}_K$ such that d_1m_1, d_2m_2 is a basis of $\mathbf{Z}[\sqrt{d}]$ and $2 = [\mathcal{O}_K : \mathbf{Z}[\sqrt{d}]] = d_1d_2$. So the only possibility is $d_1 = 1$, $d_2 = 2$. So $2\mathcal{O}_K = 2\mathbf{Z}m_1 + 2\mathbf{Z}m_2 \subseteq \mathbf{Z}m_1 + 2\mathbf{Z}m_2 = \mathbf{Z}[\sqrt{d}]$.

6.1 Other formulas for the discriminant

Proposition 6.11. *Let K be a number field and let $\tau_1, \ldots, \tau_n \in \Sigma_K$ be a enumeration of the embeddings of K. Let $\alpha_1, \ldots, \alpha_n \in K$. Then we have*[27]

$$\operatorname{disc}(\alpha_1, \ldots, \alpha_n) = \det(\tau_i(\alpha_j))^2$$

Proof. We write $A = (\operatorname{tr}_K(\alpha_i\alpha_j))$ and $T = (\tau_i(\alpha))$. We compute using Theorem 4.4:

$$\operatorname{tr}_K(\alpha_i\alpha_j) = \sum_k \tau_k(\alpha_i\alpha_j) = \sum_k \tau_k(\alpha_i)\tau_k(\alpha_j)$$

and this is the i,jth entry of T^tT. Thus $A = T^tT$ and the result follows by taking determinants. □

Example 6.12. Using this formula we compute

$$\operatorname{disc}(1, \sqrt{d}) = \det \begin{pmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{pmatrix}^2 = (-2\sqrt{d})^2 = 4d$$

You might ask, what about $\det(\tau_i(\alpha_j))$ without the square? First of all this is not really well defined because we ordered the embeddings arbitrarily, and permuting them could flip the sign. Also it is not necessarily rational or even real. Nonetheless, when $\alpha_1, \ldots, \alpha_n$ is an integral basis, the absolute value

$$|\det(\tau_i(\alpha_j))| = \sqrt{|D_K|}$$

has an important geometric interpretation that we will explain soon. This will explain why $\sqrt{|D_K|}$ appears in so many formulas in algebraic number theory, some of which we will see later in this course.

Now suppose that $K = \mathbf{Q}(\alpha)$ where α is an algebraic number with minimal polynomial

$$m_\alpha = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)$$

where $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$ are the conjugates of α. Recall then that by Proposition 4.3 we can consider an enumeration $\tau_1, \ldots, \tau_n$ of $\Sigma_{\mathbf{Q}(\alpha)}$ where $\tau_i(\alpha) = \alpha_i$.

A natural basis of $\mathbf{Q}(\alpha)$ to consider is $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ (sometimes called a power basis). We now consider its discriminant. We find:

$$\begin{aligned} \operatorname{disc}(1, \alpha, \ldots, \alpha^{n-1}) &= \det \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{pmatrix}^2 \\ &= \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2 \end{aligned}$$

Here the first equality is 6.11 and the second is the determinant of the Vandermonde matrix:

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots & \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \le i < j \le n} (x_i - x_j)$$

[27] In some texts this formula is taken as the definition of the discriminant. With this definition, it is a bit less clear that the discriminant is rational.

Exercise 6.13. Prove this formula if you don't already know it.

The quantity

$$\prod_{1\leq i<j\leq n}(\alpha_i-\alpha_j)^2$$

is called the discriminant of the polynomial m_α. From this formula it is not obvious that it is in $\mathbf{Q}$. We give one more formula which is in practice often the fastest way to compute discriminants.

Proposition 6.14. *We have*

$$\operatorname{disc}(1,\alpha,\dots,\alpha^{n-1})=(-1)^{\frac{n(n-1)}{2}}N_{\mathbf{Q}(\alpha)}(m'_\alpha(\alpha)).$$

Exercise 6.15. Check that $(-1)^{\frac{n(n-1)}{2}}$ is 1 if $n\equiv 0,1 \pmod 4$ and -1 if $n\equiv 2,3 \pmod 4$.

Proof. By Theorem 4.4 we have

$$N_{\mathbf{Q}(\alpha)}(m'_\alpha(\alpha))=\prod_{i=1}^n\tau_i(m'_\alpha(\alpha))=\prod_{i=1}^n m'_\alpha(\tau_i(\alpha))=\prod_{i=1}^n m'_\alpha(\alpha_i).$$

Now using the formula $m_\alpha=(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$ we compute using the Leibniz rule that for $i=1,\dots,n$

$$m'_\alpha=(x-\alpha_1)\cdots(x-\alpha_{i-1})(x-\alpha_{i+1})\cdots(x-\alpha_n)+(x-\alpha_i)(\cdots)$$

and so

$$m'_\alpha(\alpha_i)=\prod_{1\leq j\leq n, j\neq i}(\alpha_i-\alpha_j)$$

Thus

$$N_{\mathbf{Q}(\alpha)}(m'_\alpha(\alpha))=\prod_{1\leq i,j\leq n, i\neq j}(\alpha_i-\alpha_j)=(-1)^{\frac{n(n-1)}{2}}\prod_{1\leq i<j\leq n}(\alpha_i-\alpha_j)^2$$

□

Example 6.16. The polynomial $f=x^3+x+1$ is irreducible by the rational root test. Let $\alpha\in\mathbf{C}$ be a root. We compute

$$\begin{aligned}\operatorname{disc}(\mathbf{Z}[\alpha])&=-N_{\mathbf{Q}(\alpha)}(f'(\alpha))\\&=-N_{\mathbf{Q}(\alpha)}(3\alpha^2+1)\\&=-N_{\mathbf{Q}(\alpha)}(\alpha^{-1}(3\alpha^3+\alpha))\\&=N_{\mathbf{Q}(\alpha)}(-2\alpha-3)\\&=2^3f(-3/2)\\&=-31\end{aligned}$$

(For the formula $N_{\mathbf{Q}(\alpha)}(-2\alpha-3)=2^3f(-3/2)$, see problem 2e on the first example sheet, or compute that $N_{\mathbf{Q}(\alpha)}(-2\alpha-3)=-31$ however you like!) If you knew the formula for the discriminant of a cubic $f=x^3+ax+b$ is $-27b^2-4a^3$ you could have used that.

Now because 31 is prime, we can conclude by proposition 6.9 that $\mathcal{O}_{\mathbf{Q}(\alpha)}=\mathbf{Z}[\alpha]$ and $D_{\mathbf{Q}(\alpha)}=-31$.

Is $\mathbf{Z}[\alpha]$ a UFD? Stay tuned!

7 Lecture 7: Example class 1

1. Let $d \in \mathbf{Z}$, $d \neq 0, 1$ be squarefree, and consider the ring of integers $\mathcal{O}_d$ of $\mathbf{Q}(\sqrt{d})$.

   (a) Compute $\mathcal{O}_d$ again for yourself!

   (b) For all $d < 0$, compute $\mathcal{O}_d^\times$. (Hint: norms)

   (c) Show that $1 + \sqrt{2} \in \mathcal{O}_2^\times$. Prove that $\mathcal{O}_2^\times$ has infinite order. Do the same for $\mathcal{O}_3$ and $\mathcal{O}_5$. (Later in the course we will prove that for all $d > 0$, $\mathcal{O}_d^\times$ has infinite order. If you took elementary number theory last term, you already know this!)

2. Let K be a number field.

   (a) For $\alpha \in K$ and $c \in \mathbf{Q}$ prove the following formulas:

   $$\mathrm{tr}_K(c) = c[K : \mathbf{Q}], \quad N_K(c) = c^{[K:\mathbf{Q}]}, \quad \mathrm{tr}(c\alpha) = c\,\mathrm{tr}_K(\alpha)$$

   Deduce from the last formula that $\mathrm{tr}_K : K \to \mathbf{Q}$ is a $\mathbf{Q}$-linear map.

   (b) What can you say about $\alpha \in K$ with $N_K(\alpha) = 0$?

   (c) What can you say about $\alpha \in K$ with $\mathrm{tr}_K(\alpha) = 0$?

   (d) Prove that there exists an algebraic number $\alpha \in \mathbf{C}$ with $N_{\mathbf{Q}(\alpha)}(\alpha)$, $\mathrm{tr}_{\mathbf{Q}(\alpha)}(\alpha) \in \mathbf{Z}$, but α is not an algebraic integer.

   (e) Let α be an algebraic number with minimal polynomial m_α. Show that for $c \in \mathbf{Q}$,

   $$N_{\mathbf{Q}(\alpha)}(\alpha + b) = (-1)^{\deg(\alpha)} m_\alpha(-b)$$

   More generally, for $a, b \in \mathbf{Q}$ give a similar formula for $N_{\mathbf{Q}(\alpha)}(a\alpha + b)$.

3. Let K be a number field. An embedding $\tau : K \to \mathbf{C}$ is called a *real embedding* if $\tau(K) \subseteq \mathbf{R}$. Otherwise it is called a *complex embedding*.

   (a) Show that the number of complex embeddings of K is even. We denote by r the number of real embeddings of K and by s half the number of complex embeddings. Show that $r + 2s = [K : \mathbf{Q}]$.

   The numbers r, s are called the *signature* of K and they give a refinement of the degree.

   (b) Let $\alpha \in \mathbf{C}$ be an algebraic number with minimal monic polynomial $m_\alpha \in \mathbf{Q}[x]$. Let r, s be the signature of $\mathbf{Q}(\alpha)$. Show that r is the number of real roots of m_α in $\mathbf{C}$ and s is half the number of complex roots in $\mathbf{C}$.

   (c) What is the signature of $\mathbf{Q}(\sqrt{d})$, $d \in \mathbf{Z}$ $d \neq 0, 1$, d squarefree?

   (d) What is the signature of $\mathbf{Q}(\sqrt[3]{2})$? What about $\mathbf{Q}(\sqrt[n]{2})$ and $\mathbf{Q}(\sqrt[n]{-2})$ for all $n > 2$?

   (e) (Bonus question!) Given a number field K an embedding $\tau_0 : K \to \mathbf{C}$, and an algebraic number $\alpha \in \mathbf{C}$ with minimal polynomial over K $m_{\alpha,K} \in K[x]$, how would you figure out how many of the embeddings $\tau : K(\alpha) \to \mathbf{C}$ restricting to τ_0 are real and complex?

4. (a) Let α be an algebraic integer with the property that $|\tau(\alpha)| \leq 1$ for all embeddings $\tau : \mathbf{Q}(\alpha) \to \mathbf{C}$ (here $|x+iy| = \sqrt{x^2+y^2}$ denotes the usual complex absolute value). Show that α is a root of unity. (Hint: try to prove that the set $\{\alpha^n \mid n \geq 0\}$ must be finite by bounding the coefficients of their minimal polynomials.)

   (b) Give an example of an algebraic number α which is not a root of unity for which $|\tau(\alpha)| = 1$ for all embeddings $\tau : \mathbf{Q}(\alpha) \to \mathbf{C}$.

5. (Bonus question!) Find the integer ring of $K = \mathbf{Q}(\sqrt[3]{d})$.

8 Lecture 8: Factorization, UFDs, and integrally closed rings

Let R be a commutative ring. In this course, commutative rings always have 1 and ring homomorphisms always send 1 to 1.

- R is called an *integral domain* if $0 \neq 1$[28] in R and if $a, b \in R$ are nonzero then $ab \neq 0$. Integral domains have the cancelation property: if $ab = ac$ and $a \neq 0$ then $b = c$. It is not really reasonable to discuss factorization without this property.

- $a \in R$ is called a unit if there exists $b \in R$ with $ab = 1$. We denote the set of units by $R^\times$, and it forms a group under multiplication.

- Two elements $a, b \in R$ are said to be *associate* if there exists a unit $u \in R^\times$ with $a = ub$. This is an equivalence relation. Equivalently $Ra = Rb$, i.e. the principal ideals they generate are the same: $(a) = (b)$. We don't want to consider $15 = 3 \cdot 5 = (-3) \cdot (-5)$ a failure of unique factorization and so we only ask for factorization up to associates.[29]

From now on R is an integral domain. There are two ways of saying that $p \in \mathbf{Z}$ is prime which lead to different notions in general:

- A nonzero, non unit $\pi \in R$ is irreducible if $\pi = ab$ implies that either a or b is a unit.

- A nonzero, non unit $\pi \in R$ is prime if $\pi | ab$ implies $\pi | a$ or $\pi | b$.

Lemma 8.1. *If π is prime then π is irreducible.*

Proof. If π is prime and $\pi = ab$ then wlog $\pi | a$ and so $a = \pi c$, and so $\pi = ab = \pi bc$, hence $bc = 1$ so b is a unit. □

The converse is not true and is closely related to the failure of unique factorization.

[28] Recall there is a ring $R = \{0\}$ with $0 + 0 = 0$, $0 \cdot 0 = 0$ called the zero ring. We don't want to call it an integral domain.

[29] We can try to get around this by picking a representative for each equivalence class. In $\mathbf{Z}$ we can use positive integers and in $K[x]$ we can use monic polynomials. But in a general ring R there is not usually a natural choice.

Example 8.2. Consider $\mathcal{O}_{-5} = \mathbf{Z}[\sqrt{-5}]$. We have

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

Using norms we can see that $2, 3, (1 + \sqrt{-5}), (1 - \sqrt{-5})$ are all irreducible. Indeed, their norms are $4, 9, 6, 6$ and $N(a + b\sqrt{-5}) = a^2 + 5b^2 = 2, 3$ has no solutions. None of them are prime. For instance $2 \mid (1 + \sqrt{-5})(1 - \sqrt{-5})$ but $2 \nmid (1 + \sqrt{-5}), (1 - \sqrt{-5})$ so 2 is not prime.

Definition 8.3. R is called a unique factorization domain (UFD) if any $0 \neq a \in R$ can be written $a = u\pi_1 \cdots \pi_r$ where $u \in R^\times$ and $\pi_i \in R$ for $i = 1, \ldots, r$ are irreducible, and moreover, this expression is unique in the sense that the π_i are uniquely determined up to ordering and associates.

Remark 8.4. We tend to think of the uniqueness as being the essential part, but even the existence of factorizations into irreducibles can fail. For example, in the ring $\overline{\mathbf{Z}}$ of all algebraic integers, we can indefinitely factor

$$2 = (\sqrt{2})^2 = (\sqrt[4]{2})^4 = (\sqrt[8]{2})^8 = \cdots$$

In fact, this shows that $\overline{\mathbf{Z}}$ has no irreducible elements at all! This is related to the failure of the Noetherian property, as we will discuss later.

Proposition 8.5. *In a UFD R, irreducible elements are prime.*

Proof. Let $\pi \in R$ be irreducible and suppose that $\pi \mid ab$ so that $\pi c = ab$. We consider factorizations of a, b, c into irreducibles:

$$\pi(u''\pi_1'' \cdots \pi_t'') = (u\pi_1 \cdots \pi_r)(u'\pi_1' \cdots \pi_s')$$

Uniqueness of factorization implies that the irreducible π must be associate to one of the irreducibles on the other side. If π is associate to one of $\pi_1, \ldots, \pi_r$ then $\pi | a$ and if π is associate to one of $\pi_1', \ldots, \pi_s'$ then $\pi | b$. □

We recall that an integral domain R has a fraction field $\mathrm{Frac}(R)$. Formally it is defined as the set of fractions $\frac{a}{b}$ where $a, b \in R$, $b \neq 0$, and two fractions $\frac{a}{b}, \frac{a'}{b'}$ are declared equal if $ab' = a'b$. One checks that this defines an equivalence relation and that $\mathrm{Frac}(R)$ becomes a field when endowed with addition and multiplication defined in the usual way:

$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + a'b}{bb'}, \quad \frac{a}{b} \cdot \frac{a'}{b'} = \frac{aa'}{bb'}.$$

Moreover R may be identified with the subring of elements $\{\frac{a}{1}\} \subseteq \mathrm{Frac}(R)$.

Example 8.6. We have $\mathrm{Frac}(\mathbf{Z}) = \mathbf{Q}$, and more generally given a number field K, we have $\mathrm{Frac}(\mathcal{O}_K) = K$. Indeed we have seen that any $\alpha \in K$ may be written as $\frac{n\alpha}{n}$ for some integer $n > 0$.

More generally, given any subring $\mathcal{O} \subseteq \mathcal{O}_K$ which contains a basis for K as a $\mathbf{Q}$-vector space[30], we have that $\mathrm{Frac}(\mathcal{O}) = K$. Examples include $\mathbf{Z}[\sqrt{-3}] \subset \mathcal{O}_{-3}$, or $\mathbf{Z}[2i] \subseteq \mathbf{Z}[i]$. More generally if $K = \mathbf{Q}(\alpha)$ for an algebraic integer α, we have $\mathrm{Frac}(\mathbf{Z}[\alpha]) = K$.

[30]Such a subring is called an *order* in K.

Definition 8.7. An integral domain R is said to be integrally closed if for any monic polynomial $f \in R[x]$ and any $\alpha \in \mathrm{Frac}(R)$ with $f(\alpha) = 0$, we have $\alpha \in R$.

The importance of this definition may be hard to appreciate at first, but it clearly is connected with the notion of algebraic integers. Indeed we have already proved the following:

Proposition 8.8. *Let K be a number field. Then $\mathcal{O}_K$ is integrally closed.*

Proof. If $f \in \mathcal{O}_K[x]$ is a monic polynomial and $\alpha \in K$ with $f(\alpha) = 0$, then by Theorem 3.1 part 2, α is an algebraic integer. Hence α in an algebraic integer in K so it is in $\mathcal{O}_K$ by definition. □

The following proposition is a first hint that this is a desirable property:

Proposition 8.9. *Let R be a UFD. Then R is integrally closed.*

Proof. Let $f \in R[x]$ be monic and let $\frac{a}{b} \in \mathrm{Frac}(R)$ satisfy $f(\frac{a}{b}) = 0$. Because R is a UFD, we may assume that $\frac{a}{b}$ is a fraction in lowest terms: if $\pi \in R$ is prime we do not have both $\pi|a$ and $\pi|b$, as otherwise we can replace $\frac{a}{b}$ with $\frac{a/\pi}{b/\pi}$, and because a and b have only finitely many prime factors we can repeat this until it is the case.

Now if b is a unit then $\frac{a}{b} = \frac{ab^{-1}}{1} \in R$. If not, we can choose a prime $\pi \in R$ with $\pi|b$. Let $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. We consider $0 = b^n f(\frac{a}{b})$. We obtain:

$$a^n = -a_{n-1}ba^{n-1} - a_{n-2}b^2a^{n-2} - \cdots - a_0b^n.$$

The right hand side is a multiple of b hence of π. So $\pi|a^n$, and hence $\pi|a$ as R is a UFD. This is a contradiction. □

Corollary 8.10. *Let K be a number field and let $\mathcal{O} \subsetneq \mathcal{O}_K$ be any subring with $\mathrm{Frac}(\mathcal{O}) = K$. Then $\mathcal{O}$ is not integrally closed, and hence $\mathcal{O}$ is not a UFD.*

Proof. Any element $\alpha \in \mathcal{O}_K$, $\alpha \notin \mathcal{O}$ satisfies a monic polynomial in $\mathbf{Z}[x] \subset \mathcal{O}[x]$. □

Example 8.11. In particular, if $d \equiv 1 \pmod 4$, the ring $\mathbf{Z}[\sqrt{d}] \subsetneq \mathcal{O}_d$ is never a UFD. Similarly $\mathbf{Z}[2i]$ etc is not a UFD.

Exercise 8.12. Give an explicit example of unique factorization failing in $\mathbf{Z}[\sqrt{-3}]$ and $\mathbf{Z}[2i]$.

9 Lecture 9: Ideals

Throughout this section R denotes a commutative ring with 1.

- A subset $I \subseteq R$ is called an ideal if it is a subgroup under addition and if for all $b \in R$ and $a \in I$ we have $ba \in I$.

  We can then form the quotient ring R/I: the elements are cosets $a + I$ for $a \in R$ and addition and multiplication are defined by the formulas

  $$(a + I) + (b + I) = (a + b + I) \quad (a + I)(b + I) = (ab + I).$$

  One way you could have discovered the definition of an ideal by thinking about what condition is necessary to make this multiplication well defined.

- The ideal generated by a subset $S \subseteq R$ is the smallest ideal containing the set S, given explicitly by
$$(S) = \{\sum_{i=1}^{n} r_i s_i \mid r_i \in R, s_i \in S\}$$
the set of all R-linear combinations of elements of S. Given finitely many elements $a_1, \ldots, a_n$ we also write $(a_1, \ldots, a_n)$ for the ideal generated by $\{a_1, \ldots, a_n\}$. Explicitly we have
$$(a_1, \ldots, a_n) = \{\sum_{i=1}^{n} b_i a_i \mid b_i \in R\}$$
the set of all R-linear combinations of $a_1, \ldots, a_n$.

- We say that an ideal I is *principal* if it can be generated by a single element, i.e. $I = (a) = Ra$ for some $a \in R$.

- An integral domain is called a principal ideal domain (PID) if every ideal is principal.[31]

Remark 9.1. We have $(a) \subseteq (b)$ if and only if $a \in (b)$ if and only if $a = br$ for some $r \in R$, i.e. $b|a$. If R is an integral domain then $(a) = (b)$ if and only if a and b are associate in the sense of the last lecture. Indeed if $(a) = (b)$ then by considering both inclusions we have $a = br$ and $b = ar'$ for some $r, r' \in R$, and so $a = arr'$ and hence $rr' = 1$, and so $r, r' \in R^\times$.

Example 9.2. The fundamental examples of PIDs are $\mathbf{Z}$ and $K[x]$ for K a field. They are proved to be PIDs using the Euclidean property as we will review shortly.

In your algebra course you saw a proof of

Theorem 9.3. *If R is a PID then R is a UFD.*

We won't repeat the proof now, but later in the course we will prove a generalization of this.

Given ideals $I, J \subseteq R$ there are three ways to produce a new ideal:

- $I \cap J$
- $I + J = \{a + b \mid a \in I, b \in J\}$
- $I \cdot J = \{\sum_{i=1}^{n} a_i b_i \mid a_i \in I, b_i \in J\}$

The ideal product is going to play an extremely important role. We emphasize that $I \cdot J$ is *not* in general equal to $\{ab \mid a \in I, b \in J\}$. This set is not necessarily closed under addition.

Often ideals are expressed in terms of generators. Then you have the following formulas:

- $(a_1, \ldots, a_n) + (b_1, \ldots, b_m) = (a_1, \ldots, a_n, b_1, \ldots, b_m)$.
- $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_m) = (a_i b_j)_{1 \leq i \leq n, 1 \leq j \leq m}$.

[31] There are rings, like $\mathbf{Z}/nZ$ for n not prime, in which every ideal is principal but are not integral domains.

However there is no simple formula for generators of $I \cap J$ in terms of generators of I and J.

This is all a bit abstract, but we will soon do some very explicit examples. Perhaps the most important thing to keep in mind for now is what these operations correspond to in $R = \mathbf{Z}$, or more generally R a PID. Here is a dictionary:

1. $(a) \subseteq (b) \Leftrightarrow a \in (b) \Leftrightarrow b|a$ "to contain is to divide"
2. $(a) \cdot (b) = (ab)$ "multiplication is multiplication"
3. $(a) + (b) = (a, b) = (\gcd(a, b))$ "addition is gcd"
4. $(a) \cap (b) = (\text{lcm}(a, b))$ "intersection is lcm"

Exercise 9.4. Check all these things for $R = \mathbf{Z}$.

Example 9.5. We return to the failure of unique factorization in $\mathbf{Z}[\sqrt{-5}]$

$$2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

In the integers if $a|b \cdot c$ and $\gcd(b, c) = 1$ then $b = \gcd(a, b) \cdot \gcd(a, c)$. Let's test this formula in $\mathbf{Z}[\sqrt{-5}]$ for $(1 + \sqrt{-5}) \mid 2 \cdot 3$:

$$(2, 1 + \sqrt{-5})(3, 1 + \sqrt{-5}) = (6, 2(1 + \sqrt{-5}), 3(1 + \sqrt{-5}), (1 + \sqrt{-5})^2) = (1 + \sqrt{-5}).$$

In the first equality we just multiplied out all pairs of generators. To check the second equality, we note that all four generators are multiples of $1 + \sqrt{-5}$, proving $\subseteq$ and then note that $1 + \sqrt{-5} = 3(1 + \sqrt{-5}) - 2(1 + \sqrt{-5})$ is in the ideal on the left.

So that looks good! Being a bit more systematic let:

$$\begin{aligned}
\mathfrak{p}_2 &= (2, 1 + \sqrt{-5}) = (2, 1 - \sqrt{-5}) = (1 + \sqrt{-5}, 1 - \sqrt{-5}) \\
\mathfrak{p}_3 &= (3, 1 + \sqrt{-5}) \\
\overline{\mathfrak{p}}_3 &= (3, 1 - \sqrt{-5})
\end{aligned}$$

and then we have the formulas:

$$(2) = \mathfrak{p}_2^2, \quad (3) = \mathfrak{p}_3\overline{\mathfrak{p}}_3, \quad (1 + \sqrt{-5}) = \mathfrak{p}_2\mathfrak{p}_3, \quad (1 - \sqrt{-5}) = \mathfrak{p}_2\overline{\mathfrak{p}}_3$$

so it appears everything is factoring as intended into the three "primes" $\mathfrak{p}_2$, $\mathfrak{p}_3$, $\overline{\mathfrak{p}}_3$. Finally we note that $(6) = \mathfrak{p}_2^2\mathfrak{p}_3\overline{\mathfrak{p}}_3$.

Exercise 9.6. Check that the three ideals in the definition of $\mathfrak{p}_2$ are all equal, and then check the four formulas below (we already did $(1 + \sqrt{-5}) = \mathfrak{p}_2\mathfrak{p}_3$ and you can use conjugation to deduce $(1 - \sqrt{-5}) = \mathfrak{p}_2\overline{\mathfrak{p}}_3$.)

We now recall it what sense these ideals are prime.

Definition 9.7. An ideal $I \subsetneq R$ is called prime if for any $b, c \in R$, if $bc \in I$ then $b \in I$ or $c \in I$.[32]

[32] We do not want to call $R = (1)$ a prime ideal for the same reason we don't want to call 1 a prime number.

If R is an integral domain, then in the last lecture we defined what it means for an element $0 \neq a \in R$ to be prime. We observe that such an a is prime if and only if the ideal (a) is prime. Indeed:

- $bc \in (a)$ if and only if $a \mid bc$.
- $b \in (a)$ or $c \in (a)$ if and only if $a|b$ or $a|c$.

It is often useful to rephrase the condition that an ideal is prime in terms of the quotient ring R/I.

Proposition 9.8. *An ideal $I \subseteq R$ is prime if and only if the quotient ring R/I is an integral domain.*

Proof. The key point is that for $a \in R$, we have $a \in I$ if and only if $a + I = 0 + I$, i.e. $a+I$ is 0 in the quotient ring R/I, and so we can translate any statement about elements of R being in I to a statement about elements of R/I being zero.

Then we check two things. First $I \neq R$ if and only if $1 \notin I$[33] if and only if $1 \neq 0$ in R/I. Thus $I \subsetneq R$ if and only if $1 \neq 0$ in R/I.

Second, the statement "for all $a, b \in R$, $a, b \notin I$ implies $ab \notin I$" is equivalent to the statement "for all $a+I, b+I \in R/I$, $a+I, b+I \neq 0$ in R/I implies $(a+I)(b+I) = ab+I \neq 0$ in R/I." □

Example 9.9. Consider the ideal $\mathfrak{p}_3 = (3, 1+\sqrt{-5}) \subset \mathcal{O}_{-5}$. We will show that $\mathcal{O}_{-5}/\mathfrak{p}_3 \simeq \mathbf{F}_3 = \mathbf{Z}/3\mathbf{Z}$, and hence $\mathfrak{p}_3$ is a prime ideal. To see this we note that for $a, b \in \mathbf{Z}$:

$$a + b\sqrt{-5} = a - b + b(1 + \sqrt{-5}) = r + (3q + b(1 + \sqrt{-5}))$$

if $a - b = 3q + r$ is the unique expression with $r \in \{0, 1, 2\}$. So see that the cosets $0 + \mathfrak{p}_3, 1 + \mathfrak{p}_3, 2 + \mathfrak{p}_3$ really are distinct, it suffices to check that $1 \notin \mathfrak{p}_3$. We can check this directly:

$$1 = 3(a + b\sqrt{-5}) + (1 + \sqrt{-5})(c + d\sqrt{-5}) = (3a + c - 5d) + (3b + c + d)\sqrt{-5}.$$

Then $3b + c + d = 0$ implies $c + d \equiv 0 \pmod 3$, and hence $3a + c - 5d \equiv c + d \equiv 0 \pmod 3$. Thus $1 \notin \mathfrak{p}_3$.

10 Lecture 10: Integer rings as Euclidean domains, Dedekind domains

Now that we have reviewed UFDs and PIDs, we finally turn briefly to Euclidean domains. Here is the definition:

Definition 10.1. Let R be an integral domain. A function $f : R \to \mathbf{N}$ is said to be a *Euclidean function* if for all $a, b \in R$, $b \neq 0$, we can write

$$a = bq + r$$

for some $q, r \in R$, and either $r = 0$ or $f(r) < f(b)$.

R is said to be a *Euclidean domain* if a Euclidean function exists.

[33] If $1 \in I$ then $a \cdot 1 = a \in I$ for any $a \in R$ so $I = R$

Here are some basic consequences of this definition, that you have seen in your algebra course:

- Euclidean domains are PIDs. (Proof: if $I \subset R$ is a nonzero ideal, you take $0 \neq b \in I$ with $f(b)$ minimal among all nonzero elements of I. Then for any $a \in I$ you have $a = bq + r$ and $r = a - bq \in I$. You can't have $f(r) < f(b)$ based on how b was chosen so you must have $r = 0$. Thus $I = (b)$.)
- In a Euclidean domain you can use the Euclidean algorithm to find d with $(d) = (a, b)$, and to write $d = xa + yb$.
- The basic examples of Euclidean domains are $\mathbf{Z}$, where the Euclidean function is $|\cdot|$, and $K[x]$ for K a field, where the Euclidean function is given by the degree.

For integer rings of number fields we have a natural candidate for a Euclidean function:

Definition 10.2. Let K be a number field. We say that $\mathcal{O}_K$ is *norm Euclidean* if $|N_K| : \mathcal{O}_K \to \mathbf{N}$ is a Euclidean function.

Using the multiplicativity of the norm we can rewrite the condition of being norm Euclidean:

Lemma 10.3. *Let K be a number field. Then $\mathcal{O}_K$ is norm Euclidean if and only if for all $\alpha \in K$ there exists $\beta \in \mathcal{O}_K$ with $|N_K(\alpha - \beta)| < 1$.*

Proof. Given $a, b \in \mathcal{O}_K$ with $b \neq 0$, we take $\alpha = \frac{a}{b}$ and choose $\beta \in \mathcal{O}_K$ with $|N_K(\alpha - \beta) < 1|$. Then given $\beta \in \mathcal{O}_K$ with $|N_K(\alpha - \beta)| < 1$, we take $q = \beta$ and $r = a - bq = b(\alpha - \beta)$. Then using the multiplicativity of the norm we find

$$|N_K(r)| = |N_K(b)||N_K(\alpha - \beta)| < |N_K(b)|$$

and hence $|N_K|$ defines a Euclidean function on $\mathcal{O}_K$.

Conversely given $\alpha \in K$, we write $\alpha = \frac{a}{b}$ for $a, b \in \mathcal{O}_K$, $b \neq 0$. Then using the Euclidean property we can find $q, r \in \mathcal{O}_K$ with $a = qb + r$ and $|N_K(r)| < |N_K(b)|$. Dividing by b we see that

$$|N_K(\alpha - q)| = |N_K(b^{-1}(a - qb))| = |N_K(b)|^{-1}|N_K(r)| < 1.$$

□

Theorem 10.4. *For $d < 0$ squarefree the following are equivalent*

1. *$\mathcal{O}_d$ is norm Euclidean.*
2. *$\mathcal{O}_d$ is Euclidean.*
3. *$d = -1, -2, -3, -7, -11$.*

Proof. We sketch the proof that 1 and 3 are equivalent. For the proof that 2 and 3 are equivalent, see example sheet 2.

We view $\mathcal{O}_{-d} \subseteq \mathbf{C}$ as a lattice. For $\beta \in \mathcal{O}_{-d}$, the equation $|N(\alpha - \beta)| < 1$ defines a circle of radius 1 around β. So we are really asking: is $\mathbf{C}$ covered by the (open) circles of radius 1 about the points of $\mathcal{O}_{-d}$?

When $d \not\equiv 1 \pmod 4$ we see that the lattice points form rectangles of width 1 and height $\sqrt{-d}$, and the points farthest from lattice points are at the center of these rectangles, with distance squared $\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{-d}}{2}\right)^2 = \frac{1-d}{4}$. Thus $\mathbf{Z}[\sqrt{-d}]$ is norm Euclidean if and only if $-d < 3$, or $d = -1, -2$.

When $d \equiv 1 \pmod 4$, the lattice points form isosceles triangles with base 1 and height $\frac{\sqrt{-d}}{2}$. The farthest point from the vertices is at the center of the triangle (figure to be added...). Its distance r from the vertices thus satisfies

$$r^2 = \frac{1}{4} + (\frac{\sqrt{-d}}{2} - r)^2$$

and hence

$$r = \frac{1-d}{4\sqrt{-d}} < 1$$

which gives $d = -3, -7, -11$. □

In a few weeks, we will prove the following theorem by computing class groups:

Theorem 10.5. *The rings $\mathcal{O}_{-19}, \mathcal{O}_{-43}, \mathcal{O}_{-67}, \mathcal{O}_{-163}$ are PIDs.*

Thus these rings are examples of PIDs which are not Euclidean domains! This already shows the limitations of the theory of Euclidean domains in algebraic number theory.

What about other number fields? It turns out that beyond the case of imaginary quadratic fields, the Euclidean property becomes substantially more difficult to understand. We consider the real quadratic case, $\mathcal{O}_d$ for $d > 0$.

How can we picture $\mathcal{O}_d$? $\mathcal{O}_d \subseteq \mathbf{R}$ is dense, so that doesn't give us a very useful picture. Instead, given $\alpha \in \mathbf{Q}(\sqrt{d})$ it is better to picture $(\tau_1(\alpha), \tau_2(\alpha)) \in \mathbf{R}^2$, where

$$\tau_1(a + b\sqrt{d}) = a + b\sqrt{d}, \quad \tau_2(a + b\sqrt{d}) = a - b\sqrt{d}$$

are the two embeddings.

Then given a $\beta \in \mathcal{O}_d$, what does the set of $\alpha \in K$ satisfying $|N(\alpha - \beta)| < 1$ look like? Well letting $x = \tau_1(\alpha)$, $y = \tau_2(\alpha)$, we have

$$|N_K(\alpha - \beta)| = |\tau_1(\alpha - \beta)\tau_2(\alpha - \beta)| = |(x - \tau_2(\beta))(y - \tau_2(\beta))|$$

In other words, in $\mathbf{R}^2$, the region $N(\alpha - \beta) < 1$ looks like the interior of a hyperbola centered at $(\tau_1(\beta), \tau_2(\beta))$.

Then $\mathcal{O}_d$ is norm Euclidean if and only if the infinitely many hyperbolas centered around the points of $\mathcal{O}_d$ cover $\mathbf{R}^2$. This question is much more subtle than the circles in $\mathbf{C}$ in the imaginary quadratic case: in fact the union of these hyperbolas will always be dense in $\mathbf{R}^2$.

Just for fun here are some facts about when $\mathcal{O}_d$, $d > 0$ are (norm) Euclidean:

Theorem 10.6. *Let $d > 0$ be squarefree. Then $\mathcal{O}_d$ is norm Euclidean if and only if $d = 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73$.*

This is an extremely difficult theorem, proved over decades by many mathematicians, ending in the start of the 1950s.

We finally consider the case of $\mathcal{O}_{14} = \mathbf{Z}[\sqrt{14}]$:

- $\mathbf{Z}[\sqrt{14}]$ is a PID. We will easily be able to prove this soon by computing its class group.
- $\mathbf{Z}[\sqrt{14}]$ is not norm Euclidean. This is a problem on example sheet 2.
- $\mathbf{Z}[\sqrt{14}]$ is Euclidean! This is an extremely difficult theorem, only proved in 2004 by Malcolm Harper.

So to summarize our discussion of the Euclidean property in algebraic number theory: it is in general a quite difficult problem to determine when $\mathcal{O}_K$ is Euclidean or norm Euclidean. On the other hand if the main interest in the Euclidean property is to check that $\mathcal{O}_K$ is a PID, then ultimately it is not so useful because we will soon have a much better technique for doing this, by computing the class group.

10.1 Dedekind domains

Definition 10.7. Let R be an integral domain. R is called a *Dedekind domain* if it satisfies the following three conditions:

1. R is integrally closed (see definition 8.7).
2. Every nonzero prime ideal of R is maximal.[34]
3. R is Noetherian.

We recall the definitions of maximal ideals and Noetherian rings:

Definition 10.8. An ideal $I \subsetneq R$ is *maximal* if for any ideal $I \subseteq J \subseteq R$ for some ideal J implies $J = R$ or $J = I$.

We recall from your algebra course that an ideal $I \subseteq R$ is maximal if and only if R/I is a field,[35] and for every ideal $J \subsetneq R$ there is a maximal ideal $I \subset R$ with $I \subseteq J$.

Definition 10.9. A ring R is called *Noetherian* if it satisfies the following three equivalent conditions:

1. Given ideals
$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$
of R there is some integer n such that $I_n = I_{n+1} = I_{n+2} = \cdots$.
2. Any nonempty set of ideals in R has a maximal element.
3. Any ideal $I \subseteq R$ is finitely generated.

For completeness, we give here in the notes a proof that these conditions are equivalent. We will only use notion during the proof of the main theorem about factorization into prime ideals in Dedekind domains in the next lecture.

[34] In terminology you might have seen in commutative algebra or algebraic geometry, this is equivalent to saying that R has Krull dimension ≤ 1.

[35] The proof of this is to show that giving an ideal $I \subseteq J \subseteq R$ is the same as giving an ideal $(0) \subseteq J' \subseteq R/I$, and then a ring R is a field if and only if $(0) \subsetneq R$ are the only two ideals.

Proof. We show 1 implies 2. Given a nonempty set S of ideals we pick $I_1 \in S$. If I_1 is not maximal then we can pick $I_1 \subsetneq I_2 \in S$. If I_2 is not maximal we pick $I_2 \subsetneq I_3 \in S$. Repeating we either find a maximal element of S, or we construct a chain

$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$$

of ideals, contradicting 1.[36]

We show 2 implies 3. Given an ideal $I \subseteq R$ we consider the set S of finitely generated ideals contained in I. Since $(0) \subset I$ is finitely generated, this set is nonempty. Let J be a maximal element of S. If $J \subsetneq I$, pick $a \in I$, $a \notin J$ and consider $J \subsetneq J + (a) \subseteq I$, a strictly larger finitely generated ideal contained in I. This is a contradiction, so we must $J = I \in S$ and hence I is finitely generated.

We show 3 implies 1. Given a chain

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

we consider $I = \cup_{n=1}^{\infty} I_n$. This is an ideal of R. Indeed given $a, b \in I$ we have $a, b \in I_n$ for some n and hence $ra + sb \in I_n \subseteq I$. Now I is finitely generated, say $I = (a_1, \ldots, a_r)$. Each $a_i \in I_{n_i}$ for some n_i. Since there are only finitely many we may take n to be the maximum of the n_i and have $a_i \in I_n$ for $i = 1, \ldots, r$. Then $I_n = I_{n+1} = \cdots = I$. □

Proposition 10.10. *Let K be a number field. Then $\mathcal{O}_K$ is a Dedekind domain.*

We prepare with a lemma.

Lemma 10.11. *Let $I \subseteq \mathcal{O}_K$ be a nonzero ideal. Then I has finite index in $\mathcal{O}_K$, that is $[\mathcal{O}_K : I] = \#\mathcal{O}_K/I < \infty$.*

Proof. Pick $0 \neq \alpha \in I$. Then we claim that

$$(N_K(\alpha)) \subseteq (\alpha) \subseteq I.$$

To see that $(N_K(\alpha)) \subseteq (\alpha)$ we use an argument similar to the proof of Proposition 4.9. By Theorem 4.4 we can write

$$N_K(\alpha) = \alpha \prod_{\tau \in \Sigma_K, \tau \neq \tau_{id}} \tau(\alpha).$$

Then as in the proof of Proposition 4.9 we show that $\prod_{\tau \in \Sigma_K, \tau \neq \tau_{id}} \tau(\alpha) \in \mathcal{O}_K$. It is an algebraic integer because it is a product of algebraic integers, and it is in K because it equals $\alpha^{-1} N_K(\alpha) \in K$, hence it is in $\mathcal{O}_K$.

Now we have

$$[\mathcal{O}_K : I] \leq [\mathcal{O}_K : (N_K(\alpha))]$$

so it suffices to show $[\mathcal{O}_K : (N)] < \infty$, where $N \in \mathbf{Z}$, $N \neq 0$. Now by theorem 5.1, $\mathcal{O}_K \simeq \mathbf{Z}^n$ as abelian groups where $n = [K : \mathbf{Q}]$. Under this isomorphism, $(N) = N\mathcal{O}_K$ goes to $N\mathbf{Z}^n$. Hence

$$[\mathcal{O}_K : (N)] = [\mathbf{Z}^n : N\mathbf{Z}^n] = |N|^n < \infty.$$

□

[36] In creating this chain you need to make infinitely many choices and so this argument runs in to set theoretic difficulties. You need a weak form of the axiom of choice, called the axiom of dependent choice. You cannot prove that 1 implies 2 in ZF!

Proof of Proposition 10.10. We have seen in Proposition 8.8 that $\mathcal{O}_K$ is integrally closed.

Now let $\mathfrak{p} \subset \mathcal{O}_K$ be a nonzero prime ideal. By Lemma 10.11, $R/\mathfrak{p}$ is finite and by Proposition 9.8 $R/\mathfrak{p}$ is an integral domain. Hence $R/\mathfrak{p}$ is a field, as finite integral domains are fields. Thus $\mathfrak{p}$ is a maximal ideal.

Finally to check that $\mathcal{O}_K$ is Noetherian, let

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

be an ascending chain of ideals in $\mathcal{O}_K$. We apply 10.11 to conclude that they have finite index and furthermore the indexes satisfy

$$\infty > [\mathcal{O}_K : I_1] \geq [\mathcal{O}_K : I_2] \geq [\mathcal{O}_K : I_3] \geq \cdots$$

Hence there must be some n for which $[\mathcal{O}_K : I_n] = [\mathcal{O}_K : I_{n+1}] = \cdots$ and hence $I_n = I_{n+1} = \cdots$. □

We now give some examples of integral domains which are not Dedekind domains because they fail one of the three properties:

1. We have seen that rings like $\mathbf{Z}[2i]$ and $\mathbf{Z}[\sqrt{-3}]$ are not integrally closed, and hence aren't Dedekind domains.

2. Rings like $\mathbf{Z}[x]$ or $\mathbf{C}[x, y]$ have nonzero prime ideals which aren't maximal: e.g. $(x) \subsetneq (2, x) \subsetneq \mathbf{Z}[x]$ and $(x) \subsetneq (x, y) \subsetneq \mathbf{C}[x, y]$.[37]

3. The ring $\overline{\mathbf{Z}}$ of all algebraic integers in $\mathbf{C}$ is not Noetherian. For example we have chains of ideals like

$$(2) \subseteq (2^{1/2}) \subseteq (2^{1/4}) \subseteq (2^{1/8}) \subseteq \cdots$$

We also have

Proposition 10.12. *Let R be a PID. Then R is a Dedekind domain.*

Proof. PIDs are UFDs and hence integrally closed by Proposition 8.9.

If $(a) \subset R$ is a prime ideal, then a is prime and hence irreducible. If $(a) \subseteq (b) \subseteq R$ is another ideal, then $a = bc$ for some $c \in R$. As a is irreducible, either b is a unit, i.e. $R = (b)$, or c is a unit, i.e. $(a) = (b)$. Hence $(a) \subset R$ is a maximal ideal.

PIDs are Noetherian since each ideal is generated by a single element. □

If you have taken algebraic geometry, another large class of Dedekind domains are the coordinate rings of smooth affine algebraic curves, i.e. rings such as

$$\mathbf{C}[x, y]/(x^n + y^n - 1).$$

The point of the notion of a Dedekind domain is that it is an abstraction of the key properties of $\mathcal{O}_K$ that are needed in the proof of unique factorization into prime ideals, that we give in the next lecture.

Exercise 10.13. $\mathbf{C}[x, y]/(y^2 - x^3)$ is an example of a singular affine algebraic curve. Show that this ring is not integrally closed. (This ring should somehow be thought of as an analog of the rings $\mathbf{Z}[2i]$, $\mathbf{Z}[\sqrt{-3}]$).

[37] In the language of commutative algebra and algebraic geometry, these rings have Krull dimension 2.

11 Lecture 11: Factorization into prime ideals

The goal for today is to prove the following theorem, which will be a fundamental tool in our study of rings of integers $\mathcal{O}_K$ for the rest of the course.

Theorem 11.1. *Let R be a Dedekind domain. Any nonzero ideal $I \subseteq R$ may be written in the form*

$$I = \mathfrak{p}_1 \mathfrak{p}_2 \cdots \mathfrak{p}_r$$

where $\mathfrak{p}_i \subseteq R$ for $i = 1, \ldots, r$ are prime ideals.[38] *Moreover, this expression is unique up to permuting the factors.*

Remark 11.2. As we saw in the last lecture, PIDs are Dedekind domains. For PIDs prime ideals $\mathfrak{p}_i = (a_i)$ for a_i prime, and the theorem exactly says that PIDs are UFDs.

The proof of Theorem 11.1 is the most technical and "abstract" part of this course. For the rest of this lecture, R will denote a Dedekind domain. We begin with several lemmas.

Lemma 11.3. *Let $I \subseteq R$ be an ideal and let $x \in \mathrm{Frac}(R)$ satisfy $xI \subseteq I$. Then $x \in R$.*

This lemma is the only place we will use that R is integrally closed. The proof will use the same "Cayley-Hamilton trick" as proposition 2.9.

Proof. As R is Noetherian we have $I = (a_1, \ldots, a_n)$. Then for $i = 1, \ldots, n$ we can write

$$\alpha a_j = \sum_{i=1}^{n} c_{ij} a_i$$

and build a matrix $A = (c_{ij}) \in M_n(R)$. Then as in the proof of Proposition 2.9, we let $\chi_A \in R[x]$ be the characteristic polynomial, and show that $\chi_A(\alpha)I = \chi_A(A)I = \{0\}$ by the Cayley-Hamilton theorem, and hence $\chi_A(\alpha) = 0$. But $\chi_A \in R[x]$ is a monic polynomial. Hence $\alpha \in R$ as R is integrally closed. □

Lemma 11.4. *Let $I \subseteq R$ be a nonzero ideal. Then there are nonzero prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ of R (not necessarily distinct) with $\mathfrak{p}_1 \mathfrak{p}_2 \cdots \mathfrak{p}_r \subseteq I$.*

Proof. Let S be the set of nonzero ideals of R for which the conclusion of the lemma fails. If S is empty we are done.

Otherwise, as R is Noetherian, we can pick a maximal element I of S. Now I cannot be a prime ideal, since the conclusion of the lemma clearly holds for nonzero prime ideals. Hence we can find $a, b \in R$ with $ab \in I$ but $a, b \notin I$. As $I \subsetneq I + (a), I + (b)$, the conclusion of the lemma must hold for $I + (a)$, $I + (b)$, i.e. there are nonzero primes $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$, $\mathfrak{q}_1, \ldots, \mathfrak{q}_s$ of R with $\mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq I + (a)$, $\mathfrak{q}_1 \cdots \mathfrak{q}_r \subseteq I + (b)$. But then

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \mathfrak{q}_1 \cdots \mathfrak{q}_s \subseteq (I + (a))(I + (b)) \subseteq I$$

which is a contradiction. □

The following lemma has nothing to do with Dedekind domains, it is just an alternative point of view on the definition of prime ideals:

[38] In this statement we should either exclude $I = R$, or allow $I = R$ to be the empty product by definition.

Lemma 11.5. *Let $\mathfrak{p} \subset R$ be a prime ideal and let $I, J \subseteq R$ be ideals. Then if $IJ \subseteq \mathfrak{p}$ then $I \subseteq \mathfrak{p}$ or $J \subseteq \mathfrak{p}$.*

Note that the definition of a prime ideal is that this holds when I, J are principal.

Proof. We suppose $I \not\subset \mathfrak{p}$ and prove $J \subseteq \mathfrak{p}$. Since $I \not\subset \mathfrak{p}$ we can pick $a \in I$ with $a \notin \mathfrak{p}$. Then for all $b \in J$, $ab \in IJ \subseteq \mathfrak{p}$, so $b \in \mathfrak{p}$. Thus $J \subseteq \mathfrak{p}$. □

Lemma 11.6. *Let $I \subsetneq R$ be a proper, nonzero ideal. Then there exists $x \in \mathrm{Frac}(R)$, $x \notin R$, such that $xI \subseteq R$.*

This lemma is the main point we will use that nonzero prime ideals are maximal in R. It is also the most unintuitive lemma in the proof of Theorem 11.1.

Proof. Pick $0 \neq a \in I$. By lemma 11.4 applied to (a), there exists nonzero primes $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ of R with $\mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq (a)$. Moreover we can and do choose r to be as small as possible.

Now as I is a proper ideal, we can pick a maximal ideal $\mathfrak{m}$ of R with $I \subseteq \mathfrak{m}$. Then

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq (a) \subseteq I \subseteq \mathfrak{m}$$

By Lemma 11.5 applied to $\mathfrak{m}$, we conclude that $\mathfrak{p}_i \subseteq \mathfrak{m}$ for some i. Relabelling, we assume $\mathfrak{p}_1 \subseteq \mathfrak{m}$. Now as R is a Dedekind domain, $\mathfrak{p}_1$ is also a maximal ideal, and hence $\mathfrak{p}_1 = \mathfrak{m}$.

Now because we chose r minimal, we have $\mathfrak{p}_2 \cdots \mathfrak{p}_r \not\subset (a)$. Hence we can pick $b \in \mathfrak{p}_2 \cdots \mathfrak{p}_r$ with $b \notin (a)$. Now behold:

$$bI \subseteq b\mathfrak{p}_1 \subseteq \mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq (a)$$

and hence

$$\frac{b}{a} I \subseteq R.$$

Thus we can take $x = \frac{b}{a}$, and $x \notin R$ because $b \notin (a)$. □

Exercise 11.7. Consider the ideal $I = (3, 1 + \sqrt{-5}) \subsetneq \mathcal{O}_{-5}$. Run through the argument of the proof of Lemma 11.6 to find $x \in \mathbf{Q}(\sqrt{-5})$, $x \notin \mathcal{O}_{-5}$ with $xI \subseteq \mathcal{O}_{-5}$. Does there exist such an x of the form $x = a^{-1}$, $a \in \mathcal{O}_{-5}$?

Exercise 11.8. Consider the ring $\mathbf{C}[x, y]$ and the ideal $(x, y) \subsetneq \mathbf{C}[x, y]$. Show that if $f \in \mathrm{Frac}(\mathbf{C}[x, y])$ with $f \cdot (x, y) \subseteq \mathbf{C}[x, y]$ then $f \in \mathbf{C}[x, y]$ (hint: $\mathbf{C}[x, y]$ is a UFD). At what point does the proof of Lemma 11.6 fail?

Now we combine our lemmas to prove the following proposition, which is the main step in the proof of Theorem 11.1.

Proposition 11.9. *Let R be a Dedekind domain, let $I \subseteq R$ be a nonzero ideal, and let $0 \neq a \in I$. Then there exists an ideal $J \subseteq R$ such that $IJ = (a)$.*[39]

Proof. Let

$$J = \{b \in R \mid bI \subseteq (a)\}.$$

We claim that J is an ideal. Indeed if $b_1, b_2 \in J$ then $(b_1 + b_2)I \subseteq b_1 I + b_2 I \subseteq (a) + (a) \subseteq (a)$, and similarly if $b \in J$ and $c \in R$ then $cbI \subseteq c(a) \subseteq (a)$.

[39] In terminology of commutative algebra, we are proving that I is an invertible ideal.

Moreover we clearly have $IJ \subseteq (a)$. If this is an equality, the proposition is proved. Assume for the sake of contradiction that $IJ \subsetneq (a)$. Then $K = \frac{1}{a}IJ \subsetneq R$ is a nonzero, proper ideal of R. We apply lemma 11.6 to find $x \in \mathrm{Frac}(R)$, $x \notin R$, with $xK \subseteq R$ and hence $xIJ \subseteq (a)$.

We observe that $aJ \subseteq IJ$ and hence $J \subseteq \frac{1}{a}IJ = K$. Hence $xJ \subseteq xK \subseteq R$. We claim that in fact $xJ \subseteq J$. Indeed, if $b \in J$ then $xb \in R$ as we have just seen, but

$$xbI \subseteq xJI \subseteq (a)$$

hence $xb \in J$ by the definition of J. Thus $xJ \subseteq J$ and so by lemma 11.3, $x \in R$. This is a contradiction. □

Now you can forget about the lemmas above and just remember Proposition 11.9! We prove two corollaries:

Corollary 11.10 (Cancellation)**.** *If $I, J, K \subseteq R$ are nonzero ideals and $IJ = IK$ then $J = K$.*

Proof. By proposition 11.9 we can find an ideal I' so that $II' = (a)$ is principal with $a \neq 0$. Then multiplying both sides of $IJ = IK$ by I' we obtain $aJ = aK$ and hence $J = K$. □

Corollary 11.11 (To contain is to divide)**.** *If $I, J \subseteq R$ are nonzero ideals, then $I \subseteq J$ if and only if there exists an ideal $K \subseteq R$ with $I = JK$.*

Note that proposition 11.9 is a special case of this when $J = (a)$ is principal.

Proof. The reverse direction is easy: we have $I = JK \subseteq J$.

For the forward direction, suppose $I \subseteq J$ and apply Proposition 11.9 to J to obtain an ideal J' with $JJ' = (a)$ for some $a \in J$ nonzero. Then noting that $IJ' \subseteq JJ' = (a)$ we have $K = a^{-1}IJ' \subseteq R$ is an ideal. Moreover, $JK = a^{-1}IJJ' = a^{-1}I(a) = I$. □

Now we are ready to prove Theorem 11.1.

Proof of Theorem 11.1. We first prove every nonzero ideal is a product of prime ideals. Let S be the set of all ideals of R which are not a product of prime ideals. If S is empty we are done. Suppose not. Then as R is Noetherian we can pick a maximal element I of S. We may pick a maximal ideal $\mathfrak{p}$ with $I \subseteq \mathfrak{p}$. Then by Corollary 11.10, $I = \mathfrak{p}J$ for some ideal J. But $I = \mathfrak{p}J \subseteq J$. If $I = J$ then by Corollary 11.10, $\mathfrak{p} = R$, a contradiction. Thus $I \subsetneq J$. It follows that $J = \mathfrak{p}_1 \cdots \mathfrak{p}_r$ is a product of prime ideals, and hence so is $I = \mathfrak{p}J = \mathfrak{p}\mathfrak{p}_1 \cdots \mathfrak{p}_r$, a contradiction.

Now we prove uniqueness. Suppose $\mathfrak{p}_1 \cdots \mathfrak{p}_r = \mathfrak{q}_1 \cdots \mathfrak{q}_s$ where the $\mathfrak{p}_i$ and $\mathfrak{q}_i$ are nonzero prime ideals. We prove by induction on r that $r = s$ and the $\mathfrak{p}_i$ and $\mathfrak{q}_i$ are the same up to reordering. We have

$$\mathfrak{q}_1 \cdots \mathfrak{q}_s = \mathfrak{p}_1 \cdots \mathfrak{p}_s \subseteq \mathfrak{p}_1$$

By Lemma 11.5, we have $\mathfrak{q}_i \subseteq \mathfrak{p}_1$. After renumbering we may assume $\mathfrak{q}_1 \subseteq \mathfrak{p}_1$. As $\mathfrak{q}_1$ is a nonzero prime and hence maximal, we must have $\mathfrak{q}_1 = \mathfrak{p}_1$. By Corollary 11.10 we must then have $\mathfrak{p}_2 \cdots \mathfrak{p}_r = \mathfrak{q}_2 \cdots \mathfrak{q}_s$, and then we are done by induction. □

Remark 11.12. Here is a remark about set theory, since somebody asked and I said something slightly wrong in lecture. The question of whether or not Theorem 11.1 can be proved in ZF depends on how you define Noetherian rings, as the equivalence of the three conditions in definition 10.9 actually cannot be proved in ZF (more precisely that definition 1 implies 2 cannot). You need at least a weak form of the axiom of choice called the "axiom of dependent choice". Similarly you cannot prove a PID is a UFD in ZF, you again need at least the axiom of dependent choice. (In lecture I might have said "axiom of *countable* choice" instead of "axiom of *dependent* choice" and I also said "now we use Zorn's lemma" when we found $I \subseteq \mathfrak{m}$ for $\mathfrak{m}$ maximal in the proof of lemma 11.6 but really we don't need Zorn's lemma, we just use R is Noetherian in sense 2. again.)

None of this bothers me because you can prove Theorem 11.1 for $R = \mathcal{O}_K$ in ZF, and that is really all I care about! In fact, you can convince yourself that all the awkward arguments where we made use of R being Noetherian of the form "let S be the set of ideals where the result fails..." can be replaced with clearer direct arguments by induction on the index $[\mathcal{O}_K : I]$.

12 Lecture 12: Class groups, CRT, and Dedekind domain miscellany

We are now ready to define the class group, as teased in the first lecture.

Definition 12.1. Let R be a Dedekind domain. We say that nonzero ideals $I, J \subseteq R$ are in the same ideal class, and write $I \sim J$, if there exist $\alpha \in \mathrm{Frac}(R)^\times$ with $\alpha I = J$. This defines an equivalence relation on the set of nonzero ideals of R, and denote by $\mathrm{Cl}(R)$ the set of equivalence classes. For a nonzero ideal I we denote by $[I] \in \mathrm{Cl}(R)$.

We note that $[I] = [R]$ if and only if $I = \alpha R = (\alpha)$, i.e. I is principal. Thus the principal ideals form an ideal class, which will turn out to be the identity element of the class group.

We define a binary operation on $\mathrm{Cl}(R)$ by $[I][J] = [IJ]$. This is clearly well defined, associative and commutative.

Proposition 12.2. *The set* $\mathrm{Cl}(R)$ *together with this operation has the structure of an abelian group.*

Proof. The class of principal ideals $[R]$ is an identity: $[R][I] = [RI] = [I]$. We need to check the existence of inverses.

Given an ideal I and $0 \neq a \in I$, by Proposition 11.9 from last lecture we can find an ideal J with $IJ = (a)$. Thus $[I][J] = [IJ] = [(a)] = [R]$. □

Example 12.3. We return to our example in $\mathbf{Z}[\sqrt{-5}]$. We found ideals $\mathfrak{p}_2, \mathfrak{p}_3, \overline{\mathfrak{p}}_3$ which satisfy formulas $\mathfrak{p}_2^2 = (2)$, $\mathfrak{p}_3\overline{\mathfrak{p}}_3 = (3)$, $\mathfrak{p}_2\mathfrak{p}_3 = (1+\sqrt{-5})$, $\mathfrak{p}_2\overline{\mathfrak{p}}_3 = (1-\sqrt{-5})$. These formulas give us relations in the class group:

$$[\mathfrak{p}_2]^2 = [\mathfrak{p}_3][\overline{\mathfrak{p}}_3] = [\mathfrak{p}_2][\mathfrak{p}_3] = [\mathfrak{p}_2][\overline{\mathfrak{p}}_3] = 1$$

and hence $[\mathfrak{p}_2] = [\mathfrak{p}_3] = [\overline{\mathfrak{p}}_3]$ and moreover this class has order 2. You might guess that $\mathrm{Cl}(\mathbf{Z}[\sqrt{-5}]) = \{1, [\mathfrak{p}_2]\}$, and is hence isomorphic to $\mathbf{Z}/2\mathbf{Z}$. This turns out to be the case, but for the moment we have no way to show that there aren't any more ideal classes.

Proposition 12.4. *Let R be a Dedekind domain. Then the following are equivalent:*

1. $\mathrm{Cl}(R) = \{1\}$, *i.e. the class group is trivial.*

2. *R is a PID.*

3. *R is a UFD.*

Proof. The equivalence of 1 and 2 follows from the fact that $[I] = [R]$ if and only if I is principal. Thus $[R]$ is the only ideal class, if and only if every ideal is principal.

We know 2 implies 3 in general from algebra. Now we prove 3 implies 2. We first show that if $\mathfrak{p} \subseteq R$ is a nonzero prime ideal then $\mathfrak{p}$ is principal. So let $0 \neq a \in \mathfrak{p}$. As R is a UFD, we can write

$$a = u\pi_1 \cdots \pi_r$$

where $u \in R^\times$ and π_i are irreducible, and hence prime by Proposition 8.5. Now since $\mathfrak{p}$ is prime and $u\pi_1 \cdots \pi_r \in \mathfrak{p}$ and $u \notin \mathfrak{p}$ we have $\pi_i \in R$ for some i. Then

$$(\pi_i) \subseteq \mathfrak{p}$$

are nonzero prime ideals, hence both maximal and hence equal. Thus $\mathfrak{p} = (\pi_i)$ is principal.

Now for a general nonzero ideal $I \subseteq R$, using Theorem 11.1 write $I = \mathfrak{p}_1 \cdots \mathfrak{p}_r$ with $\mathfrak{p}_i$ prime. We've just proved each $\mathfrak{p}_i$ is principal and hence their product I is also principal. □

Remark 12.5. In particular this shows that if K is a number field, then $\mathcal{O}_K$ is a PID if and only if it is a UFD. Moreover if we are able to compute the class group $\mathrm{Cl}(\mathcal{O}_K)$, we will know if this holds!

When we work with prime factorizations of integers, it is usually helpful to group equal prime factors into prime powers, i.e. write $n = \pm p_1^{a_1} \cdots p_r^{a_r}$ where p_i are distinct primes. We can do the same thing with a nonzero ideal I of a Dedekind domain R: write $I = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$ where $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are distinct nonzero primes of R. Sometimes we insist that the $a_i > 0$, so that all the $\mathfrak{p}_i$ are really factors of R, but other times it is convenient to allow $a_i = 0$, for instance when you want to express two or more ideals in terms of the same prime ideals, as in the following proposition.

Proposition 12.6. *Let R be a Dedekind domain and let $I, J \subseteq R$ be nonzero ideals. Suppose $I = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$ and $J = \mathfrak{p}_1^{b_1} \cdots \mathfrak{p}_r^{b_r}$ where $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are distinct nonzero prime ideals of R and $a_i, b_i \geq 0$. Then:*

1. *$I \subseteq J$ if and only if $a_i \geq b_i$.*

2. *$IJ = \mathfrak{p}_1^{a_1+b_1} \cdots \mathfrak{p}_r^{a_r+b_r}$.*

3. *$I + J = \mathfrak{p}_1^{\min(a_1,b_1)} \cdots \mathfrak{p}_r^{\min(a_r,b_r)}$*

4. *$I \cap J = \mathfrak{p}_1^{\max(a_1,b_1)} \cdots \mathfrak{p}_r^{\max(a_r,b_r)}$*

You should compare these formulas with the case of $R = \mathbf{Z}$ (or more generally R a PID) which I reviewed in lecture 9. Note that these formulas for $I + J$ and $I \cap J$ exactly match with how you compute the gcd or lcm of integers from prime power factorizations. It is a good exercise to try to prove these yourself to build dexterity with using unique factorization into prime ideals! As you get good at working with factorizations into prime ideals, you will start to apply these results, especially point 1, without even thinking!

Proof. Point 2 is clear.

We prove 1. By Corollary 11.11 we have that $I \subseteq J$ if and only if there exists an ideal $K \subseteq R$ with $I = JK$. We also consider the prime factorization $K = \mathfrak{p}_1^{c_1} \cdots \mathfrak{p}_r^{c_r}$ (possibly expanding our list $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ of primes to include any new factors of K, but in fact there won't be any!) Then by 2 the equation $I = JK$ reads $a_i = b_i + c_i$. Hence if K_i exists, $a_i \geq b_i$, and conversely if $a_i \geq b_i$, $K = \mathfrak{p}_1^{a_i - b_i} \cdots \mathfrak{p}_r^{a_i - b_i}$ satisfies $I = JK$.

For the proofs of 3 and 4, we ponder the implication of 1 on "$I + J$ is the smallest ideal containing I and J" and "$I \cap J$ is the largest ideal containing $I \cap J$" on the prime factorizations of $I + J$ and $I \cap J$ and the proofs are almost automatic:

We prove 3. Let $I + J = \mathfrak{p}_1^{c_1} \cdots \mathfrak{p}_r^{c_r}$. Then $I, J \subseteq I + J$ give $c_i \leq a_i, b_i$ by 1, and hence $c_i \leq \min(a_i, b_i)$. On the other hand 1 also implies

$$I, J \subseteq \mathfrak{p}_1^{\min(a_1,b_1)} \cdots \mathfrak{p}_r^{\min(a_r,b_r)}$$

since $\min(a_i, b_i) \leq a_i, b_i$, and hence

$$I + J \subseteq \mathfrak{p}_1^{\min(a_1,b_1)} \cdots \mathfrak{p}_r^{\min(a_r,b_r)}$$

and hence $c_i \geq \min(a_i, b_i)$. Thus $c_i = \min(a_i, b_i)$.

We prove 4. Let $I \cap J = \mathfrak{p}_1^{c_1} \cdots \mathfrak{p}_r^{c_r}$. Then $I \cap J \subseteq I, J$ gives $a_i, b_i \leq c_i$ by 1, and hence $\max(a_i, b_i) \leq c_i$. On the other hand 1 also implies

$$\mathfrak{p}_1^{\max(a_1,b_1)} \cdots \mathfrak{p}_r^{\max(a_r,b_r)} \subseteq I, J$$

and hence

$$\mathfrak{p}_1^{\max(a_1,b_1)} \cdots \mathfrak{p}_r^{\max(a_r,b_r)} \subseteq I \cap J$$

and so 1 again implies $\max(a_i, b_i) \geq c_i$. Thus $c_i = \max(a_i, b_i)$. □

Example 12.7. We return to our ideals $\mathfrak{p}_2, \mathfrak{p}_3, \overline{\mathfrak{p}_3}$ and check some of these formulas. For instance

$$(3, 1 + \sqrt{-5}) = (3) + (1 + \sqrt{-5}) = \mathfrak{p}_3\overline{\mathfrak{p}}_3 + \mathfrak{p}_2\mathfrak{p}_3 = \mathfrak{p}_3$$

where in the last step we are using part 3 of the proposition. Of course now we have gone in circle because we originally defined $\mathfrak{p}_3 = (3, q + \sqrt{-5})$.

Similarly,

$$\mathfrak{p}_3\overline{\mathfrak{p}}_3 \cap \mathfrak{p}_2\mathfrak{p}_3 = \mathfrak{p}_2\mathfrak{p}_3\overline{\mathfrak{p}}_3$$

by part 4 of the proposition, and so we see that $(3) \cap (1 + \sqrt{-5}) = \mathfrak{p}_2\mathfrak{p}_3\overline{\mathfrak{p}}_3$ is not a principal ideal. (Why is it not a principal ideal? Can you find generators for it?)

Exercise 12.8. Try to compute $(a) + (b)$ and $(a) \cap (b)$ for various $a, b \in \mathcal{O}_{-5}$ among $2, 3, 1 + \sqrt{-5}, 1 - \sqrt{-5}$ or otherwise.

12.1 The Chinese remainder theorem

We now review a bit of general commutative ring theory which we could have discussed before, but makes the most sense to consider now. You have probably seen this before multiple times, but it is really important!

Proposition 12.9 (CRT)**.** *Let R be a commutative ring. Let $I, J \subset R$ be ideals. Then if $I + J = R$*[40] *we have $I \cap J = IJ$ and there is an isomorphism of rings*

$$R/IJ \to R/I \times R/J$$
$$a + IJ \mapsto (a + I, a + J)$$

Proof. To begin with, without any hypothesis on I and J there is always a homomorphism

$$\phi : R \to R/I \times R/J$$
$$a \mapsto (a + I, a + J)$$

Moreover $\phi(a) = 0$ if and only if $a + I = 0 + I$ and $a + J = 0 + J$ or equivalently if $a \in I$ and $a \in J$, or again equivalently if $a \in I \cap J$. Hence $\ker(\phi) = I \cap J$.

Now we consider the hypothesis $I + J = R$ and prove that ϕ is surjective. Since $1 \in R$, $I + J = R$ implies we can find $x \in I$ and $y \in J$ with $x + y = 1$. Hence

$$\phi(x) = (0 + I, 1 + J), \quad \phi(y) = (1 + I, 0 + J)$$

and hence for $a, b \in R$,

$$\phi(ay + bx) = (a + I, b + J)$$

and so ϕ is surjective.

Finally we show $I \cap J = IJ$. Indeed we have

$$IJ \subseteq I \cap J = (I \cap J)(I + J) \subseteq (I \cap J)I + (I \cap J)J \subseteq IJ.$$

Now we've shown that ϕ is surjective and its kernel is $IJ = I \cap J$, so the first isomorphism theorem gives the isomorphism in the statement of the theorem. □

Exercise 12.10. Prove $I + J = R$ implies $IJ = I \cap J$ in the special case that R is a Dedekind domain, using Proposition 12.6.

Remark 12.11. The most important part of the Chinese remainder theorem is the statement that if $I + J = R$ then the homomorphism

$$\phi : R \to R/I \times R/J$$

is surjective. We can view this as a statement about solving systems of congruences: given $a, b \in R$ we can always find a $c \in R$ with

$$c \equiv a \pmod{I}$$
$$c \equiv b \pmod{J}$$

Here when I write $c \equiv a \pmod{I}$ I really just mean $c + I = a + I$, or $c - a \in I$. Moreover the proof of the Chinese remainder theorem just gives you a formula for c as long as you can find $x \in I$, $y \in J$ with $x + y = 1$. Indeed $c = ay + bx$ then works.

Thus the proof of the CRT is constructive as long as you can find $x + y = 1$. If R is a Euclidean domain then you could use the Euclidean algorithm.

Using induction one can prove a more general statement:

[40] In commutative ring theory we say I and J are comaximal, but in the case of PIDs or Dedekind domains where we view $I + J$ as being the "gcd" of I and J, we might also say I and J are coprime.

Corollary 12.12 (More general CRT)**.** *Let R be a commutative ring and let $I_1, \dots, I_r \subseteq R$ be ideals. Suppose that $I_i + I_j = R$ for all $1 \leq i, j \leq r$, $i \neq j$. Then $I_1 \cap I_2 \cap \cdots \cap I_r = I_1 I_2 \cdots I_r$ and there is an isomorphism of rings*

$$R/I_1 \cdots I_r \to R/I_1 \times R/I_2 \times \cdots \times R/I_r$$
$$a + I_1 \cdots I_r \mapsto (a + I_1, a + I_2, \dots, a + I_r)$$

Proof. This is proved by induction on r. We claim that $I_1 + I_i = R$ for $i = 2, \dots, r$ implies that $I_1 + I_2 I_3 \cdots I_r = R$. In fact we have

$$R = (I_1 + I_2)(I_1 + I_3) \cdots (I_1 + I_r) \subseteq I_1 + I_2 \cdots I_r.$$

Now we have isomorphisms:

$$R/I_1 \cdots I_r \to R/I_1 \times R/I_2 \cdots I_r \to R/I_1 \times R/I_2 \times \cdots \times R/I_r$$

where the first is proposition 12.9 applied to the ideals $I = I_1$ and $J = I_2 \cdots I_r$ and the second is by induction. □

Returning to the situation of a Dedekind domain R, if we have an ideal $\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$ where $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ are distinct nonzero prime ideals of R, then we have $\mathfrak{p}_i^{a_i} + \mathfrak{p}_j^{a_j} = R$ for $i \neq j$ by Proposition 12.6. Hence by the Chinese remainder theorem we have an isomorphism

$$R/\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r} \to R/\mathfrak{p}_1^{a_1} \times R/\mathfrak{p}_2^{a_2} \times \cdots \times R/\mathfrak{p}_r^{a_r}$$

As we remarked above, we are often most interested in the fact that the homomorphism

$$R \to R/\mathfrak{p}_1^{a_1} \times R/\mathfrak{p}_2^{a_2} \times \cdots \times R/\mathfrak{p}_r^{a_r}$$

is surjective, and we like to think about this as the statement that a system of congruences on $x \in R$

$$\begin{aligned} x &\equiv x_1 \pmod{\mathfrak{p}_1^{a_1}} \\ x &\equiv x_2 \pmod{\mathfrak{p}_2^{a_2}} \\ &\cdots \\ x &\equiv x_r \pmod{\mathfrak{p}_r^{a_r}} \end{aligned}$$

is solvable for all $x_1, \dots, x_r \in R$.

As an illustration of using factorization into prime ideals and the CRT, we prove the following result, which is interesting but ultimately not that important.

Proposition 12.13. *Let R be a Dedekind domain and let $I \subseteq R$ be a nonzero ideal, and let $0 \neq a \in I$. Then there exists $b \in I$ with $I = (a, b)$. In particular every ideal in R can be generated by two elements.*

Proof. Let

$$I = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$$

and

$$(a) = \mathfrak{p}_1^{b_1} \cdots \mathfrak{p}_r^{b_r} \mathfrak{q}_1^{c_1} \cdots \mathfrak{q}_s^{c_s}$$

where $\mathfrak{p}_1, \dots, \mathfrak{p}_r, \mathfrak{q}_1, \dots, \mathfrak{q}_s$ are distinct nonzero prime ideals of R and $b_i \geq a_i$ for $i = 1, \dots, r$ as $(a) \subseteq I$.

We will explain below how to use the CRT to construct $b \in R$ with the following properties:

- $b \in \mathfrak{p}_i^{a_i}$ and $b \notin \mathfrak{p}_i^{a_i+1}$ for $i = 1, \ldots, r$
- $b \notin \mathfrak{q}_i$ for $i = 1, \ldots, s$.

Assuming we have such a b for the moment, we now consider the factorization of (b) into prime ideals. We claim we have

$$(b) = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r} \mathfrak{r}_1^{d_1} \cdots \mathfrak{r}_t^{d_t}$$

where $\mathfrak{r}_1, \ldots, \mathfrak{r}_t$ are some new primes pairwise distinct as well as distinct from the $\mathfrak{p}_i$ and $\mathfrak{q}_i$. Indeed the point is that $\mathfrak{p}_i$ occurs a_i times in the factorization as a result of $b \in \mathfrak{p}_i^{a_i}$ but $b \notin \mathfrak{p}_i^{a_i+1}$ (interpret $(b) \subseteq \mathfrak{p}_i^{a_i}$, $(b) \subsetneq \mathfrak{p}_i^{a_i+1}$ via Proposition 12.6 1). Similarly $\mathfrak{q}_i$ does not occur because $b \notin \mathfrak{q}_i$.

Now we compute

$$\begin{aligned}(a, b) &= (a) + (b) \\ &= \mathfrak{p}_1^{b_1} \cdots \mathfrak{p}_r^{b_r} \mathfrak{q}_1^{c_1} \cdots \mathfrak{q}_s^{c_s} + \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_1^{a_r} \mathfrak{r}_1^{d_1} \cdots \mathfrak{r}_t^{d_t} \\ &= \mathfrak{p}_1^{\min(a_1,b_1)} \cdots \mathfrak{p}_r^{\min(a_r,b_r)} \\ &= \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r} = I.\end{aligned}$$

Finally we explain how to construct b.

- For $i = 1, \ldots, r$ we have $\mathfrak{p}_i^{a_i+1} \subsetneq \mathfrak{p}_i^{a_i}$ and so we can pick $x_i \in \mathfrak{p}_i^{a_i}$, $x_i \notin \mathfrak{p}_i^{a_i+1}$.
- For $i = 1, \ldots, s$ we have $\mathfrak{q}_i \subsetneq R$ so we can pick $y_i \in R$, $y_i \notin \mathfrak{q}_i$.

Now we apply the Chinese remainder theorem to $\mathfrak{p}_1^{a_1+1}, \ldots, \mathfrak{p}_r^{a_r+1}, \mathfrak{q}_1, \ldots, \mathfrak{q}_s$ to produce $b \in R$ with:

- $b \equiv x_i \pmod{\mathfrak{p}_i^{a_i+1}}$ for $i = 1, \ldots, r$ (which implies $b \in \mathfrak{p}_i^{a_i}$, $b \notin \mathfrak{p}_i^{a_i+1}$)
- $b \equiv y_i \pmod{\mathfrak{q}_i}$ for $i = 1, \ldots, s$ (which implies $b \notin \mathfrak{q}_i$)

so we are done! □

You might find the proof of Proposition 12.13 to be totally crazy the first time you see it, but once you internalize proposition 12.6, it should seem much more reasonable!

Remark 12.14. This is just for fun: we tend to think that $\mathbf{C}[x, y]$ is a really nice ring! It is Noetherian so every ideal is finitely generated, however there is no bound on the number of generators required:

$$(x^n, x^{n-1}y, \ldots, xy^{n-1}, y^n) \subset \mathbf{C}[x, y]$$

cannot be generated by fewer than $n+1$ elements. Knowing this might make Proposition 12.13 a bit more surprising!

13 Lecture 13: Example Class 2

1. (a) In lecture 6 we saw three formulas for the discriminant: $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = \det(\mathrm{tr}_K(\alpha_i\alpha_j)) = \det(\tau_i(\alpha_j))^2$, and $\mathrm{disc}(1, \alpha, \ldots, \alpha^{n-1}) = (-1)^{\frac{n(n-1)}{2}} N_{\mathbf{Q}(\alpha)}(m'_\alpha(\alpha))$. Compute $\mathrm{disc}(1, \sqrt[3]{2}, \sqrt[3]{4})$ directly using all three.

   (b) Suppose that $f = x^3 + ax + b \in \mathbf{Q}[x]$ is irreducible and let $\alpha \in \mathbf{C}$ be a root, and consider $K = \mathbf{Q}(\alpha)$. Show that $\mathrm{disc}(1, \alpha, \alpha^2) = -27b^2 - 4a^3$ (you might do this as in the computation in Example 6.16 in the notes.). Do the same for $f = x^n + ax + b$.

   (c) Let $f = x^3 + x - 3$. Prove that f is irreducible. Then, letting $K = \mathbf{Q}(\alpha)$ where $\alpha \in \mathbf{C}$ a root of f, compute $\mathcal{O}_K$ and D_K.

2. Explain directly from the definition why the rings $\mathbf{Z}[2i]$ and $\mathbf{Z}[\sqrt{-3}]$ aren't integrally closed. For each ring, give an explicit example of nonunique factorizations into irreducibles.

3. Consider the ideals in $\mathcal{O}_{-5} = \mathbf{Z}[\sqrt{-5}]$ from Example 9.5 in Lecture 9:

$$\mathfrak{p}_2 = (2, 1 + \sqrt{-5}) = (2, 1 - \sqrt{-5}) = (1 + \sqrt{-5}, 1 - \sqrt{-5})$$
$$\mathfrak{p}_3 = (3, 1 + \sqrt{-5})$$
$$\overline{\mathfrak{p}}_3 = (3, 1 - \sqrt{-5})$$

   (a) Check that the three ideals asserted to be equal in the definition of $\mathfrak{p}_2$ are actually equal.

   (b) Check the formulas $(2) = \mathfrak{p}_2^2$, $(3) = \mathfrak{p}_3\overline{\mathfrak{p}}_3$.

   (c) Show that $\mathfrak{p}_3^2$ is principal and find a generator.

   (d) What is $\mathcal{O}_{-5}/\mathfrak{p}_2$?

   (e) Prove that $(7) \subset \mathcal{O}_{-5}$ is *not* a prime ideal.

4. This problem is about Euclidean domains.

   (a) Show that $\mathcal{O}_2$ and $\mathcal{O}_3$ are norm Euclidean. If you are feeling braver you might try to show that $\mathcal{O}_6$ and $\mathcal{O}_7$ are norm Euclidean. The remaining $d > 0$, $d \not\equiv 1 \pmod 4$ for which $\mathcal{O}_d$ is norm Euclidean are $d = 11, 19$. Treating these without a very good strategy and/or using a computer is quite difficult!

   (b) Show that the ring $\mathcal{O}_{14} = \mathbf{Z}[\sqrt{14}]$ is not norm Euclidean by showing that there are no $q, r \in \mathcal{O}_{14}$ with $1 + \sqrt{14} = 2q + r$ and $|N_{\mathbf{Q}(\sqrt{14})}(r)| < |N_{\mathbf{Q}(\sqrt{14})}(2)| = 4$. (Hint: show that this is equivalent to finding odd numbers $a, b \in \mathbf{Z}$ with $|a^2 - 14b^2| < 4$, and try to show that this is impossible using congruences mod 8 and 7.)

   (c) Let $d < -3$ be squarefree. Suppose that $f : \mathcal{O}_d \to \mathbf{N}$ is a Euclidean function. Let $b \in \mathcal{O}_d$ be a nonzero nonunit element such that $f(b)$ is minimal among all nonzero nonunit elements of $\mathcal{O}_d$. Show that for any $a \in R$ there exists $q \in \mathcal{O}_d$ and $r \in \{-1, 0, 1\}$ with $a = qb + r$. Deduce that the cardinality of $\mathcal{O}_d/(b)$ is 2 or 3.

   (d) We will prove soon that for $b \in \mathcal{O}_d$, $\#(\mathcal{O}_d/(b)) = |N_{\mathbf{Q}(\sqrt{d})}(b)|$. Using this and the previous part, show that $\mathcal{O}_d$ is not a Euclidean domain for $d < 0$ unless $d = -1, -2, -3, -7, -11$.

14 Lecture 14: Splitting of primes I

Let K be a number field with integer ring $\mathcal{O}_K$. We now know that any ideal of $\mathcal{O}_K$ can be factored into prime ideals. How can we describe all the prime ideals of $\mathcal{O}_K$?

Lemma 14.1. *Let K be a number field and let $\mathfrak{p} \subset \mathcal{O}_K$ be a nonzero prime ideal. Then there is a unique prime number p with $p \in \mathfrak{p}$.*

Proof. The uniqueness part is easy: if $p, q \in \mathfrak{p}$ for $p \neq q$ prime numbers, then we can write $1 = ap + bq$ for $a, b \in \mathbf{Z}$ and so $1 \in \mathfrak{p}$ and hence $\mathfrak{p} = R$ isn't prime.

We prove existence. We first claim that there is some integer $n \neq 0$ with $n \in \mathfrak{p}$. We prove this as in the proof of Lemma 10.11. Pick $0 \neq \alpha \in \mathfrak{p}$. We claim that $n = N_K(\alpha) \in \mathfrak{p}$. Indeed by Theorem 4.4 we have

$$N_K(\alpha) = \alpha \prod_{\tau \neq \tau_{id}} \tau(\alpha) = \alpha\beta$$

We recall that while we need not have $\tau(\alpha) \in K$ for each embedding $\tau : K \to \mathbf{Q}$, we do have $\beta = \prod_{\tau \neq \tau_{id}} \tau(\alpha) \in K$ because $\beta = N_K(\alpha)\alpha^{-1} \in K$. We also have that β is an algebraic integer because each $\tau(\alpha)$ is an algebraic integer. Hence $\beta \in \mathcal{O}_K$. Hence since $\alpha \in \mathfrak{p}$, $n = N_K(\alpha) = \alpha\beta \in \mathfrak{p}$.

Now we can write $n = \pm p_1 \cdots p_r$ where p_i are prime numbers. Because $\mathfrak{p}$ is prime, we conclude that $p_i \in \mathfrak{p}$ for some i. □

Remark 14.2. Here is an alternative point of view on this lemma: we know by Lemma 10.11 that $R/\mathfrak{p}$ is a finite field. It follows that it is a field of characteristic p for a unique prime p (i.e. $p = 0$ in $R/\mathfrak{p}$) and hence $p \in \mathfrak{p}$.

Definition 14.3. We say that a nonzero prime ideal $\mathfrak{p} \subset \mathcal{O}_K$ *lies over* the prime p if $p \in \mathfrak{p}$.

Proposition 14.4. *Let p be a prime number and suppose $(p) = p\mathcal{O}_K = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2} \cdots \mathfrak{p}_r^{e_r}$ with $\mathfrak{p}_1, \ldots, \mathfrak{p}_r \subseteq \mathcal{O}_K$ distinct prime ideals. Then the primes $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are exactly the primes which lie above p.*

Proof. This is an immediate consequence of "to contain is to divide". Indeed we have $p \in \mathfrak{p}_i$ for $i = 1, \ldots, r$, so each prime $\mathfrak{p}_i$ occurring in the factorization of (p) lies over p. Conversely, if $\mathfrak{p} \subseteq \mathcal{O}_K$ is a prime ideal with $p \in \mathfrak{p}$ then we have $(p) \subseteq \mathfrak{p}$ and hence $(p) = \mathfrak{p}I$ by Corollary 11.11, and hence $\mathfrak{p}$ must be $\mathfrak{p}_i$ for some i by unique factorization into prime ideals. □

Thus we see that to find all the prime ideals of $\mathcal{O}_K$, we just have to factor each of the ideals $(2), (3), (5), (7), \ldots$ into prime ideals. We call the factorization of (p) into prime in $\mathcal{O}_K$ "the splitting of p in $\mathcal{O}_K$".

Example 14.5. We determine the splitting of the first several primes in the Gaussian integers $\mathbf{Z}[i]$. We recall that $\mathbf{Z}[i]$ is a PID because it is Euclidean, and hence if $\mathfrak{p}$ is a prime ideal lying over p, we have $\mathfrak{p} = (\pi)$ is principal and $\pi | p$. Taking norms we conclude that $N(\pi) | p^2$, so $N(\pi) = p$ or $N(\pi) = p^2$.

We observe that $N(x + iy) = x^2 + y^2 \equiv 0, 1, 2 \pmod 4$. Hence if $p \equiv 3 \pmod 4$ there cannot exist $\pi \in \mathbf{Z}[i]$ with $N(\pi) = p$, and hence $(p) \subset \mathbf{Z}[i]$ must be prime. With this in mind we can list the first splitting of the first several primes:

- $(2) = (1+i)(1-i) = (1+i)^2$ since $i(1-i) = 1+i$.
- (3) is prime.
- $(5) = (2+i)(2-i)$
- (7) is prime.
- (11) is prime.
- $(13) = (3+2i)(3-2i)$

We note that if $p = x^2 + y^2$ then $p = (x+iy)(x-iy)$ and $x+iy$ and $x-iy$ have norm p so they must be prime. They might or might not generate the same ideal. This happens if and only if $x+iy = u(x-iy)$ for a unit u. Noting that $x, y \neq 0$ we cannot have $x+iy = \pm(x-iy)$ and $x+iy = \pm i(x-iy)$ implies $x = \pm y$. But when p is odd, one of x and y is even and one is odd, and so this cannot happen. Thus for all odd primes, either (p) is prime, or $(p) = \mathfrak{p}_1\mathfrak{p}_2$ splits as the product of two distinct primes. Meanwhile $(2) = (1+i)^2$ is the square of a prime.

Example 14.6. Similarly we have considered some splittings for $\mathcal{O}_{-5} = \mathbf{Z}[\sqrt{-5}]$:

- $(2) = (2, 1+\sqrt{-5})^2$
- $(3) = (3, 1+\sqrt{-5})(3, 1-\sqrt{-5})$
- $(5) = (\sqrt{-5})^2$
- $(7) = (7, 3+\sqrt{-5})(7, 3-\sqrt{-5})$
- (11) is prime.

We note that by contrast to the case of $\mathbf{Z}[i]$, $\mathbf{Z}[\sqrt{-5}]$ is not a PID. Thus p not being a norm does not imply that (p) is prime. Indeed, $2, 3, 7 = x^2 + 5y^2$ has no solutions, but we see (2), (3), (7) do split. The only thing we can conclude in this case is that the factors cannot be principal ideals.

How then do we see that (11) is prime? Well, if there existed a prime ideal $\mathfrak{p} \subsetneq (p)$ then we would have $[\mathcal{O}_{-5} : \mathfrak{p}] \mid [\mathcal{O}_{-5} : (p)] = p^2$. Hence $\#\mathcal{O}_{-5}/\mathfrak{p} = p$, so $\mathcal{O}_{-5}/\mathfrak{p} \simeq \mathbf{F}_p$. It would follow then that there exists some $a \in \mathbf{Z}$ such that $\sqrt{-5} + \mathfrak{p} = a + \mathfrak{p}$. Squaring this, we obtain $a^2 \equiv -5 \pmod{p}$. This explains how we might guess some of the primes in the factorization above, and also proves that (11) is prime: $a^2 \equiv -5 \pmod{11}$ has no solutions as you can check: $(\pm 1)^2 = 1, (\pm 2)^2 = 4, (\pm 3)^2 = 9, (\pm 4)^2 = 16 \equiv 5, (\pm 5)^2 \equiv 3$.

Definition 14.7. Let K be a number field, and let $\mathfrak{p} \subseteq \mathcal{O}_K$ be a nonzero prime ideal lying over p.

1. The *ramification index* $e(\mathfrak{p})$ of $\mathfrak{p}$ is the multiplicity with which $\mathfrak{p}$ occurs in the factorization of (p) into prime ideals in $\mathcal{O}_K$. Thus

$$(p) = \prod_{p \in \mathfrak{p}} \mathfrak{p}^{e(\mathfrak{p})}.$$

2. The quotient $\mathcal{O}_K/\mathfrak{p}$ is a finite field of characteristic p. The *inertial degree* of $\mathfrak{p}$ is

$$f(\mathfrak{p}) = [\mathcal{O}_K/\mathfrak{p} : \mathbf{F}_p].$$

Hence by definition, $\mathcal{O}_K/\mathfrak{p}$ is a $\mathbf{F}_p$-vector space of dimension $f(\mathfrak{p})$, and so $\#\mathcal{O}_K/\mathfrak{p} = [\mathcal{O}_K : \mathfrak{p}] = p^{f(\mathfrak{p})}$.[41]

Theorem 14.8. *Let K be a number field and let p be a prime. We have*

$$\sum_{p \in \mathfrak{p}} e(\mathfrak{p}) f(\mathfrak{p}) = [K : \mathbf{Q}]$$

where the sum is taken over prime ideals of $\mathcal{O}_K$ lying over p.

Example 14.9. Let us understand what this theorem says when $K = \mathbf{Q}(\sqrt{d})$, is a quadratic field. We see that there are three possible ways a prime p can behave in $\mathcal{O}_d$:

1. We have $(p) = \mathfrak{p}_1\mathfrak{p}_2$ where $\mathfrak{p}_i$ are distinct prime ideals, and $e(\mathfrak{p}_i) = f(\mathfrak{p}_i) = 1$ for $i = 1, 2$. We say p *splits* in $\mathcal{O}_d$.

2. We have $(p) = \mathfrak{p}^2$ where $\mathfrak{p}$ is prime. We have $e(\mathfrak{p}) = 2$, $f(\mathfrak{p}) = 1$. We say p *ramifies* in $\mathcal{O}_d$.

3. We have $(p) = \mathfrak{p}$ is prime. Then we have $e(\mathfrak{p}) = 1$, $f(\mathfrak{p}) = 2$. We say p is *inert* in $\mathcal{O}_d$.

The formula in the theorem reads $1 \cdot 1 + 1 \cdot 1 = 2$ in the first case, $2 \cdot 1 = 2$ in the second case, and $1 \cdot 2 = 2$ in the third case. We saw that all three possibilities occurred in $\mathbf{Z}[i], \mathbf{Z}[\sqrt{-5}]$

We introduce some more notation in preparation for proving Theorem 14.8.

Definition 14.10. Let $I \subseteq \mathcal{O}_K$ be a nonzero ideal. We write $\|I\| = [\mathcal{O}_K : I] = \#\mathcal{O}_K/I$ and call it the norm of the ideal I.[42] We saw that this is finite in Lemma 10.11.

Example 14.11. If $\mathfrak{p}$ is a prime ideal with $p \in \mathfrak{p}$ then we have $\|\mathfrak{p}\| = p^{f(\mathfrak{p})}$ by definition of the inertial degree $f(\mathfrak{p})$.

As another example, for $0 \neq n \in \mathbf{Z}$ we have $\|(n)\| = |n|^{[K:\mathbf{Q}]}$. Indeed picking an integral basis we see that the index of $n\mathcal{O}_K \subseteq \mathcal{O}_K$ is the same as the index of $n\mathbf{Z}^{[K:\mathbf{Q}]} \subseteq \mathbf{Z}^{[K:\mathbf{Q}]}$ (we made this computation in the proof of Lemma 10.11).

We will prove later that in fact for all $0 \neq \alpha \in \mathcal{O}_K$, $\|(\alpha)\| = |N_K(\alpha)|$.

Proposition 14.12. *If I and J are nonzero ideals then we have*

$$\|IJ\| = \|I\| \cdot \|J\|.$$

[41] Warning: there is a terrible potential notational conflict here. When $K \subset L$ are fields we write $[L : K]$ for the *degree*, i.e. the dimension of L as a K vector space. When $A \subset B$ are groups we write $[B : A]$ for the index, i.e. the number of cosets of A in B. In particular when we write $[\mathcal{O}_K/\mathfrak{p} : \mathbf{F}_p]$ we mean the degree of a field extension, not the index of abelian groups!

[42] A more common notation is just $N(I)$. We won't use this to avoid confusion with the norm of the element $N_K(\alpha)$. Nonetheless we will soon see that there is a close relation.

Proof. It suffices to prove that for distinct nonzero primes $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$,

$$\|\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}\| = \|\mathfrak{p}_1\|^{a_1} \cdots \|\mathfrak{p}_r\|^{a_r}.$$

By the Chinese remainder theorem we have

$$\mathcal{O}_K/\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r} \simeq \mathcal{O}_K/\mathfrak{p}_1^{a_1} \times \cdots \times \mathcal{O}_K/\mathfrak{p}_r^{a_r}.$$

Taking cardinalities gives

$$\|\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}\| = \|\mathfrak{p}_1^{a_1}\| \cdots \|\mathfrak{p}_r^{a_r}\|$$

thus it remains to prove that for $\mathfrak{p} \subset \mathcal{O}_K$ a nonzero prime, $\|\mathfrak{p}^k\| = \|\mathfrak{p}\|^k$.

We will prove this by induction on k. We observe that we have inclusions

$$\mathfrak{p}^{k+1} \subseteq \mathfrak{p}^k \subseteq \mathcal{O}_K$$

and by Lagrange's theorem[43]

$$\|\mathfrak{p}^{k+1}\| = [\mathcal{O}_K : \mathfrak{p}^{k+1}] = [\mathcal{O}_K : \mathfrak{p}^k][\mathfrak{p}^k : \mathfrak{p}^{k+1}] = \|\mathfrak{p}^k\| \cdot [\mathfrak{p}^k : \mathfrak{p}^{k+1}]$$

so by induction it suffices to prove $[\mathfrak{p}^k : \mathfrak{p}^{k+1}] = \|\mathfrak{p}\|$, i.e. $\#\mathfrak{p}^k/\mathfrak{p}^{k+1} = \#\mathcal{O}_K/\mathfrak{p}$.

We have $\mathfrak{p}^{k+1} \subsetneq \mathfrak{p}^k$ by Corollary 11.10, so we can pick $b \in \mathfrak{p}^k$, $b \notin \mathfrak{p}^{k+1}$. We will prove that the map

$$\begin{aligned} \mathcal{O}_K/\mathfrak{p} &\to \mathfrak{p}^k/\mathfrak{p}^{k+1} \\ a + \mathfrak{p} &\mapsto ab + \mathfrak{p}^{k+1} \end{aligned}$$

is a well defined isomorphism of abelian groups.

- Well defined: if $a' = a + x$ with $x \in \mathfrak{p}$, then $ba' = ba + bx$ and $bx \in \mathfrak{p}^{k+1}$ so $ba' + \mathfrak{p}^{k+1} = ba + \mathfrak{p}^{k+1}$.
- Injective: if $a \in \mathcal{O}_K$, $a \notin \mathfrak{p}$ we prove that $ab \notin \mathfrak{p}^{k+1}$. Indeed since $\mathfrak{p} \subset \mathcal{O}_K$ is maximal, $(a) + \mathfrak{p} = \mathcal{O}_K$, i.e. we can write $1 = ax + y$ for $x \in \mathcal{O}_K$, $y \in \mathfrak{p}$. Then if $ab \in \mathfrak{p}^{k+1}$ we have $b = (ax + y) \in \mathfrak{p}^{k+1}$, contradicting how we chose b.
- Surjective: we first note that $b \in \mathfrak{p}^k$, $b \notin \mathfrak{p}^{k+1}$ implies that we have $(b) = \mathfrak{p}^k I$ where I is coprime to $\mathfrak{p}$. Thus $(b) + \mathfrak{p}^{k+1} = \mathfrak{p}^k$. It follows that for $x \in \mathfrak{p}^k$ we have $x = ab + y$ for $a \in \mathcal{O}_K$ and $y \in \mathfrak{p}^{k+1}$. Then $a + \mathfrak{p} \mapsto x + \mathfrak{p}^{k+1}$ and so the map is surjective.

□

Exercise 14.13. Show that if $\mathfrak{p} = (2, x) \subseteq \mathbf{Z}[x]$ then $[\mathbf{Z}[x] : \mathfrak{p}] = 2$ while $[\mathfrak{p} : \mathfrak{p}^2] = 4$. Where do things go wrong if you try to produce an isomorphism $\mathbf{Z}[x]/\mathfrak{p} \to \mathfrak{p}/\mathfrak{p}^2$ as in the above proof?

[43]Here I am calling Lagrange's theorem the statement that if $A \subset B \subset C$ are abelian groups then $[C : A] = [C : B][B : A]$. You can prove this directly by showing that if $B = \coprod_i b_i + A$ and $C = \coprod_j c_j + B$ then $C = \coprod_{i,j} b_i + c_j + A$. Alternatively you can prove it like this: $B/A \subseteq C/A$ is a subgroup and so $\#C/A = \#B/A \cdot \#((C/A)/(B/A))$ and $(C/A)/(B/A) \simeq C/B$ by the third isomorphism theorem.

15 Lecture 15: Splitting of primes II

We give the proof of Theorem 14.8 stated last time.

Proof of Theorem 14.8. Let $(p) = \mathfrak{p}_1^{e(\mathfrak{p}_1)} \cdots \mathfrak{p}_r^{e(\mathfrak{p}_r)}$ be the factorization of p. We compute:

$$\begin{aligned} p^{[K:\mathbf{Q}]} &= \|(p)\| \\ &= \|\mathfrak{p}_1^{e(\mathfrak{p}_1)} \cdots \mathfrak{p}_r^{e(\mathfrak{p}_r)}\| \\ &= \|\mathfrak{p}_1\|^{e(\mathfrak{p}_1)} \cdots \|\mathfrak{p}_r\|^{e(\mathfrak{p}_r)} \\ &= p^{e(\mathfrak{p}_1)f(\mathfrak{p}_1)+\cdots+e(\mathfrak{p}_r)f(\mathfrak{p}_r)} \end{aligned}$$

and hence

$$[K : \mathbf{Q}] = \sum_{i=1}^{r} e(\mathfrak{p}_i) f(\mathfrak{p}_i).$$

□

To finish our discussion of the ideal norm, we clarify the relation between the ideal norm and the "element" norm.

Proposition 15.1. *Let K be a number field and let $0 \neq \alpha \in \mathcal{O}_K$. Then we have*

$$\|(\alpha)\| = |N_K(\alpha)|.$$

Proof. We first explain a quick proof in the case that $K = \mathbf{Q}(\sqrt{d})$ is quadratic. Then conjugation defines an automorphism of $\mathcal{O}_K$, and so $\|(\alpha)\| = \|(\overline{\alpha})\|$. Then we compute

$$\|(\alpha)\|^2 = \|(\alpha)\| \cdot \|(\overline{\alpha})\| = \|(\alpha\overline{\alpha})\| = \|(N_K(\alpha))\| = |N_K(\alpha)|^2.$$

where in the middle equality we have used Proposition 14.12 and the last equality holds because $N_K(\alpha) \in \mathbf{Z}$.

You might hope we could argue in the same way for a general number field K, using the formula $N_K(\alpha) = \prod_{\tau \in \Sigma_K} \tau(\alpha)$. The problem is that the embeddings $\tau : K \to \mathbf{C}$ don't in general define automorphisms of K.[44]

Instead we will use the following lemma, whose proof uses the structure theory of finitely generated abelian groups. We take an integral basis of $\mathcal{O}_K$ and let A be the matrix of multiplication by α with respect to this basis. Then $\det(A) = N_K(\alpha)$ by the definition of the norm, and so $[\mathcal{O}_K : \alpha\mathcal{O}_K] = |N_K(\alpha)|$. □

Lemma 15.2. *Let $A \in M_n(\mathbf{Z})$ be a matrix with $\det(A) \neq 0$. Then $[\mathbf{Z}^n : A\mathbf{Z}^n] = |\det(A)|$.*[45]

Proof. By the structure theorem for finitely generated abelian groups applied to $A\mathbf{Z}^n \subseteq \mathbf{Z}^n$ we have that there is a basis $m_1, \ldots m_n$ for $\mathbf{Z}^n$ and integers $d_1, \ldots, d_n$ so that $d_1 m_1, \ldots, d_n m_n$

[44] In fact, this argument will work exactly when $K/\mathbf{Q}$ is a Galois extension. There is a way to reduce to this case, but instead we give a totally different argument.

[45] The proof of this lemma is similar to the proof of Proposition 6.8. Later when we talk about lattices and the geometry of numbers, I will try to explain a different argument that avoids using the structure theory of finitely generated abelian groups.

is a basis for $A\mathbf{Z}^n$. Then $[\mathbf{Z}^n : A\mathbf{Z}^n] = d_1 \cdots d_n$. On the other hand this also implies that there are matrices $S, S' \in M_n(\mathbf{Z})^\times$ such that

$$A = S \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} S'.$$

Then $\det(S), \det(S') \in \{\pm 1\}$. Taking determinants we obtain $|\det(A)| = d_1 \cdots d_n$. □

Remark 15.3.

- If $I \subseteq \mathcal{O}_K$ is an ideal and $\|I\| = p$ is prime, then I must be prime. Indeed I must be maximal because if $I \subseteq J \subseteq \mathcal{O}_K$ we have $[\mathcal{O}_K : J][J : I] = p$ so either $I = J$ or $J = \mathcal{O}_K$. Alternatively we can consider the factorization of I into prime ideals and consider the norm.
- Now if $\alpha \in \mathcal{O}_K$ and $N_K(\alpha) = \pm p$, then we observed before that this implies π is irreducible. But actually α must be prime as $\|(\alpha)\| = |N_K(\alpha)| = p$ by the proposition.

Theorem 15.4. *Let $d \neq 0, 1$ be a squarefree integer and let p be an odd prime. Then p splits in $\mathcal{O}_d$ as follows:*

- *If $p|d$ then p ramifies in $\mathcal{O}_d$:*
$$(p) = (p, \sqrt{d})^2.$$
- *If $p \nmid d$ and there there exists $a \in \mathbf{Z}$ with $a^2 \equiv d \pmod{p}$ then p splits in $\mathcal{O}_d$:*
$$(p) = (p, a + \sqrt{d})(p, a - \sqrt{d}).$$
- *If $p \nmid d$ and there does not exist $a \in \mathbf{Z}$ with $a^2 \equiv d \pmod{p}$ then p is inert $\mathcal{O}_d$ (i.e. (p) is prime.)*

Proof. We first check all the claimed factorizations:

- If $p \mid d$ then
$$(p, \sqrt{d})^2 = (p^2, p\sqrt{d}, d) \subseteq (p)$$
and the ideal on the left contains $\gcd(p^2, d) = p$. (d square free implies that $p^2 \nmid d$.)
- If $p \nmid d$ and $a^2 \equiv d \pmod{p}$ then
$$(p, a + \sqrt{d})(p, a - \sqrt{d}) = (p^2, p(a + \sqrt{-d}), p(a - \sqrt{-d}), a^2 - d) \subseteq (p)$$
where $a^2 - d$ is a multiple of p because $a^2 \equiv d \pmod{p}$. Moreover p is in the ideal on the left as $2pa = p(a + \sqrt{-d}) + p(a - \sqrt{-d})$ is, and hence so is $\gcd(2pa, p^2) = p$. (Note that $p \nmid a$ and $p^2 \equiv d \not\equiv 0 \pmod{p}$.)

Now we check these are really prime factorizations:

- When $p|d$ we have $p^2 = \|(p)\| = \|(p, \sqrt{d})\|^2$ and hence $\|(p, \sqrt{d})\| = p$, so $(p, \sqrt{d})$ is a prime ideal.

- When $p \nmid d$ and $a^2 \equiv d \pmod{p}$ then we have $\overline{(p, a+\sqrt{d})} = (p, a-\sqrt{d})$ so these ideals have the same norm. Moreover the product of their norms is p^2, so they have norm p and hence are prime. We should check they are really distinct (this is what we mean when we say p splits rather than ramifies!).

  To see they are distinct, note that if $(p, a+\sqrt{d}) = (p, a-\sqrt{d})$ then they contain $a+\sqrt{d}+a-\sqrt{d} = 2a$, and hence they contain $\gcd(2a, p) = 1$, which contradicts that they have norm p.

- When $p \nmid d$ and $a^2 \equiv d \pmod{p}$ has no solutions, we need to check (p) is prime. We use the same strategy as we used to show $(11) \subseteq \mathcal{O}_{-5}$ is prime in Example 14.6.

  Indeed if (p) is not prime, there exists $(p) \subsetneq \mathfrak{p}$ prime. By considering norms we must have $\|\mathfrak{p}\| = p$. But then we have $\mathcal{O}_d/\mathfrak{p} \simeq \mathbf{F}_p$. Moreover we have $(\sqrt{d}+\mathfrak{p})^2 = d+\mathfrak{p}$ i.e. there exists an element $x \in \mathbf{F}_p$ with $x^2 = d$, or in other words there exists $a \in \mathbf{Z}$ with $a^2 \equiv d \pmod{p}$.

□

16 Lecture 16: Splitting of primes III

We begin with introducing some notation from elementary number theory.

Definition 16.1. Let p be an odd prime number. We say that $m \in \mathbf{Z}$ with $p \nmid m$ is a *quadratic residue*[46] if there exists an integer $a \in \mathbf{Z}$ solving the congruence $a^2 \equiv m \pmod{p}$. If $p \nmid m$ and this congruence does not have a solution then we say a is a quadratic non-residue. Note that if $p \mid m$ then $0^2 \equiv m \pmod{p}$, but we don't call it either a quadratic residue or nonresidue.

There is a shorthand notation called, the Legendre symbol, which will be useful when we discuss quadratic reciprocity later on:

$$\left(\frac{m}{p}\right) = \begin{cases} 1 & \text{if } m \text{ is a quadratic residue mod } p \\ -1 & \text{if } m \text{ is a quadratic nonresidue mod } p \\ 0 & \text{if } p \mid m \end{cases}$$

Exercise 16.2. If you haven't taken elementary number theory, try to show that

$$\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right).$$

This formula is the main reason for introducing the Legendre symbol. One case of this is harder than the rest: if m and n are quadratic nonresidues then mn is a quadratic residue. For this you will need to use that $(\mathbf{Z}/p\mathbf{Z})^\times \simeq \mathbf{Z}/(p-1)\mathbf{Z}$ is cyclic.

Using this shorthand we can restate Theorem 15.4 from last lecture:

Theorem 16.3. *Let $d \neq 0, 1$ be a squarefree integer and let p be an odd prime. Then p factors in $\mathcal{O}_d$ in the following way:*

[46] Quadratic=square, residue=remainder.

1. *If* $\left(\frac{d}{p}\right) = 1$ *then* p *splits in* $\mathcal{O}_d$*:*

$$(p) = (p, a + \sqrt{d})(p, a - \sqrt{d})$$

where $a \in \mathbf{Z}$ *satisfies* $a^2 \equiv d \pmod{p}$.

2. *If* $\left(\frac{d}{p}\right) = -1$ *then* p *is inert in* $\mathcal{O}_d$.

3. *If* $\left(\frac{d}{p}\right) = 0$ *then* p *ramifies in* $\mathcal{O}_d$*:*

$$(p) = (p, \sqrt{d})^2.$$

Now we can complete the story by discussing how 2 factors in quadratic integer rings:

Theorem 16.4. *Let* $d \neq 0, 1$ *be a squarefree integer. Then* 2 *factors in* $\mathcal{O}_d$ *in the following way:*

1. *If* $d \not\equiv 1 \pmod{4}$ *then* 2 *ramifies in* $\mathcal{O}_d$*. More precisely if* $2 \mid d$ *then*

$$(2) = (2, \sqrt{d})^2$$

while if $d \equiv 3 \pmod{4}$ *then*

$$(2) = (2, 1 + \sqrt{d})^2.$$

2. *If* $d \equiv 1 \pmod{8}$ *then* 2 *splits in* $\mathcal{O}_d$*:*

$$(2) = (2, \frac{1 + \sqrt{d}}{2})(2, \frac{1 - \sqrt{d}}{2})$$

3. *If* $d \equiv 5 \pmod{8}$ *then* 2 *is inert in* $\mathcal{O}_d$ *(i.e.* (2) *is prime).*

You should try to check this yourself! Here is the proof for completeness.

Proof. We first check all the claimed factorizations:

- If $2 \mid d$ then exactly as when p is odd, we have

$$(p, \sqrt{d})^2 = (p^2, p\sqrt{d}, d) \subseteq (p)$$

and the ideal on the left contains $\gcd(p^2, d) = p$. (d square free implies that $p^2 \nmid d$.)

- If $d \equiv 3 \pmod{4}$ then

$$(2, 1 + \sqrt{d})^2 = (4, 2(1 + \sqrt{d}), 1 + d + 2\sqrt{d}) = (2)$$

where 2 is in the ideal on the left because $1 + d + 2\sqrt{d} - 2(1 + \sqrt{d}) = d - 1 \equiv 2 \pmod{4}$ is and hence so is $2 = \gcd(4, d - 1)$.

- If $d \equiv 1 \pmod 8$ then

$$(2, \frac{1+\sqrt{d}}{2})(2, \frac{1-\sqrt{d}}{2}) = (4, 1+\sqrt{d}, 1-\sqrt{d}, \frac{1-d}{4}) = (2)$$

  where $\frac{1-d}{4}$ is even because $d \equiv 1 \pmod 8$, and in the other direction, $1+\sqrt{d}+1-\sqrt{d} = 2$ is in the ideal on the left.

By considering norms we see that all the ideals occurring in the factorization have norm 2 and thus are prime. We should also check that the two ideals in the case $d \equiv 1 \pmod 8$ are actually different: If they were equal they would contain $\frac{1+\sqrt{d}}{2} + \frac{1-\sqrt{d}}{2} = 1$.

Finally we should check that when $d \equiv 5 \pmod 8$ then (2) is prime. For this, note that $\frac{1+\sqrt{d}}{2}$ satisfies the polynomial $x^2 - x + \frac{1-d}{4}$, whose reduction mod 2 is $x^2+x+1 \in \mathbf{F}_2[x]$. No element of $\mathbf{F}_2$ satisfies this polynomial. On the other hand, if (2) were not prime, there would be a prime ideal $(2) \subset \mathfrak{p} \subset \mathcal{O}_d$ with $\|\mathfrak{p}\| = 2$ and hence $\mathcal{O}_d/\mathfrak{p} \simeq \mathbf{F}_2$. But then $\frac{1+\sqrt{d}}{2} + \mathfrak{p}$ would give an element of $\mathbf{F}_2$ satisfying the polynomial $x^2 + x + 1$, which is impossible. □

We recall that the discriminant $D_{\mathbf{Q}(\sqrt{d})}$ is given by (see Example 6.2)

$$D_{\mathbf{Q}(\sqrt{d})} = \begin{cases} 4d & d \not\equiv 1 \pmod 4 \\ d & d \equiv 1 \pmod 4 \end{cases}$$

Hence we have the following consequence of Theorems 15.4, 16.4:

Corollary 16.5. *A prime p ramifies in $\mathcal{O}_d$ if and only if $p \mid D_{\mathbf{Q}(\sqrt{d})}$.*

Proof. By Theorem 15.4 an odd prime ramifies if and only if it divides d, while by Theorem 16.4 2 ramifies if and only if $d \not\equiv 1 \pmod 4$. □

Remark 16.6. In particular, only finitely many primes ramify in $\mathcal{O}_d$, and so ramification is quite rare! On the other hand it can be shown that p split and p inert both happen infinitely often. In fact, they both happen "half" the time.[47]

These results suggest a connection between the factorization of (p) in $\mathcal{O}_d$, and the factorization mod p of the minimal polynomial $x^2 - d$ of $\sqrt{d}$ (or the minimal polynomial $x^2 - x + \frac{1-d}{4}$ of $\frac{1+\sqrt{d}}{2}$.) We are going to now explore this connection. We begin with a lemma

Lemma 16.7. *Let $\alpha \in \mathbf{C}$ be a nonzero algebraic integer. Let $m_\alpha \in \mathbf{Z}[x]$ be the minimal monic polynomial of α. Then there is an isomorphism $\mathbf{Z}[\alpha] \simeq \mathbf{Z}[x]/(m_\alpha)$.*[48]

We have been using throughout the course that $\mathbf{Q}(\alpha) \simeq \mathbf{Q}[x]/(m_\alpha)$. The proof of this lemma is similar, but we give it carefully to emphasize where we use that m_α is monic.

[47] More precisely $\lim_{X\to\infty} \frac{\#\{p<X | p \text{ splits in } \mathcal{O}_d\}}{\#\{p<X\}} = \frac{1}{2}$. This requires a bit of analytic number theory to prove and is slightly beyond the scope of this course.

[48] If α is just an algebraic number, then we can choose $d \in \mathbf{Z}$ to make $dm_\alpha \in \mathbf{Z}[x]$ and the gcd of the coefficients is 1 (we dm_α say is a primitive polynomial.) Then you can check $\mathbf{Z}[\alpha] \simeq \mathbf{Z}[x]/(dm_\alpha)$. For the proof you will need to use a form of Gauss' lemma discussed in the footnote of 2.2.

Proof. We have a homomorphism

$$\begin{aligned} \phi : \mathbf{Z}[x] &\to \mathbf{Z}[\alpha] \\ f &\mapsto f(\alpha) \end{aligned}$$

Essentially by definition, ϕ is surjective, so to prove the lemma using the first isomorphism theorem, we need to show $\ker(\phi) = (m_\alpha)$. We clearly have $(m_\alpha) \subseteq \ker(\phi)$ as $m_\alpha(\alpha) = 0$. We need to check the other inclusion.

If $f \in \mathbf{Z}[x]$ with $f(\alpha) = 0$, then we know that $f = gm_\alpha$ for $g \in \mathbf{Q}[x]$. (At this point one can use Gauss' lemma to conclude that $g \in \mathbf{Z}[x]$, but not quite in the form we proved in Lemma 2.2 since g is not necessarily monic.) Thus if $\deg(f) < \deg(m_\alpha)$, $f = 0$. We now prove that $f \in (m_\alpha)$ by induction on $\deg(f)$, with the base case being that $\deg(f) < \deg(m_\alpha)$.

For the inductive step, if $f = ax^n + \cdots$ with $n \geq \deg(m_\alpha)$, write

$$f = ax^{n-\deg(m_\alpha)}m_\alpha + g$$

where $\deg(g) < n$. Then plugging in α we see that $g(\alpha) = 0$ also, so $g = hm_\alpha$ for $h \in \mathbf{Z}[x]$, and hence $f = (ax^{n-\deg(m_\alpha)} + h)m_\alpha \in (m_\alpha)$. □

Remark 16.8. The proof we just gave essentially boils down to polynomial long division, making crucial use of the fact that m_α is monic. More generally, if R is any commutative ring and $g \in R[x]$ is a *monic* polynomial, then for any $f \in R[x]$ we can uniquely write $f = qg + r$ where $r \in R[x]$ has $\deg(r) < \deg(g)$. This has important consequences in algebra, including that $R[x]/(g)$ has $1, x, \ldots, x^{n-1}$ as an R-module basis.

Now suppose $d \not\equiv 1 \pmod 4$ so that $\mathcal{O}_d = \mathbf{Z}[\sqrt{d}]$. By Lemma 16.7 we have an isomorphism

$$\mathcal{O}_d \simeq \mathbf{Z}[x]/(x^2 - d).$$

Using this we can try to study the quotient ring $\mathcal{O}_d/(p)$:

$$\mathcal{O}_d/(p) \simeq \mathbf{Z}[x]/(p, x^2 - d) \simeq \mathbf{F}_p[x]/(x^2 - \overline{d})$$

Here we write $\overline{d} \in \mathbf{F}_p$ for the reduction mod p of d, so $x^2 - \overline{d} \in \mathbf{F}_p[x]$ is the reduction mod p of the polynomial $x^2 - d \in \mathbf{Z}[x]$.

Now we have three possibilities for the polynomial $\overline{f} = x^2 - \overline{d} \in \mathbf{F}_p[x]$:

- $\overline{f} = (x - a)(x - b)$ where $a, b \in \mathbf{F}_p$, $a \neq b$.
- $\overline{f} = (x - a)^2$ for $a \in \mathbf{F}_p$.
- $\overline{f}$ is irreducible.

In each case we can describe the quotient ring $\mathbf{F}_p[x]/(x^2 - \overline{d})$:

- $\mathbf{F}_p[x]/(x^2 - \overline{d}) \simeq \mathbf{F}_p[x]/(x - a) \times \mathbf{F}_p[x]/(x - b) \simeq \mathbf{F}_p \times \mathbf{F}_p$ (the first isomorphism is by CRT.)
- $\mathbf{F}_p[x]/(x^2 - \overline{d}) \simeq \mathbf{F}_p/((x - a)^2)$.
- $\mathbf{F}_p[x]/(x^2 - \overline{d})$ is a field.

On the other hand we can also describe $\mathcal{O}_d/(p)$ according to whether p splits, ramifies, or is inert:

- $\mathcal{O}_d/(p) \simeq \mathcal{O}_d/(\mathfrak{p}_1\mathfrak{p}_2) \simeq \mathcal{O}_d/\mathfrak{p}_1 \times \mathcal{O}_d/\mathfrak{p}_2 \simeq \mathbf{F}_p \times \mathbf{F}_p$ if p splits, again using CRT.
- $\mathcal{O}_d/(\mathfrak{p}^2)$ if p ramifies.
- $\mathcal{O}_d/(p)$ is a field if p is inert.

To complete this discussion we should explain why $\mathbf{F}_p[x]/((x-a)^2)$ and $\mathcal{O}_d/\mathfrak{p}^2$ are neither fields nor $\mathbf{F}_p \times \mathbf{F}_p$, and thus must match. In fact, these rings contain an element $r \neq 0$ with $r^2 = 0$ (such an r is called nilpotent.) Neither a field nor $\mathbf{F}_p \times \mathbf{F}_p$ can contain such an element. On the other hand in $\mathbf{F}_p[x]/((x-a)^2)$ we have $r = x - a$ while in in $\mathcal{O}_d/\mathfrak{p}^2$ we can take any $r \in \mathfrak{p}$, $r \notin \mathfrak{p}^2$.

To summarize this discussion, we have proved that we can determine how (p) factors in $\mathcal{O}_d$ in terms of how $\overline{f}$ factors in $\mathbf{F}_p[x]$:

- p splits if and only if $\overline{f} = (x-a)(x-b)$ with $a \neq b$.
- p ramifies if and only if $\overline{f} = (x-a)^2$.
- p is inert if and only if $\overline{f}$ is irreducible.

We now state a massive generalization of this:

Theorem 16.9 (Dedekind's factorization criteria)**.** *Let $K = \mathbf{Q}(\alpha)$ for $\alpha \in \mathbf{C}$ a nonzero algebraic integer with minimal polynomial f. Let p be a prime with $p \nmid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$.*

Then if

$$\overline{f} = \overline{g}_1^{e_1} \cdots \overline{g}_r^{e_r}$$

is the factorization of $\overline{f} \in \mathbf{F}_p[x]$ with $\overline{g}_1, \dots, \overline{g}_r$ pairwise distinct, monic, irreducible, then

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$$

where

$$\mathfrak{p}_i = (p, g_i(\alpha))$$

are pairwise distinct primes of $\mathcal{O}_K$, and moreover $f(\mathfrak{p}_i) = \deg(g_i)$.[49]

17 Lecture 17: Splitting of primes IV

We begin with some remarks on Dedekind's factorization criteria. First we address the condition $p \nmid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$.

- If we are lucky, we already know $\mathcal{O}_K = \mathbf{Z}[\alpha]$ (e.g. because we computed $\operatorname{disc}(\mathbf{Z}[\alpha])$ and found it was squarefree) and then the criteria works for all primes.
- On the other hand, we might not even know what $\mathcal{O}_K$ is, but recall the formula of Proposition 6.9

  $$\operatorname{disc}(\mathbf{Z}[\alpha]) = [\mathcal{O}_K : \mathbf{Z}[\alpha]]^2 \cdot D_K.$$

  This implies that if $p \mid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$ then $p^2 \mid \operatorname{disc}(\mathbf{Z}[\alpha])$.

  This is great because we might not easily be able to calculate $\mathcal{O}_K$, D_K, and $[\mathcal{O}_K : \mathbf{Z}[\alpha]]$, but we can more easily calculate $\operatorname{disc}(\mathbf{Z}[\alpha])$.

[49] Recall $f(\mathfrak{p}_i)$ is the inertial degree of the prime $\mathfrak{p}_i$, i.e. $\|\mathfrak{p}_i\| = p^{f(\mathfrak{p}_i)}$. The f here has nothing to do with the polynomial f. The notation is not great, sorry!

- The primitive element theorem states that any number field K has the form $\mathbf{Q}(\alpha)$ for some algebraic integer α. You might then wonder if in fact $\mathcal{O}_K$ always has the form $\mathbf{Z}[\alpha]$ for some $\alpha \in \mathcal{O}_K$. Not only is this not the case, it can happen that there is a prime p such that $p \mid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$ for all $\alpha \in \mathcal{O}_K$ with $K = \mathbf{Q}(\alpha)$ (see example sheet 3 for a famous example due to Dedekind.) Thus we cannot necessarily factor all primes using Dedekind's criteria. Nonetheless, a more optimistic point of view is that we can always factor all but finitely many primes using Dedekind's criteria.

Example 17.1. Let's use Dedekind's criteria to factor an odd prime p in $\mathcal{O}_d$ and thus reprove Theorem 15.4. We will let $\alpha = \sqrt{d}$. Then

$$[\mathcal{O}_d : \mathbf{Z}[\sqrt{d}]] = \begin{cases} 1 & d \not\equiv 1 \pmod 4 \\ 2 & d \equiv 1 \pmod 4 \end{cases}$$

Since p is odd, $p \nmid [\mathcal{O}_d : \mathbf{Z}[\sqrt{d}]]$ and so Dedekind's criteria will apply.

We have to factor the minimal polynomial $f = x^2 - d \bmod p$. The factorization is given by

$$\overline{f} = \begin{cases} (x - \overline{a})(x + \overline{a}) & \left(\frac{d}{p}\right) = 1 \\ x^2 - \overline{d} & \left(\frac{d}{p}\right) = -1 \\ x^2 & \left(\frac{d}{p}\right) = 0 \end{cases}$$

where in the first case $\overline{a} \in \mathbf{F}_p$ satisfies $\overline{a}^2 = \overline{d}$.

In the first case, we can choose a lift $a \in \mathbf{Z}$ of $\overline{a}$ (so $a^2 \equiv d \pmod p$) and use $x - a$ and $x + a$ as our lifts to $\mathbf{Z}[x]$ of the two factors, so we then get

$$(p) = (p, \sqrt{d} - a)(p, \sqrt{d} + a).$$

In the second case we can just use $f = x^2 - d$ as our lift of the single irreducible factor, and we get $(p, f(\sqrt{d})) = (p)$ is prime. In the third case we get that $(p) = (p, \sqrt{d})^2$.

Exercise 17.2. Use Dedekind's criteria to factor $p = 2$ in $\mathcal{O}_d$ and thus prove Theorem 16.4. When $d \not\equiv 1 \pmod 4$ you should factor $x^2 - d$, the minimal polynomial of $\sqrt{d}$, while if $d \equiv 1 \pmod 4$ you should factor $x^2 - x + \frac{1-d}{4}$, the minimal polynomial of $\frac{1+\sqrt{d}}{2}$.

Example 17.3. We do a higher degree example. Consider $K = \mathbf{Q}(\alpha)$ where $\alpha = \sqrt[3]{2}$ with minimal polynomial $x^3 - 2$. You computed on example sheet 2 that $\operatorname{disc}(\mathbf{Z}[\alpha]) = -27 \cdot 4 = -108$. Thus Dedekind's criteria will apply except possibly for $p = 2, 3$. In fact, it turns out that $\mathcal{O}_K = \mathbf{Z}[\alpha]$ (see the bonus problem of example sheet 1 where we checked this by brute force) and so Dedekind's criteria actually applies for all p. Now we compute:

- $x^3 - 2 \equiv x^3 \pmod 2$, so $(2) = (2, \sqrt[3]{2})^3 = (\sqrt[3]{2})^3$.
- $x^3 - 2 \equiv (x + 1)^3 \pmod 3$, so $(3) = (3, \sqrt[3]{2} + 1)^3 = (\sqrt[3]{2} + 1)^3$.[50]
- $x^3 - 2 \equiv (x - 3)(x^2 + 3x - 1) \pmod 5$ so $(5) = (5, \sqrt[3]{2} - 3)(5, \sqrt[3]{2}^2 + 3\sqrt[3]{2} - 1)$.

[50] To see the last equality, if $\beta = \sqrt[3]{2} + 1$ then β satisfies the polynomial $(x-1)^3 - 2 = x^3 - 3x^2 + 3x - 3$, so $3 = \beta^3 - 3\beta^2 + 3\beta \in (\beta)$.

- $x^3 - 2$ is irreducible mod 7, so (7) is prime.

Exercise 17.4. Here are a few exercises exploring this example further.

1. Show that if $p \equiv 2 \pmod 3$ is an odd prime, then
$$\begin{aligned} \mathbf{F}_p^\times &\to \mathbf{F}_p^\times \\ x &\mapsto x^3 \end{aligned}$$
is a bijection (Hint: what is the order of $\mathbf{F}_p^\times$?) and deduce that the congruence $a^3 \equiv 2 \pmod p$ has a unique solution. Deduce that $(p) = \mathfrak{p}_1\mathfrak{p}_2$ where $f(\mathfrak{p}_1) = 1$ and $f(\mathfrak{p}_2) = 2$.

2. Show that if $p \equiv 1 \pmod 3$ then either $(p) = \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$ is a product of three distinct primes, or (p) is prime, according to whether or not the congruence $a^3 = 2 \pmod p$ has a solution (Hint: show that if it has a solution then it has three distinct solutions by considering elements of order 3 in $\mathbf{F}_p^\times$.) $31, 43, 109, 127, ...$ are the first few primes which split into three distinct factors.[51]

3. Show that the prime factors of (5) found above are actually principal, by finding generators. (In fact, $\mathbf{Z}[\sqrt[3]{2}]$ is a PID as we will prove later. It is even norm-Euclidean!)

Now we prove Dedekind's factorization criteria. We begin with a lemma.

Lemma 17.5. *Suppose $p \nmid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$. Then the map*
$$\begin{aligned} \mathbf{Z}[\alpha]/(p) &\to \mathcal{O}_K/(p) \\ x + p\mathbf{Z}[\alpha] &\mapsto x + p\mathcal{O}_K \end{aligned}$$
is an isomorphism.

Proof. This is a homomorphism between rings of finite order $p^{[K:\mathbf{Q}]}$, so we can check it is an isomorphism by showing that it is surjective. By Lagrange's theorem, for any $x+\mathbf{Z}[\alpha] \in \mathcal{O}_K/\mathbf{Z}[\alpha]$, $[\mathcal{O}_K : \mathbf{Z}[\alpha]](x+\mathbf{Z}[\alpha]) = 0+\mathbf{Z}[\alpha]$, or in other words, $[\mathcal{O}_K : \mathbf{Z}[\alpha]]\mathcal{O}_K \subseteq \mathbf{Z}[\alpha]$. Also since $p \nmid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$ we can write $1 = a[\mathcal{O}_K : \mathbf{Z}[\alpha]] + bp$ for $a, b \in \mathbf{Z}$. Now given $x + p\mathcal{O}_K \in \mathcal{O}_K/(p)$, we have
$$x + p\mathcal{O}_K = (a[\mathcal{O}_K : \mathbf{Z}[\alpha]] + bp)x + p\mathcal{O}_K = a[\mathcal{O}_K : \mathbf{Z}[\alpha]]x + p\mathcal{O}_K$$
and $a[\mathcal{O}_K : \mathbf{Z}[\alpha]]x \in \mathbf{Z}[\alpha]$. So the map is surjective an hence an isomorphism. □

Example 17.6. Suppose $d \equiv 1 \pmod 4$ so that $\mathbf{Z}[\sqrt{d}] \subsetneq \mathcal{O}_d = \mathbf{Z}[\frac{1+\sqrt{d}}{2}]$. You can think about Lemma 17.5 and its proof like this: $\mathbf{Z}[\sqrt{d}]/(p)$ for p an odd prime really already contains $\frac{1+\sqrt{d}}{2}$ because 2 is invertible mod p.

[51] I can't resist telling you that the primes $p \equiv 1 \pmod 3$ for which the congruence $a^3 \equiv 2 \pmod 3$ has a solution are exactly the primes of the form $p = x^2 + 27y^2$. This was conjectured by Euler (I have no idea how he noticed this!) and proved by Gauss using cubic reciprocity. We won't discuss higher reciprocity laws in this course, but this is well within reach to learn about if you are interested.

Proof of Theorem 16.9. By Lemma 16.7 from last time, we have an isomorphism $\mathbf{Z}[x]/(f) \to \mathbf{Z}[\alpha]$. Combining this with Lemma 17.5 we have isomorphisms

$$\mathbf{F}_p[x]/(\overline{f}) \simeq \mathbf{Z}[\alpha]/(p) \simeq \mathcal{O}_K/(p).$$

It follows that

$$\mathcal{O}_K/(p, g_i(\alpha)) \simeq \mathbf{Z}[\alpha]/(p, g_i(\alpha)) \simeq \mathbf{F}_p[x]/(\overline{f}, \overline{g}_i).$$

As $\overline{g}_i$ is a factor of $\overline{f}$, $\mathbf{F}_p[x]/(\overline{f}, \overline{g}_i) = \mathbf{F}_p[x]/(\overline{g}_i)$ and as $\overline{g}_i$ is irreducible, it is a field of degree $\deg(\overline{g}_i)$ over $\mathbf{F}_p$. It follows that $\mathfrak{p}_i = (p, g_i(\alpha))$ is prime and $f(\mathfrak{p}_i) = \deg(\overline{g}_i)$.

Next we check that $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are distinct. Indeed for $i \neq j$ we check

$$\mathcal{O}_K/(\mathfrak{p}_i + \mathfrak{p}_j) \simeq \mathcal{O}_K/(p, g_i(\alpha), g_j(\alpha)) \simeq \mathbf{F}_p[x]/(\overline{g}_i, \overline{g}_j) = \{0\}$$

since $(\overline{g}_i, \overline{g}_j) = \mathbf{F}_p[x]$ as $\overline{g}_i, \overline{g}_j$ are coprime. Hence $\mathfrak{p}_i + \mathfrak{p}_j = \mathcal{O}_K$, and hence $\mathfrak{p}_i \neq \mathfrak{p}_j$.

Finally we observe that

$$\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r} \subseteq (p, g_1^{e_1}(\alpha) \cdots g_r^{e_r}(\alpha)) \subseteq (p)$$

where the first inclusion just follows from multiplying out all the generators of the first product, and the second inclusion follows from $g_1^{e_1} \cdots g_r^{e_r} \equiv f \pmod{p}$ and hence

$$(g_1^{e_1} \cdots g_r^{e_r})(\alpha) = f(\alpha) + (g_1^{e_1} \cdots g_r^{e_r} - f)(\alpha) \in p\mathcal{O}_K$$

since $f(\alpha) = 0$. Now we take norms. We have

$$\|\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}\| = p^{\deg(f)} = p^{[K:\mathbf{Q}]} = \|(p)\|$$

and hence the inclusion $\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r} \subseteq (p)$ is an equality. □

18 Lecture 18: Ramification, Eisenstein polynomials

Definition 18.1. We say that the prime p *ramifies* in K if in the unique factorization of (p) into prime ideals of $\mathcal{O}_K$ there is a repeated factor. Equivalently there exists some prime $\mathfrak{p}$ lying above p for which the ramification index $e(\mathfrak{p}) > 1$.

We are not going to prove the following theorem in this course, but we have seen it is true in the case of quadratic fields in Corollary 16.5.

Theorem 18.2. *Let K be a number field. Then p ramifies in K if and only if $p \mid D_K$.*

An important consequence is that ramification is "rare":

Corollary 18.3. *Only finitely many primes ramify in K.*

We want to explain this corollary from two points of view. First of all, we have seen that if $K = \mathbf{Q}(\alpha)$ for α an algebraic integer, then all but finitely many primes can be factored by factoring the minimal polynomial f of α mod p. Thus we can view Corollary 18.3 as a very concrete statement about how the reductions mod p of f factor:

Proposition 18.4. *Let $f \in \mathbf{Z}[x]$ be an irreducible monic polynomial.*[52] *Then for all but finitely many primes p, $\overline{f} \in \mathbf{F}_p[x]$ has no repeated irreducible factors.*

[52] More generally this is true with the same proof if you just assume f has no repeated factors.

This proposition, combined with Dedekind's factorization criteria implies Corollary 18.3 for $K = \mathbf{Q}(\alpha)$ (and hence for all number fields if you use the primitive element Theorem.)

Proof. We recall that if F is a field and $f \in F[x]$ is a polynomial and $g^2 \mid f$ for some $g \in F[x]$, then taking derivatives we conclude that $g \mid f'$. Indeed if $f = g^2 h$ then by Leibniz, $f' = 2gg'h + g^2 h' = g(2gh + gh')$. As a consequence, if $(f, f') = F[x]$ then f cannot have a repeated factor.

Now f, f' are coprime in $\mathbf{Q}[x]$ because f is irreducible and f' is nonzero but of degree smaller than f. Hence we can write

$$1 = gf + hf'$$

for $g, h \in \mathbf{Q}[x]$. But now we can clear denominators, i.e. choose $d > 0$ such that $dg, dh \in \mathbf{Z}[x]$, and conclude that

$$d = (dg)f + (dh)f'.$$

Now if $p \nmid d$, reducing this formula mod p shows that $(\overline{f}, \overline{f}') = (\overline{d}) = \mathbf{F}_p[x]$. It follows that $\overline{f}$ cannot have a repeated irreducible factor. □

Remark 18.5. In fact d, in the proof can be taken to be the discriminant of f.[53]

We now sketch a proof of one direction of Theorem 18.2, namely that p ramified implies that $p|D_K$. Note that this also implies Corollary 18.3.

The idea is to make use of something we observed in lecture 16: if $(p) = \mathfrak{p}^2\mathfrak{p}_2 \cdots \mathfrak{p}_r$ is a ramified prime, then we can pick $x \in \mathfrak{p}\mathfrak{p}_2 \cdots \mathfrak{p}_r$, $x \notin (p)$. Then $x^2 \in (p)$. Hence $x + (p) \in \mathcal{O}_K/(p)$ is a nilpotent element. In fact, one way to say that p ramifies is to say that $\mathcal{O}_K/(p)$ contains nonzero nilpotents.

We now recall from linear algebra that if $A \in M_n(F)$, F a field, is a matrix, then we say A is nilpotent if $A^r = 0$ for some $r > 0$. This implies that all the eigenvalues of A are 0, and hence $\chi_A = x^n$, and so in particular $\mathrm{tr}(A) = 0$. Consequently, if $A \in M_n(\mathbf{Z})$ is a matrix whose reduction mod p, $\overline{A} \in M_n(\mathbf{F}_p)$ is nilpotent, then $\mathrm{tr}(A) \equiv 0 \pmod{p}$.

Now we pick an integral basis $\alpha_1, \ldots, \alpha_n$ for $\mathcal{O}_K$ with the property that $\alpha_1 + (p) \in \mathcal{O}_K/(p)$ is nilpotent.[54] The same is then true of $\alpha_1\alpha_i$ for $i = 1, \ldots, n$, and hence $p \mid$

[53] With a bit more effort, you can even prove Theorem 18.2 in the case that $\mathcal{O}_K = \mathbf{Z}[\alpha]$, using Dedekind's factorization criteria. But this won't work in general.

[54] This is really the only step in this "sketch" which I didn't justify, so let me do it here in the footnote. We start with any integral basis $\alpha_1, \ldots, \alpha_n$. We write $\overline{\alpha_1}, \ldots, \overline{\alpha_n} \in \mathcal{O}_K/(p)$ for the reductions mod p. This is a basis for $\mathcal{O}_K/(p)$ as a $\mathbf{F}_p$-vector space. Now choose a nonzero nilpotent $x \in \mathcal{O}_K/(p)$. We can write $x = \overline{c}_1\overline{\alpha}_1 + \cdots + \overline{c}_n\overline{\alpha}_n$. By possibly reordering we may assume $\overline{c}_1 \neq 0$, and replacing x with $\overline{c}_1^{-1}x$ which is still nilpotent, we may assume that $\overline{c}_1 = 1$. Now take $\alpha_1' = \alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$, where $c_2, \ldots, c_r \in \mathbf{Z}$ are any lifts of $\overline{c}_2, \ldots, \overline{c}_n$. Then α_1' reduces to the nilpotent x, and $\alpha_1', \alpha_2, \ldots, \alpha_n$ is also an integral basis, because $\alpha_1 = \alpha_1' - c_2\alpha_2 - \cdots - c_r\alpha_n$. I find this a bit awkward but there is nothing fancy going on!

In the lecture I mumbled something else which might have been confusing, so let me clarify it here for anyone who is interested. To be very concrete we consider $\mathbf{Z}^n$ and $\mathbf{Z}^n/p\mathbf{Z}^n = \mathbf{F}_p^n$. You can ask, given a basis $\overline{v}_1, \ldots, \overline{v}_n$ of $\mathbf{F}_p^n$, is there a basis $v_1, \ldots, v_n \in \mathbf{Z}^n$ lifting it? Perhaps somewhat surprisingly, the answer is not always! One way to say that $\overline{v}_1, \ldots, \overline{v}_n$ is a basis is that the matrix with columns $\overline{v}_1, \ldots, \overline{v}_n$ is invertible. From this point of view the question is: is the reduction map $M_n(\mathbf{Z})^\times \to M_n(\mathbf{F}_p)^\times$ surjective? The answer is no if $p > 3$, any element of the image will have determinant ± 1. But this is the only obstruction: $\mathrm{SL}_n(\mathbf{Z}) \to \mathrm{SL}_n(\mathbf{F}_p)$ is surjective, where SL_n denotes matrices with determinant 1. If you like linear algebra, one way to prove this is to show that the "elementary matrices" $I + aE_{ij}$ generate $\mathrm{SL}_n(\mathbf{F}_p)$ where E_{ij} denotes the matrix with a 1 in the ith row and jth column, and $i \neq j$. Going back to the original problem, this means that you can at least always lift $c\overline{v}_1, v_2, \ldots, \overline{v}_n$ for some nonzero $c \in \mathbf{F}_p$.

$\mathrm{tr}_K(\alpha_1\alpha_i)$. It follows that $p \mid D_K = \det(\mathrm{tr}(\alpha_i\alpha_j))$, since this matrix has integer entries and the first column consists of multiples of p.

Exercise 18.6. The goal of this exercise (which is just for fun!) is to refine the above sketched argument to prove that if $(p) = \mathfrak{p}_1^{e(\mathfrak{p}_1)} \cdots \mathfrak{p}_r^{e(\mathfrak{p}_r)}$ then $p^k \mid D_K$ where $k = (e(\mathfrak{p}_1) - 1)f(\mathfrak{p}_1) + \cdots + (e(\mathfrak{p}_r) - 1)f(\mathfrak{p}_r)$.[55]

1. Show that the set of nilpotent elements of $\mathcal{O}_k/(p)$ is exactly $\mathfrak{p}_1 \cdots \mathfrak{p}_r/(p)$.[56]
2. Show that $\dim_{\mathbf{F}_p}(\mathfrak{p}_1 \cdots \mathfrak{p}_r/(p)) = k$ (Hint: first show $\#\mathfrak{p}_1 \cdots \mathfrak{p}_r/(p) = p^k$ using ideal norms.)
3. Show that there is an integral basis $\alpha_1, \ldots, \alpha_n$, $n = [K : \mathbf{Q}]$ for $\mathcal{O}_K$ with the property that $\alpha_1, \ldots, \alpha_k$ are nilpotent mod (p). Conclude that $p^k \mid D_K$.

18.1 Eisenstein polynomials

Definition 18.7. A monic polynomial $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbf{Z}[x]$ is called Eisenstein for a prime p if $p|a_i$ for $i = 0, \ldots, n-1$ and $p^2 \nmid a_0$.

Eisenstein's criteria states that an Eisenstein polynomial is irreducible. Our goal now is to understand Eisenstein polynomials a bit better from the point of view of algebraic number theory.

Example 18.8.
1. $x^n - p$ is Eisenstein for p.
2. Recall the pth cyclotomic polynomial $\Phi_p(x) = \frac{x^p-1}{x-1} = x^{p-1} + x^{p-2} + \cdots + 1$ which has $\zeta_p = e^{e\pi i/p}$. If we consider instead the polynomial with $\zeta_p - 1$ as a root:

$$\Phi_p(x+1) = \frac{(x+1)^p - 1}{x} = x^{p-1} + px^p + \binom{p}{2}x^{p-2} + \cdots + \binom{p}{2}x + p$$

this is Eisenstein at p. Hence $\Phi_p(x+1)$ and then also $\Phi_p(x)$ are irreducible.

Proposition 18.9. *Let $K = \mathbf{Q}(\alpha)$ for a nonzero algebraic integer α, and suppose that the minimal polynomial f of α is Eisenstein at p. Then $p \nmid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$, and*

$$(p) = (p, \alpha)^{[K:\mathbf{Q}]}.$$

Remark 18.10. We say that a prime p is *totally ramified* in a number field K if $(p) = \mathfrak{p}^{[K:\mathbf{Q}]}$, i.e. there is a single prime $\mathfrak{p}$ of $\mathcal{O}_K$ lying over p, and $e(\mathfrak{p}) = [K : \mathbf{Q}]$, $f(\mathfrak{p}) = 1$.

Proof. We write $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ for the minimal monic polynomial of α.

If $p \mid [\mathcal{O}_K : \mathbf{Z}[\alpha]]$ there exists $x \in \mathcal{O}_K$, $x \notin \mathbf{Z}[\alpha]$ but with $px \in \mathbf{Z}[\alpha]$ (indeed choose an element $x + \mathbf{Z}[\alpha] \in \mathcal{O}_k/\mathbf{Z}[\alpha]$ of order p)

Then we can write

$$x = \frac{c_0}{p} + \frac{c_1}{p}\alpha + \frac{c_2}{p}\alpha^2 + \cdots + \frac{c_{n-1}}{p}\alpha^{n-1} \in \mathcal{O}_K$$

[55] In fact, p^k is the largest power of p dividing D_K if and only if none of the $e(\mathfrak{p}_i)$ are multiples of p, but this is harder to prove. We say that p is *tamely ramified* if none of $e(\mathfrak{p}_i)$ is a multiple of p and *wildly ramified* otherwise.

[56] In commutative algebra terminology, this is the nilradical.

for $c_0, \ldots, c_{n-1} \in \mathbf{Z}$ and not all $c_i \in p\mathbf{Z}$.

Now let $i \geq 0$ be the smallest number for which $c_i \notin p\mathbf{Z}$. Then

$$\frac{c_i}{p}\alpha^i + \cdots + \frac{c_{n-1}}{p}\alpha^{n-1} \in \mathcal{O}_K.$$

We multiply this by α^{n-i-1} and conclude

$$\frac{c_i}{p}\alpha^{n-1} + \frac{\alpha^n}{p}(c_{i+1} + c_{i+2}\alpha + \cdots + c_{n-1}\alpha^{n-i-2}) \in \mathcal{O}_K$$

But

$$\alpha^n = -a_0 - a_1\alpha - \cdots - a_{n-1}\alpha^{n-1} \in p\mathcal{O}_K$$

so $\frac{\alpha^n}{p} \in \mathcal{O}_K$. Hence we conclude that

$$\frac{c_i}{p}\alpha^{n-1} \in \mathcal{O}_K.$$

Now we take the norm of this. We conclude

$$N_K(\frac{c_i}{p}\alpha^{n-1}) = \frac{c_i^n}{p^n}N_K(\alpha)^{n-1} = \pm\frac{c_i^n}{p^n}a_0^{n-1} \in \mathbf{Z}$$

Now $p|a_0$ but $p^2 \nmid a_0$ by the definition of an Eisenstein polynomial. Hence $p^n \nmid a_0^{n-1}$. It follows that $p \mid c_i^n$ and hence $p \mid c_i$, which is a contradiction.

For the last statement, we can apply Dedekind's criteria: reducing f mod p we have $\overline{f} = x^n$. Hence

$$(p) = (p, \alpha)^n.$$

□

If you are lucky, Proposition 18.9 can be help you easily compute rings of integers in number fields. At the same time, it explains why it can't always be used: unless p is totally ramified in K we can not expect to find an $\alpha \in \mathcal{O}_K$ which satisfies an Eisenstein polynomial for p. Here is an example:

Example 18.11. Let's prove that if $K = \mathbf{Q}(\sqrt[3]{2})$ then $\mathcal{O}_K = \mathbf{Z}[\sqrt[3]{2}]$. We recall that $\mathrm{disc}(\mathbf{Z}[\sqrt[3]{2}]) = -3^3 \cdot 2^2$ (see example sheet 2). Thus it suffices to prove that $2, 3 \nmid [\mathcal{O}_K : \mathbf{Z}[\sqrt[3]{2}]]$.

But $x^3 - 2$ is an Eisenstein polynomial at 2 so $2 \nmid [\mathcal{O}_K : \mathbf{Z}[\sqrt[3]{2}]]$ by Proposition 18.9. What about 3? Well as in the case of the cyclotomic polynomial, there is an art to finding Eisenstein polynomials! We consider the minimal polynomial of $\sqrt[3]{2} - 2$:

$$(x+2)^3 - 2 = x^3 + 6x^2 + 12x + 6$$

which is Eisenstein at 3. Hence $3 \nmid [\mathcal{O}_K : \mathbf{Z}[\sqrt[3]{2}+1]]$ by Proposition 18.9, but $\mathbf{Z}[\sqrt[3]{2}+1] = \mathbf{Z}[\sqrt[3]{2}]$

In this course I am choosing to not emphasize this kind of calculation very much, but if you want to try more examples, here is an exercise:

Exercise 18.12. 1. Show that $\mathcal{O}_{\mathbf{Q}(\sqrt[4]{2})} = \mathbf{Z}[\sqrt[4]{2}]$.

2. Let $d \neq \pm 1$ be squarefree and consider the cubic field $K = \mathbf{Q}(\sqrt[3]{d})$. Show that the discriminant of $\mathbf{Z}[\sqrt[3]{d}]$ is $-27d^2$, and hence for each prime factor of the discriminant except possibly 3, $x^3 - d$ is Eisenstein at p.

   To handle $p = 3$ show that if $d \neq \pm 1 \pmod 9$ then the minimal polynomial of $\sqrt[3]{d} - d$ is Eisenstein at 3, and deduce that in this case $\mathcal{O}_K = \mathbf{Z}[\sqrt[3]{d}]$. On the other hand, show that if $d \equiv \pm 1 \pmod 9$ then $\frac{(\sqrt[3]{d}-d)^2}{3}$ is an algebraic integer.

3. Let p be an odd prime. Show that if $2^{p-1} \not\equiv 1 \pmod{p^2}$ then the minimal polynomial of $\sqrt[p]{2} - 2$ is Eisenstein for p. Using this prove that $\mathcal{O}_{\mathbf{Q}(\sqrt[p]{2})} = \mathbf{Z}[\sqrt[p]{2}]$. Can you see how to prove that if $2^{p-1} \equiv 1 \pmod{p^2}$ then $\frac{(\sqrt[p]{2}-2)^{p-1}}{p}$ is an algebraic integer?[57]

Finally here is another problem just for fun:

Exercise 18.13. 1. Let K be a number field of degree $[K : \mathbf{Q}] = n$. Suppose that $(p) = \mathfrak{p}^n$ (i.e. p is totally ramified.) Show that if $\alpha \in \mathfrak{p}$, $\alpha \notin \mathfrak{p}^2$, then the minimal monic polynomial of α is an Eisenstein polynomial of degree n. This is a sort of converse to Proposition 18.9. (Hint: it may first be helpful to think about what is the largest r such that $c\alpha^k \in \mathfrak{p}^r$ for $c \in \mathbf{Z}$.)

2. Show that $p^{n-1} \mid D_K$ and moreover this is the largest power of p dividing D_K if and only if $p \nmid n$ (Hint: again it might be helpful to think about the largest r such that $f'(\alpha) \in \mathfrak{p}^r$.)

19 Lecture 19: Example class 3

1. Give the factorization of $2, 3, 5, 7, 11$ in $\mathcal{O}_{-10} = \mathbf{Z}[\sqrt{-10}]$. Determine which of the prime ideals you found are principal ideals. (If you'd like you can try to show that all the non principal ideals you found lie in the same ideal class.)

2. Let $\mathcal{O}_d$ be a quadratic integer ring. Prove that if $I \subseteq \mathcal{O}_d$ is a nonzero ideal, then

$$I\overline{I} = (\|I\|).$$

   (Hint: try to prove this when $I = \mathfrak{p}$ is prime first.). Deduce that in $\mathrm{Cl}(\mathcal{O}_d)$, $[I]^{-1} = [\overline{I}]$.

3. Let $\mathcal{O}_d$ be a quadratic integer ring and let p be a prime. How many ideals $I \subseteq \mathcal{O}_d$ are there with $\|I\| = p^k$? The answer should only depend on whether p is inert, split, or ramified, think about the factorization of I into prime ideals. Can you give a formula for the number of ideals of norm $n = p_1^{k_1} \cdots p_r^{k_r}$?

4. Let $K/\mathbf{Q}$ be a cubic field (i.e. $[K : \mathbf{Q}] = 3$). List all the possible ways that a prime p can split in $\mathcal{O}_K$ (i.e. how many factors are there, and what are $e(\mathfrak{p})$ and $f(\mathfrak{p})$ for each factor.)

[57] A prime p with the property that $2^{p-1} \equiv 1 \pmod{p^2}$ is called a Wieferich prime. Like many things in this subject, they originally arose in connection with Fermat's last theorem: Wieferich proved in 1909 that if $x^p + y^p = z^p$ and p is not a Wieferich prime, then $p \mid xyz$. Only two are known: 1093 and 3511. It is conjectured that there are infinitely many, but even finding the third remains elusive: the distributed computing project PrimeGrid has shown that the third Wieferich prime is at least 2^{64}.

5. Let α be a root of the polynomial x^3+x+1 and consider $K=\mathbf{Q}(\alpha)$. We computed in Example 6.16 in the notes that $\mathrm{disc}(\mathbf{Z}[\alpha])=-31$ and hence $\mathcal{O}_K=\mathbf{Z}[\alpha]$. Use Dedekind's criteria to factor $2,3,5$ in $\mathcal{O}_K$. (Bonus: can you factor 31?)

6. This problem is about a famous example of Dedekind of a number field K such that $\mathcal{O}_K$ is not of the form $\mathbf{Z}[\alpha]$ for any $\alpha\in\mathcal{O}_K$. Part a is conceptual but the rest is mostly computational and just for fun.

   (a) Show that if $[K:\mathbf{Q}]=3$ and $(2)=\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$ factors as a product of three distinct primes in $\mathcal{O}_K$, then $\mathcal{O}_K\neq\mathbf{Z}[\alpha]$ for any $\alpha\in\mathcal{O}_K$. (Hint: can a polynomial $f\in\mathbf{F}_2[x]$ factor as a product of 3 distinct linear factors?)

   (b) Show that the polynomial x^3-x^2-2x-8 is irreducible and let $\theta\in\mathbf{C}$ be a root. Let $K=\mathbf{Q}(\theta)$. Show that $\theta'=4/\theta\in\mathcal{O}_K$. (Hint: try dividing the equation $\theta^3-\theta^2-2\theta-8=0$ by θ^3 to find the minimal polynomial of θ'.)

   (c) Prove the formulas $\theta^2=\theta+2+2\theta'$ and $(\theta')^2=-\theta'-2+2\theta$ and deduce that $\mathbf{Z}\cdot 1+\mathbf{Z}\cdot\theta+\mathbf{Z}\cdot\theta'\subseteq\mathcal{O}_K$ is a ring.

   (d) Compute $\mathrm{disc}(1,\theta,\theta')=-503$ is prime and deduce that $1,\theta,\theta'$ is an integral basis for $\mathcal{O}_K$.

   (e) Prove that $\mathcal{O}_K/(2)\simeq\mathbf{F}_2\times\mathbf{F}_2\times\mathbf{F}_2$. (Hint: this ring has the form $\mathbf{F}_2\cdot 1\oplus\mathbf{F}_2\cdot\theta\oplus\mathbf{F}_2\cdot\theta'$ with multiplication given by $\theta^2=\theta$, $\theta'^2=\theta'$, $\theta\theta'=0$)

   (f) Conclude that $\mathcal{O}_K\neq\mathbf{Z}[\alpha]$ for any $\alpha\in\mathcal{O}_K$.

20 Lecture 20: Cyclotomic fields and quadratic reciprocity

Today we let p be an odd prime and consider the cyclotomic field $\mathbf{Q}(\zeta_p)$ where $\zeta_p=e^{2\pi i/p}$. As we have seen, the minimal polynomial of ζ_p is $\Phi_p(x)=\frac{x^p-1}{x-1}=1+x+\cdots+x^{p-1}$, and the minimal polynomial of ζ_p-1 is

$$\Phi_p(x+1)=\frac{(x+1)^p-1}{x}=x^{p-1}+px+\cdots+p$$

which is an Eisenstein polynomial for p.

Theorem 20.1. *We have $\mathcal{O}_{\mathbf{Q}(\zeta_p)}=\mathbf{Z}[\zeta_p]$, $D_{\mathbf{Q}(\zeta_p)}=(-1)^{p-1}p^{p-2}$, and*

$$(p)=(\zeta_p-1)^{p-1}$$

is totally ramified.

Proof. We compute the discriminant of $\mathbf{Z}[\zeta_p]$. We start by computing $\Phi_p'(\zeta_p)$. Differentiating

$$(x-1)\Phi_p(x)=x^p-1$$

we obtain

$$\Phi_p(x)+(x-1)\Phi'(x)=px^{p-1}$$

and hence plugging in ζ_p,

$$\Phi'(\zeta_p)=\frac{p\zeta_p^{-1}}{\zeta_p-1}$$

From the minimal polynomials we compute (noting that $p-1$ is even so there is no sign):

$$N_{\mathbf{Q}(\zeta_p)}(\zeta_p) = 1, \qquad N_{\mathbf{Q}(\zeta_p)}(\zeta_p - 1) = p$$

Hence by Proposition 6.14, and noting that since p is odd, $(-1)^{\frac{(p-1)(p-2)}{2}} = (-1)^{\frac{p-1}{2}}$

$$\begin{aligned}\operatorname{disc}(\mathbf{Z}[\zeta_p]) &= (-1)^{\frac{p-1}{2}} N_{\mathbf{Q}(\zeta_p)}(\Phi_p'(\zeta_p)) \\ &= (-1)^{\frac{p-1}{2}} N_{\mathbf{Q}(\zeta_p)}(p) \frac{N_{\mathbf{Q}(\zeta_p)}(\zeta_p^{-1})}{N_{\mathbf{Q}(\zeta_p)}(\zeta_p - 1)} \\ &= (-1)^{\frac{p-1}{2}} p^{p-1} \frac{1}{p} = (-1)^{p-1} p^{p-2}\end{aligned}$$

It follows that if $q \mid [\mathcal{O}_{\mathbf{Q}(\zeta_p)} : \mathbf{Z}[\zeta_p]]$, $q = p$. But Proposition 18.9 and the fact that $\Phi_p(x+1)$ is Eisenstein for p imply that $p \nmid [\mathcal{O}_{\mathbf{Q}(\zeta_p)} : \mathbf{Z}[\zeta_p]]$. Hence $\mathcal{O}_{\mathbf{Q}(\zeta_p)} = \mathbf{Z}[\zeta_p]$ and $D_{\mathbf{Q}(\zeta_p)} = \operatorname{disc}(\mathbf{Z}[\zeta_p]) = (-1)^{\frac{p-1}{2}} p^{p-2}$. Moreover

$$(p) = (p, \zeta_p - 1)^{p-1}.$$

But $\zeta_p - 1 \mid N_{\mathbf{Q}(\zeta_p)}(\zeta_p - 1) = p$, so $(p, \zeta_p - 1) = (\zeta_p - 1)$. □

Exercise 20.2. Show that $\Phi_{p^r}(x) = \frac{x^{p^r}-1}{x^{p^{r-1}}-1}$ is irreducible, $\mathcal{O}_{\mathbf{Q}(\zeta_{p^r})} = \mathbf{Z}[\zeta_{p^r}]$ and compute $D_{\mathbf{Q}(\zeta_{p^r})}$.

Remark 20.3. It turns out that for any integer n, $\mathcal{O}_{\mathbf{Q}(\zeta_n)} = \mathbf{Z}[\zeta_n]$ but we won't prove this.

Now we describe the splitting of other primes in $\mathbf{Q}(\zeta_p)$.

Theorem 20.4. *Let $q \neq p$ be prime. Let f be the order of $\overline{q} \in (\mathbf{Z}/p\mathbf{Z})^\times$ and let $r = \frac{p-1}{f}$. Then*

$$(q) = \mathfrak{q}_1 \cdots \mathfrak{q}_r$$

where $\mathfrak{q}_1, \ldots, \mathfrak{q}_r$ are distinct primes of $\mathbf{Z}[\zeta_p]$, with $f(\mathfrak{q}_i) = f$ for $i = 1, \ldots, r$.

Proof. We first note that

$$p = -p(x^p - 1) + x(x^p - 1)'$$

and so $x^p - 1$ and hence Φ_p cannot have repeated factors mod q. Hence we have a factorization:

$$\overline{\Phi}_p(x) = g_1 \cdots g_s$$

where the irreducible factors $g_1, \ldots, g_s$ are distinct. We need to show $\deg(g_i) = f$.

We consider the finite field $\mathbf{F} = \mathbf{F}_q[x]/(g_i)$ which has cardinality $q^{\deg(g_i)}$. Hence for any nonzero element $a \in \mathbf{F}^\times$, $a^{q^{\deg(g_i)}-1} = 1$. On the other hand $\mathbf{F}$ contains the element $a = x + (g_i)$ which satisfies the polynomial g_i and hence $\overline{\Phi}_p$, so $a^p = 1$. It follows that $p \mid q^{\deg(g_i)} - 1$, i.e. $q^{\deg(g_i)} \equiv 1 \pmod{p}$. Hence $f \mid \deg(g_i)$.

On the other hand we observe that

$$\mathbf{F}' = \{b \in \mathbf{F} \mid b^{q^f} = b\}$$

is a subfield, and it contains a so it must be $\mathbf{F}$. It follows that $\mathbf{F}$ contains $\leq q^f$ elements. Hence $\deg(g_i) \leq f$. It follows that $\deg(g_i) = f$.

Thus we have shown that each irreducible factor g_i has degree f, and so there must be $r = \frac{p-1}{f}$ factors. The theorem now follows from Dedekind's factorization criteria. □

Remark 20.5. This theorem is the "gold standard" of having a simple description of the splitting behavior of all primes in a number field. We see that we can easily describe the splitting of an enormous prime q in $\mathbf{Q}(\zeta_p)$, since it only depends on $q \pmod p$.

By contrast for the moment we would have trouble understanding if an enormous prime q is split or inert in a quadratic field $\mathbf{Q}(\sqrt{d})$ (at least if you don't know quadratic reciprocity!)

Lemma 20.6. *We have* $\mathbf{Q}(\sqrt{(-1)^{\frac{p-1}{2}}p}) \subseteq \mathbf{Q}(\zeta_p)$.

Proof. The roots of $\Phi_p(x)$ are $\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1}$, and so the embeddings $\tau_1, \ldots, \tau_{p-1} : \mathbf{Q}(\zeta_p) \to \mathbf{C}$ are given by $\tau_i(\zeta_p) = \zeta_p^i$. We also have that $\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1}$ is an integral basis for $\mathbf{Z}[\zeta_p]$

Now by Proposition 6.11 we have

$$(-1)^{\frac{p-1}{2}}p^{p-2} = D_{\mathbf{Q}(\zeta_p)} = \det(\tau_i(\zeta_p^j))^2 = \det(\zeta_p^{ij})^2$$

Hence

$$\frac{\det(\zeta_p^{ij})}{p^{\frac{p-3}{2}}} = \pm\sqrt{(-1)^{\frac{p-1}{2}}p} \in \mathbf{Q}(\zeta_p)$$

□

There are a lot of fun things to say about this lemma. You might try the following exercise:

Exercise 20.7.

1. Show that if $K \subseteq \mathbf{Q}(\zeta_p)$ is a quadratic field, then the only prime that ramifies in K must be p. Deduce that K must be $\mathbf{Q}(\sqrt{(-1)^{\frac{p-1}{2}}p})$.
2. If you've taken Galois theory, use Galois theory to show that $\mathbf{Q}(\zeta_p)$ contains a unique quadratic subfield. You can then use the previous part to determine what this quadratic field is, giving an alternate proof of the lemma.
3. If you've taken elementary number theory you might have seen the quadratic Gauss sum:
$$g_p = \sum_{i=0}^{p-1} \zeta_p^{i^2} \in \mathbf{Z}[\zeta_p].$$
Show that $g_p^2 = (-1)^{\frac{p-1}{2}}p$, thus giving another proof of the lemma.

We now introduce some standard notation: we let $p^* = (-1)^{\frac{p-1}{2}}p$. Then we can rephrase lemma 20.6 as saying that $\sqrt{p^*} \in \mathbf{Q}(\zeta_p)$. We note that $p^* \equiv 1 \pmod 4$.

We now turn towards giving a proof of quadratic reciprocity. We note that

- q splits in the quadratic field $\mathbf{Q}(\sqrt{p^*})$ if and only if $\left(\frac{p^*}{q}\right) = 1$. This is a condition on $p^* \pmod q$.
- The splitting of q in $\mathbf{Q}(\zeta_p)$ is determined by $q \pmod p$ by Theorem 20.4.

Thus we might hope that by comparing the splitting of q in the two fields $\mathbf{Q}(\sqrt{p^*}) \subseteq \mathbf{Q}(\zeta_p)$ we will obtain some crazy relation between $p \pmod q$ and $q \pmod p$. This will turn out to be Gauss' law of quadratic reciprocity!

We begin with the following lemma, which as we will see is closely related to quadratic reciprocity:

Lemma 20.8. *Let f be the order of $q \pmod p$ and let $r = \frac{p-1}{f}$ be the number of prime factors of $q\mathbf{Z}[\zeta_p]$.*

1. *If q splits in $\mathbf{Q}(\sqrt{p^*})$ then r is even.*
2. *If $p \equiv 3 \pmod 4$ then the converse also holds: q splits in $\mathbf{Q}(\sqrt{p^*})$ if and only if r is even.*

Proof. We prove 1. We can write:

$$q\mathcal{O}_{p^*} = \mathfrak{q} \cdot \overline{\mathfrak{q}}$$

and hence:

$$q\mathbf{Z}[\zeta_p] = (\mathfrak{q} \cdot \overline{\mathfrak{q}}) \cdot \mathbf{Z}[\zeta_p] = (\mathfrak{q} \cdot \mathbf{Z}[\zeta_p])(\overline{\mathfrak{q}} \cdot \mathbf{Z}[\zeta_p])$$

We would like to argue that $\mathfrak{q} \cdot \mathbf{Z}[\zeta_p]$ and $\overline{\mathfrak{q}} \cdot \mathbf{Z}[\zeta_p]$ have the same number of prime factors, and hence that the number r of prime factors of $q\mathbf{Z}[\zeta_p]$ is even. Intuitively this must be the case because there is no way to tell $\mathfrak{q}, \overline{\mathfrak{q}}$ apart!

To prove this formally, we start the embedding $\tau_0 : \mathbf{Q}(\sqrt{p^*}) \to \mathbf{C}$ defined by $\tau_0(x) = \overline{x}$. By Corollary 4.7 there is an embedding $\tau : \mathbf{Q}(\zeta_p) \to \mathbf{C}$ whose restriction to $\mathbf{Q}(\sqrt{p^*})$ is τ_0. As we saw in the course of proving Lemma 20.6, we must have $\tau(\zeta) = \zeta^i$ for $1 \leq i \leq p-1$, and so in particular $\tau : \mathbf{Q}(\zeta) \to \mathbf{Q}(\zeta)$ is an automorphism (its inverse is given by $\tau^{-1}(\zeta) = \zeta^j$ where $ij \equiv 1 \pmod p$). Moreover τ, τ^{-1} preserve algebraic integers, so $\tau : \mathbf{Z}[\zeta] \to \mathbf{Z}[\zeta]$ is also an automorphism. But then we have

$$\tau(\mathfrak{q} \cdot \mathbf{Z}[\zeta_p]) = \tau(\mathfrak{q}) \cdot \mathbf{Z}[\zeta_p] = \tau_0(\mathfrak{q}) \cdot \mathbf{Z}[\zeta_p] = \overline{\mathfrak{q}} \cdot \mathbf{Z}[\zeta_p].$$

Because τ is an automorphism of $\mathbf{Z}[\zeta_p]$, $\mathfrak{q} \cdot \mathbf{Z}[\zeta_p]$ and $\overline{\mathfrak{q}} \cdot \mathbf{Z}[\zeta_p]$ have the same number of prime factors.

We now prove 2. We will first prove (without any condition on p) that if q is inert in $\mathbf{Q}(\sqrt{p^*})$ then f is even. Fix a prime $\mathfrak{q}$ of $\mathbf{Z}[\zeta_p]$ lying over q. We then have inclusions of fields:

$$\mathbf{F}_q = \mathbf{Z}/q\mathbf{Z} \subseteq \mathcal{O}_{p^*}/q\mathcal{O}_{p^*} \subseteq \mathbf{Z}[\zeta_p]/\mathfrak{q}.$$

By the tower law for degrees of field extensions (Lemma 4.5) we have that

$$2 = [\mathcal{O}_{p^*}/q\mathcal{O}_{p^*} : \mathbf{F}_q] \mid [\mathbf{Z}[\zeta_p]/\mathfrak{q} : \mathbf{F}_q] = f(\mathfrak{q}) = f$$

Now assuming that $p \equiv 3 \pmod 4$ we see that $p-1 = 2k$ with k odd. Since $rf = p-1$ we conclude that if f is even then r is odd. Thus we have shown that if $p \equiv 3 \pmod 4$ then: q is inert in $\mathbf{Q}(\sqrt{p^*})$ implies r is odd, and hence the contrapositive, r even implies q splits in $\mathbf{Q}(\sqrt{p^*})$. □

Remark 20.9. We will see in the next lecture that the converse also holds when $p \equiv 1 \pmod 4$.

21 Lecture 21: Finishing quadratic reciprocity, starting finiteness of class groups

Recall that for an odd prime p, the Legendre symbol is defined by:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue mod } p \\ -1 & \text{if } a \text{ is a quadratic nonresidue mod } p \\ 0 & \text{if } p \mid a \end{cases}$$

We begin by proving some of its elementary properties:

Lemma 21.1. *Let p be an odd prime. Then we have:*

1. *For $a, b \in \mathbf{Z}$,*
$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$
2. $\left(\frac{a}{p}\right) = 1$ *if and only if* $\frac{p-1}{f}$ *is even, where f is the order of $a \in (\mathbf{Z}/p\mathbf{Z})^\times$*
3. $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

Proof. The key point for all of these properties is that $(\mathbf{Z}/p\mathbf{Z})^\times \simeq \mathbf{Z}/(p-1)\mathbf{Z}$ is a cyclic group of order $p-1$. In particular this implies that the squares $((\mathbf{Z}/p\mathbf{Z})^\times)^2 \subset (\mathbf{Z}/p\mathbf{Z})^\times$ form a subgroup of index 2. It follows that there is a homomorphism

$$(\mathbf{Z}/p\mathbf{Z})^\times \to \{\pm 1\}$$

sending $((\mathbf{Z}/p\mathbf{Z})^\times)^2$ to 1 and the non trivial coset of $((\mathbf{Z}/p\mathbf{Z})^\times)^2$ to -1, and this homomorphism is exactly the Legendre symbol, by definition. This proves 1.

To prove 2. we recall from group theory that the order of $\overline{r} \in \mathbf{Z}/(p-1)\mathbf{Z}$ is $f = \frac{p-1}{\gcd(r,p-1)}$. In particular $\frac{p-1}{f} = \gcd(r, p-1)$ is even if and only if r is.

Point 3 is an application of point 2: the order of -1 is 2, so -1 is a quadratic residue mod p if and only if $\frac{p-1}{2}$ is even, i.e. if and only if $p \equiv 1 \pmod 4$. □

We now state the law of quadratic reciprocity:

Theorem 21.2 (Gauss' Law of quadratic reciprocity)**.** *Let p, q be distinct odd primes. Then*

$$\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right).$$

Equivalently

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Or equivalently again, $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ *unless* $p \equiv q \equiv 3 \pmod 4$*, in which case* $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$

To see why these statements are equivalent, note that

$$\left(\frac{p^*}{q}\right) = \left(\frac{(-1)^{\frac{p-1}{2}}p}{q}\right) = \left(\frac{(-1)^{\frac{p-1}{2}}}{q}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)$$

where we have used the definition of p^* in the first equality, the multiplicativity of the Legendre symbol in the second equality, and the rule for $\left(\frac{-1}{q}\right)$ in the third equality. In other words, $\left(\frac{p^*}{q}\right) = \left(\frac{p}{q}\right)$ unless $p, q \equiv 3 \pmod 4$ in which case $\left(\frac{p^*}{q}\right) = -\left(\frac{p}{q}\right)$

In order to prove quadratic reciprocity, we recall the setup of the lemma of last time. There we are considering the splitting of a prime q in the two fields $\mathbf{Q}(\sqrt{p^*}) \subseteq \mathbf{Q}(\zeta_p)$. What we know is that:

- q splits in $\mathbf{Q}(\sqrt{p^*})$ if and only if $\left(\frac{p^*}{q}\right) = 1$.

- q splits into $r = \frac{p-1}{f}$ primes in $\mathbf{Q}(\zeta_p)$, where f is the order of $\overline{q} \in (\mathbf{Z}/p\mathbf{Z})^\times$. In particular r is even if and only if $\left(\frac{q}{p}\right) = 1$ by part 2 of Lemma 21.1.

Making these translations, Lemma 20.8 becomes the statement:

1. $\left(\frac{p^*}{q}\right) = 1$ implies $\left(\frac{q}{p}\right) = 1$.
2. If $p \equiv 3 \pmod 4$ then $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right)$.

Proof of quadratic reciprocity. Proof when p, q are not both 1 (mod 4): after possibly swapping p and q we may assume $p \equiv 3 \pmod 4$. Then point 2 of the lemma gives us quadratic reciprocity in this case.

Proof when p, q are not both 3 (mod 4): in this case point 1 of the lemma gives us that $\left(\frac{p}{q}\right) = 1$ implies $\left(\frac{q}{p}\right) = 1$, but swapping p and q and applying the lemma again we get $\left(\frac{q}{p}\right) = 1$ implies $\left(\frac{p}{q}\right) = 1$. So $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$. □

Remark 21.3. Quadratic reciprocity is an extremely deep and difficult theorem and it is the cornerstone of elementary number theory. Euler discovered special cases of it but was unable to prove them. Legendre formulated the statement but was unable to give a correct proof of it. Finally Gauss managed to give the first proof around 1800, and proceeded to give several more proofs.

While there are many "elementary" proofs of quadratic reciprocity, they tend to be completely unmotivated and unenlightening. The proof we have given here is far from the easiest or most elementary, but it has the advantage of having a clear idea behind it: we analyzed how a prime q splits in the fields $\mathbf{Q}(\sqrt{p^*}) \subseteq \mathbf{Q}(\zeta_p)$. If you asked a random number theorist on the street for a proof of quadratic reciprocity you would almost certainly get some variation of this proof, unless they have just taught a course on elementary number theory!

To complete our discussion of quadratic reciprocity we should also give the formula for $\left(\frac{2}{p}\right)$. In fact:

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod 8 \\ -1 & \text{if } p \equiv \pm 3 \pmod 8 \end{cases}$$

We can prove this using a similar strategy but now using the cyclotomic field $\mathbf{Q}(\zeta_8)$. The proof will be given in the following exercise:

Exercise 21.4. 1. Prove that $\Phi_8(x) = \frac{x^8-1}{x^4-1} = x^4 + 1$ is irreducible and hence is the minimal polynomial of $\zeta_8 = e^{2\pi i/8}$. Show that $\mathcal{O}_{\mathbf{Q}(\zeta_8)} = \mathbf{Z}[\zeta_8]$.

2. Show that if $p \equiv 1 \pmod 8$ then $p\mathbf{Z}[\zeta_8]$ is the product of four distinct primes. (Hint: you just need to show that $x^8 - 1$ has 8 roots in $\mathbf{F}_p$). If you'd like you can also try to describe the factorization of the other p but we won't need this (the answer is that for $p \equiv 3, 5, 7 \pmod 8$, $p\mathbf{Z}[\zeta_8] = \mathfrak{p}_1\mathfrak{p}_2$ where each $\mathfrak{p}_i$ has inertial degree 2.)

3. Show that $\sqrt{2}, \sqrt{-2}, \sqrt{-1} \in \mathbf{Q}(\zeta_8)$.

4. Deduce from the previous two parts that if $p \equiv 1 \pmod 8$ then p splits in $\mathcal{O}_2$, $\mathcal{O}_{-2}$, and $\mathcal{O}_{-1}$, and in particular $\left(\frac{2}{p}\right) = 1$.

5. Show that if $p \equiv \pm 3 \pmod 8$ then p is inert in $\mathcal{O}_2$, and hence $\left(\frac{2}{p}\right) = -1$. (Hint: use that $\mathbf{Z}[\sqrt{2}]$ is a PID and norms to reduce this to showing that $x^2 - 2y^2 = \pm p$ has no solutions, and then show this has no solutions mod 8.)

6. The remaining case is $p \equiv -1 \pmod 8$. Do this by writing $\left(\frac{2}{p}\right) = \left(\frac{-2}{p}\right)\left(\frac{-1}{p}\right)$ and show that both of the factors are -1.

For our purposes, our main interest in quadratic reciprocity is that it helps us understanding how primes split in quadratic fields. Consider the following examples:

Example 21.5. We know that an odd prime p splits in $\mathcal{O}_5$ if $\left(\frac{5}{p}\right) = 1$. But by quadratic reciprocity $\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right)$. Since the quadratic residues mod 5 are $1, 4$ and non residues are $2, 3$, we conclude that an odd prime p splits in $\mathcal{O}_5$ if $p \equiv 1, 4 \pmod 5$, and is inert if $p \equiv 2, 3 \pmod 5$.

For instance 2011 is prime. It would take you quite a while (by hand at least!) to solve the congruence $a^2 \equiv 5 \pmod{2011}$. But since $2011 \equiv 1 \pmod 5$, quadratic reciprocity implies it has a solution, and so 2011 splits in $\mathcal{O}_5$. Note however that quadratic reciprocity just tells us that a factorization of 2011 exists but it doesn't tell us how to find the factors!

Example 21.6. We consider another example. When does an odd prime p split in $\mathcal{O}_3$? Again it is if $\left(\frac{3}{p}\right) = 1$, but how can we write this in a more elementary way using quadratic reciprocity? We can write:

$$\left(\frac{3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{3}\right).$$

So in order to have $\left(\frac{3}{p}\right) = 1$ we either have:

- $\left(\frac{-1}{p}\right) = \left(\frac{p}{3}\right) = 1$, i.e. $p \equiv 1 \pmod 4$ and $p \equiv 1 \pmod 3$, or in other words $p \equiv 1 \pmod{12}$.
- $\left(\frac{-1}{p}\right) = \left(\frac{p}{3}\right) = -1$, i.e. $p \equiv -1 \pmod 4$ and $p \equiv -1 \pmod 3$, or in other words $p \equiv -1 \pmod{12}$.

In summary, p splits in $\mathcal{O}_3$ if and only if $p \equiv \pm 1 \pmod{12}$. (Note that $12 = D_{\mathbf{Q}(\sqrt{3})}$ and this is not a coincidence!)

21.1 Towards finiteness of the class group

Our next goal is to develop some tools to compute class groups of number fields. In particular we will prove the following theorem:

Theorem 21.7. *Let K be a number field. Then $\mathrm{Cl}(\mathcal{O}_K)$ is finite.*

We start out with a lemma:

Lemma 21.8. *For any $C \in \mathbf{R}$, the number of ideals $I \subseteq \mathcal{O}_K$ with $\|I\| \leq C$ is finite.*[58]

[58]It is quite an interesting problem to estimate the number $\#\{I \subseteq \mathcal{O}_K \mid \|I\| \leq C\}$ as a function of C. This is connected to the "class number formula" which is slightly beyond the scope of this course.

Proof. Consider the factorization $I = \mathfrak{p}_1 \cdots \mathfrak{p}_r$ into primes. We will bound both the number r of prime factors as well as the number of possible prime factors. Since $\|p_i\| \geq 2$ see see that $2^r \leq \|I\| \leq C$, so $r \leq \log_2(C)$.

Now if $\mathfrak{p}$ is a prime occurring in the factorization of I, we must have $\|\mathfrak{p}\| \leq C$. Suppose $\mathfrak{p}$ lies over the rational prime p. Then $C \geq \|\mathfrak{p}\| = p^{f(\mathfrak{p})} \geq p$. In particular there are only finitely many possibilities for p, and there are only finitely many primes $\mathfrak{p}$ lying above each p. □

Exercise 21.9. Show that the number of subgroups $M \subseteq \mathbf{Z}^n$ with $[\mathbf{Z}^n : M] \leq C$ is finite. This gives another approach to proving the lemma, which doesn't use anything about ideals.

22 Lecture 22: Minkowski's theorem

We now outline our strategy for proving Theorem 21.7. We prove the following lemma:

Lemma 22.1. *Let K be a number field. Suppose there is some constant C such that:*

- *For any nonzero ideal $I \subseteq \mathcal{O}_K$, there exists $0 \neq \alpha \in I$, with $|N_K(\alpha)| < C\|I\|$.*

Then:

- *For all classes $[I] \in \mathrm{Cl}(\mathcal{O}_K)$ there exists an ideal I' with $[I'] = [I]$ and $\|I'\| < C$.*

Proof. Pick an ideal J in the class $[I]^{-1}$. Then there exists $0 \neq \alpha \in J$ with $|N_K(\alpha)| \leq C\|J\|$. Then by Proposition 11.9 we have an ideal I' with $JI' = (\alpha)$. In the class group we have $[I'] = [J]^{-1} = [I]$, while taking norms we have

$$\|J\| \cdot \|I'\| = \|(\alpha)\| = |N_K(\alpha)|$$

and hence

$$\|I'\| = \frac{|N_K(\alpha)|}{\|J\|} \leq C.$$

□

We recall from example sheet 1 that an embedding $\tau : K \to \mathbf{C}$ called *real* if $\tau(K) \subseteq \mathbf{R}$ and is called *complex* otherwise. The complex embeddings come in pairs $\tau, \overline{\tau}$, and we denote by r the number of real embeddings and s the number of pairs of complex embeddings, so that $n = [K : \mathbf{Q}] = r + 2s$.

The main theorem we will prove is now:

Theorem 22.2. *Let K be a number field and let*

$$C_K = \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|D_K|}$$

where $n = [K : \mathbf{Q}]$ is the degree of K and s is the number of pairs of complex embeddings. Then for any nonzero ideal $I \subseteq \mathcal{O}_K$ there exists $0 \neq \alpha \in I$ with $|N_K(\alpha)| \leq C_K\|I\|$, and consequently by Lemma 22.1, any ideal class is represented by an ideal I with $\|I\| \leq C_K$

Example 22.3. We are mostly interested in applying Minkowski's bound when $K = \mathbf{Q}(\sqrt{d})$ is a quadratic field so $n = 2$. In the imaginary quadratic case $d < 0$, we have $(r, s) = (0, 1)$ and so Minkowski's constant is $C_K = \frac{2}{\pi}\sqrt{|D_K|}$, while in the real quadratic case $d > 0$ we have $(r, s) = (2, 0)$ and so Minkowski's constant is $C_K = \frac{1}{2}\sqrt{|D_K|}$.

We illustrate Minkowski's bound with some examples.

Example 22.4. Let's at long last compute the class group of $\mathcal{O}_{-5}$. We have $D_{\mathbf{Q}(\sqrt{-5})} = -20$ so the Minkowski bound is $C_K = \frac{2}{\pi}\sqrt{20} < 3$. Thus any ideal class is represented by an ideal of norm 1 or 2. Of course the only ideal of norm 1 is $\mathcal{O}_K$, while since $(2) = (2, 1 + \sqrt{-5})^2$, the only ideal of norm 2 is $\mathfrak{p}_2 = (2, 1 + \sqrt{-5})$. Moreover this ideal is not principal, because $x^2+5y^2 = 2$ has no solutions, and hence $\mathrm{Cl}(\mathcal{O}_5) = \{1, [\mathfrak{p}_2]\} \simeq \mathbf{Z}/2\mathbf{Z}$.

Let's also now prove the following theorem, which we mentioned in lecture 10:

Theorem 22.5. *The rings $\mathcal{O}_{-19}, \mathcal{O}_{-43}, \mathcal{O}_{-67}, \mathcal{O}_{-163}$ are PIDs.*

Proof. We do $\mathcal{O}_{-163}$, and leave the others as an exercise. The Minkowski constant for $\mathbf{Q}(\sqrt{-163})$ is $\frac{2}{\pi}\sqrt{163} < 9$. We need to look for non principal ideals of norm less than 9. Since their prime factors would also have norm less than 9, we just need to factor the prime numbers less than 9, i.e. 2,3,5,7.

2 is inert because $-163 \equiv 5 \pmod 8$. 3 is inert because $-163 \equiv 2 \pmod 3$. 5 is inert because $-163 \equiv 2 \pmod 5$. 7 is inert because $-163 \equiv 5 \pmod 7$. □

Exercise 22.6. Show that $\mathcal{O}_{-19}, \mathcal{O}_{-43}, \mathcal{O}_{-67}$ are PIDs ($C_K < 3, < 5, < 6$ in each case, to save you the trouble.)

22.1 Lattices and Minkowski's theorem

For the moment we completely forget about algebraic integers and ideals. x

Definition 22.7. A subgroup $\Lambda \subset \mathbf{R}^n$ is called a lattice if it has a basis $v_1, \ldots, v_n$ as an abelian group which is also a basis of $\mathbf{R}^n$ as an $\mathbf{R}$-vector space.

Example 22.8. $\mathbf{Z}^n \subset \mathbf{R}^n$ is a lattice. So is $\mathcal{O}_d \subset \mathbf{C} \simeq \mathbf{R}^2$ for $d < 0$.

Definition 22.9. Let $\Lambda \subset \mathbf{R}^n$ be a lattice with basis $v_1, \ldots, v_n$. Then the *fundamental parallelotope* is

$$P(v_1, \ldots, v_n) = \{c_1 v_1 + \cdots + c_n v_n \mid 0 \leq c_i < 1\}.$$

It is a *fundamental domain* for the lattice in the sense that

$$\mathbf{R}^n = \cup_{v \in \Lambda} v + P(v_1, \ldots, v_n)$$

and the sets in the union are pairwise disjoint. We also write

$$\mathbf{R}^n = \coprod_{v \in \Lambda} v + P(v_1, \ldots, v_n)$$

to indicate that the union is disjoint. In other words we can subdivide $\mathbf{R}^n$ into translates of the fundamental parallelotope.

By linear algebra, we have[59]

$$\mathrm{vol}(P(v_1, \ldots, v_n)) = |\det(A)|$$

where A is the matrix whose columns are $v_1, \ldots, v_n$. We prove that this number is in fact independent of the basis:

Lemma 22.10. *If $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ are two bases for a lattice $\Lambda \subset \mathbf{R}^n$ then*

$$\mathrm{vol}(P(v_1, \ldots, v_n)) = \mathrm{vol}(P(w_1, \ldots, w_n)).$$

Proof. Let A be the matrix with columns $v_1, \ldots, v_n$ and let B be the matrix with columns $w_1, \ldots, w_n$. We can write $B = SA$ for a change of basis matrix S. As the w_i are integer linear combinations of the v_i, $S \in M_n(\mathbf{Z})$. But as the v_i are integer linear combinations of the w_i, $S^{-1} \in M_n(\mathbf{Z})$, and hence $S \in M_n(\mathbf{Z})^\times$ has $\det(S) = \pm 1$. Hence $|\det(A)| = |\det(B)|$. □

Remark 22.11. More generally, if $D_1, D_2 \subset \mathbf{R}^n$ are two (measurable) fundamental domains for Λ then $\mathrm{vol}(D_1) = \mathrm{vol}(D_2)$, and the lemma is a special case of this.

To prove this, since D_1 is a fundamental domain

$$D_2 = \coprod_{v \in \Lambda} (D_1 + v) \cap D_2,$$

and since D_2 is a fundamental domain,

$$D_1 = \coprod_{v \in \Lambda} (D_2 + v) \cap D_1.$$

Hence

$$\begin{aligned} \mathrm{vol}(D_1) &= \sum_{v \in \Lambda} \mathrm{vol}((D_2 + v) \cap D_1) \\ &= \sum_{v \in \Lambda} \mathrm{vol}(D_2 \cap (D_1 - v)) \\ &= \sum_{v \in \Lambda} \mathrm{vol}(D_2 \cap (D_1 + v)) \\ &= \mathrm{vol}(D_2) \end{aligned}$$

In second equality we use translation invariance of volume, and in the third equality we just reindexed the sum replacing v with $-v$. The first and last equalities use countable additivity of volume.

Definition 22.12. For a lattice $\Lambda \subseteq \mathbf{R}^n$, we define the covolume of Λ to be

$$\mathrm{covol}(\Lambda) = \mathrm{vol}(P(v_1, \ldots, v_n))$$

for any basis $v_1, \ldots, v_n$.

[59] If you have taken measure theory, volume here can be taken to mean Lebesgue measure. If not, then you can use the multivariable calculus definition of volume. All the subsets $S \subset \mathbf{R}^n$ that we will take the volume of will be nice enough that the function 1_S which is 1 on S and 0 otherwise will be Riemann integrable, and so there will be no question of what the volume means.

More generally by the remark above we could have defined it as the volume of any fundamental domain for Λ. Alternatively we could have defined it as $\mathrm{vol}(\mathbf{R}^n/\Lambda)$, if this is something that makes sense to you![60]

Example 22.13. The covolume of $\mathbf{Z}^n \subseteq \mathbf{R}^n$ is the volume of the hypercube $[0,1)^n$, namely 1.

Lemma 22.14 (Blichfeldt's Lemma)**.** *Let $\Lambda \subset \mathbf{R}^n$ be a lattice and let $S \subseteq \mathbf{R}^n$ be a measurable subset. Then if $\mathrm{vol}(S) > \mathrm{covol}(\Lambda)$, there exists $x, y \in S$ with $0 \neq x - y \in \Lambda$.*

Proof. We explain the proof in two ways, with the first being more intuitive by requiring more work to make rigorous. If $S \to \mathbf{R}^n/\Lambda$ is injective then $\mathrm{vol}(S) \leq \mathrm{covol}(\Lambda)$. Thus by the hypothesis it is not injective, and so there exists $x, y \in S$ with $x \neq y$ and $x - y \in \Lambda$.[61]

Alternatively fix a basis $v_1, \ldots, v_n$ for Λ and let $P = P(v_1, \ldots, v_n)$ be the corresponding fundamental parallelotope. We have

$$S = \coprod_{v \in \Lambda} (v + P) \cap S$$

and hence

$$\mathrm{vol}(S) = \sum_{v \in \Lambda} \mathrm{vol}((v + P) \cap S) = \sum_{v \in \Lambda} \mathrm{vol}(P \cap (S - v))$$

In the first equality we used countable additivity of volume, and in the second we used translation invariance of volume.

Now if the sets $P \cap (S - v)$ for $v \in \Lambda$ were pairwise disjoint, since they are all contained P we would have

$$\mathrm{vol}(S) = \sum_{v \in \Lambda} \mathrm{vol}(P \cap (S - v)) = \mathrm{vol}(\bigcup_{v \in \Lambda} \mathrm{vol}(P \cap (S - v))) \leq \mathrm{vol}(P).$$

Since we assumed that $\mathrm{vol}(S) > \mathrm{covol}(\Lambda)$, we must have $v \neq w \in \Lambda$ such that

$$\emptyset \neq (P \cap (S - v)) \cap (P \cap (S - w)) \subseteq (S - v) \cap (S - w)$$

In other words, there exists a point $a \in (S - v) \cap (S - w)$, which is to say that $x = a + v \in S$, $y = a + w \in S$, and $0 \neq x - y = v - w \in \Lambda$.[62] □

As a consequence we deduce:

Theorem 22.15 (Minkowski's convex body theorem)**.** *Let $\Lambda \subset \mathbf{R}^n$ be a lattice and let $S \subseteq \mathbf{R}^n$ be a measurable subset which is:*

1. *Convex: $x, y \in S$ implies $tx + (1 - t)y \in S$ for $t \in [0, 1]$.*

[60] $\mathbf{R}^n/\Lambda$ is a torus. From this point of view, we see that we should think of the covolume as something like the index $[\mathbf{R}^n : \Lambda]$, except that as $\mathbf{R}^n/\Lambda$ is an uncountable set, we can better measure its size by taking volume.

[61] Blichfeldt's lemma may be thought of as a version of the pigeonhole principle but for volume: instead of saying that a map $X \to Y$ with $\#X > \#Y$ cannot be injective we are saying that a map $X \to Y$ with $\mathrm{vol}(X) > \mathrm{vol}(Y)$ cannot be injective.

[62] See the wikipedia page for a nice illustration of this proof: we are just decomposing the set S into its intersections with all the translates of P, $(v + P) \cap S$, and then considering their translations back to P: $P \cap (S - v)$.

2. Symmetric about 0: $x \in S$ implies $-x \in S$.

Then if $\mathrm{vol}(S) > 2^n \mathrm{covol}(\Lambda)$ *there exists* $0 \neq x \in S \cap \Lambda$.

Proof. Apply Blichfeldt's lemma to $\frac{1}{2}S$. Then as $\mathrm{vol}(\frac{1}{2}S) = \frac{1}{2^n}\mathrm{vol}(S) > \mathrm{covol}(\Lambda)$, there exists $x, y \in \frac{1}{2}S$ with $0 \neq x - y \in \Lambda$.

But then $2x, 2y \in S$. Hence $-2y \in S$ by central symmetry, and hence $\frac{1}{2}(2x) + \frac{1}{2}(-2y) \in S$ by convexity. But $\frac{1}{2}(2x) + \frac{1}{2}(-2y) = x - y \in \Lambda$. □

Minkowski's theorem may be used to give some truly magical but (in my view) not very intuitive proofs of some results in elementary number theory. You might try the following exercise for an example.

Exercise 22.16. In this exercise you will prove Lagrange's theorem that any prime number p can be expressed as a sum of four squares. The key is always to come up with a clever choice of Λ and S...

1. Show that there are integers $a, b \in \mathbf{Z}$ with $a^2 + b^2 + 1 \equiv 0 \pmod{p}$ (Hint: use the pigeonhole principal.)
2. Consider the lattice $\Lambda \subseteq \mathbf{R}^4$ with basis
$$\begin{pmatrix} p \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ p \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} a \\ b \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -b \\ a \\ 0 \\ 1 \end{pmatrix}.$$
Show that $\mathrm{covol}(\Lambda) = p^2$ and for any $v \in \Lambda$, $p \mid \|v\|^2$ (Hint: show that the dot product of any two of the basis vectors is a multiple of p.)
3. Consider $S = \{v \in \mathbf{R}^4 \mid \|v\|^2 < 2p\}$. Compute its volume and conclude using Minkowski's theorem (the volume a ball of radius r in 4 dimensions is $\frac{1}{2}\pi^2 r^4$).

23 Lecture 23: Proving the Minkowski bound

We begin with a lemma.

Lemma 23.1. *Let $\Lambda \subset \mathbf{R}^n$ be a lattice and let $\Lambda' \subseteq \Lambda$ be a subgroup of finite index. Then Λ' is also a lattice and*
$$\mathrm{covol}(\Lambda') = [\Lambda : \Lambda'] \cdot \mathrm{covol}(\Lambda).$$

Proof. We have
$$[\Lambda : \Lambda']\Lambda \subseteq \Lambda' \subseteq \Lambda.$$
As Λ' is sandwiched between free abelian groups of rank n, it must itself be a free abelian group of rank n, by Proposition 5.4.

Now let $v_1, \ldots, v_n$ be a basis for Λ, and let $w_1, \ldots, w_n$ be a basis for Λ'. If A is the matrix with $w_1, \ldots, w_n$ as columns, and B is the matrix with $v_1, \ldots, v_n$ as columns, then we have
$$A = SB$$
for $S \in M_n(\mathbf{Z})$. Taking absolute values of determinants we obtain
$$\mathrm{covol}(\Lambda') = |\det(S)| \cdot \mathrm{covol}(\Lambda).$$
But by Lemma 15.2, we have $|\det(S)| = [\Lambda : \Lambda']$. □

Remark 23.2. If $L \subset M \subset N$ are groups we have the formula

$$[N : L] = [M : L][N : M].$$

We should think of Lemma 23.1 as a variant of this, where instead of considering $[\mathbf{R}^n : \Lambda] = \#\mathbf{R}^n/\Lambda$ which is uncountable, we consider the $\mathrm{covol}(\Lambda) = \mathrm{vol}(\mathbf{R}^n/\Lambda)$.

In fact, we indicate an alternative proof of 23.1 without using the structure theory of finitely generated abelian groups (and we can even go back and reprove Lemma 15.2 and Proposition 6.8 without using the structure theory.)

Pick coset representatives

$$\Lambda = \coprod_{i=1}^{[\Lambda:\Lambda']} v_i + \Lambda'$$

and pick a (measurable) fundamental domain P for Λ (e.g. it could be a fundamental paralellotope.) Then we can check that

$$P' = \coprod v_i + P$$

is a fundamental domain for Λ', and hence

$$\mathrm{covol}(\Lambda') = \mathrm{vol}(P') = [\Lambda : \Lambda']\mathrm{vol}(P) = [\Lambda : \Lambda']\mathrm{covol}(\Lambda)$$

Note that this proof is now exactly like the proof of Lagrange's theorem, except that we have replaced counting with taking volume.

23.1 Minkowski for Imaginary quadratic fields

Fix $d < 0$ squarefree and let $K = \mathbf{Q}(\sqrt{d})$. We fix the bijection $C \simeq \mathbf{R}^2$ sending $z \in \mathbf{C}$ to $(\mathrm{Re}(z), \mathrm{Im}(z))$.

Lemma 23.3. *$\mathcal{O}_d \subseteq \mathbf{C} \simeq \mathbf{R}^2$ is a lattice with covolume $\frac{|D_K|}{2}$.*

Proof. $\mathbf{Z}[\sqrt{d}] \subseteq \mathbf{C} \simeq \mathbf{R}^2$ is clearly a lattice with basis $(1, 0)$, $(0, \sqrt{|d|})$, hence with covolume $\sqrt{|d|}$.

Now if $d \not\equiv 1 \pmod 4$, $O_d = \mathbf{Z}[\sqrt{d}]$ and $D_K = 4d$ and so the covolume is $\frac{\sqrt{|D_K|}}{2}$.

On the other hand if $d \equiv 1 \pmod 4$ then $[O_d : \mathbf{Z}[\sqrt{d}]] = 2$ and so by Lemma 23.1, $\mathrm{covol}(\mathcal{O}_K) = \frac{\sqrt{|d|}}{2} = \frac{\sqrt{|D_K}}{2}$ as $D_K = d$ in this case. □

Corollary 23.4. *For any nonzero ideal $I \subseteq \mathcal{O}_d$, $I \subseteq \mathbf{C} \simeq \mathbf{R}^2$ is a lattice with covolume $\frac{\sqrt{|D_K|}}{2}\|I\|$.*

Proof. This follows from Lemma 23.1 and the fact that $\|I\| = [\mathcal{O}_K : I]$ by definition. □

Theorem 23.5 (Minkowski's bound for IQF)**.** *For any nonzero ideal $I \subset \mathcal{O}_d$, $d < 0$ squarefree, there exists $0 \neq \alpha \in I$ with*

$$N_K(\alpha) \leq \frac{2}{\pi}\sqrt{|D_K|}\|I\|.$$

Proof. Let

$$S_R = \{(x, y) \mid x^2 + y^2 \leq R\} \subset \mathbf{R}^2.$$

This is a circle of radius $\sqrt{R}$ with $\mathrm{vol}(S_R) = \pi R$. It is convex and centrally symmetric.

Moreover we have

$$\mathrm{vol}(S_R) > 2^2 \mathrm{covol}(I)$$

if and only if

$$\pi R > 4\frac{\sqrt{|D_K|}}{2}\|I\|$$

i.e. if and only if

$$R > \frac{2}{\pi}\sqrt{|D_K|}\|I\|.$$

If this inequality holds, Minkowski's theorem applies and produces $0 \neq \alpha \in I \cap S_R$. But $\alpha \in S_R$ implies

$$N_K(\alpha) = \mathrm{Re}(\alpha)^2 + \mathrm{Im}(\alpha)^2 \leq R.$$

Hence if we pick R with $\lfloor R \rfloor = \lfloor \frac{2}{\pi}\sqrt{|D_K|}\|I\| \rfloor$, then since $N_K(\alpha) \in \mathbf{Z}$, $N_K(\alpha) \leq R$ implies $N_K(\alpha) \leq \frac{2}{\pi}\sqrt{|D_K|}\|I\|$.[63] □

23.2 Minkowski for real quadratic fields

Now let $d > 1$ be squarefree and let $K = \mathbf{Q}(\sqrt{d})$. We recall that $\mathcal{O}_d \subseteq \mathbf{R}$ is dense and not a lattice! It is better to consider

$$\begin{aligned} \iota : \mathcal{O}_d &\to \mathbf{R}^2 \\ \alpha &\mapsto (\tau_1(\alpha), \tau_2(\alpha)) \end{aligned}$$

where $\tau_1(\alpha) = \alpha$ and $\tau_2(\alpha) = \overline{\alpha}$ are the two embeddings.

Lemma 23.6. *1. $\iota(\mathcal{O}_d) \subset \mathbf{R}^2$ is a lattice with covolume $\sqrt{D_K}$.*

2. For any nonzero ideal $I \subset \mathcal{O}_d$, $\iota(I) \subseteq \mathbf{R}^2$ is a lattice with covolume $\sqrt{D_K}\|I\|$.

Proof. Let α_1, α_2 be an integral basis. Then $\iota(\mathcal{O}_d)$ is spanned by the vectors $\begin{pmatrix}\tau_1(\alpha_1)\\ \tau_2(\alpha_1)\end{pmatrix}, \begin{pmatrix}\tau_1(\alpha_2)\\ \tau_2(\alpha_2)\end{pmatrix}$, and so

$$\mathrm{covol}(\iota(\mathcal{O}_d)) = \left|\det \begin{pmatrix}\tau_1(\alpha_1) & \tau_1(\alpha_2)\\ \tau_2(\alpha_1) & \tau_2(\alpha_2)\end{pmatrix}\right| = \sqrt{|D_K|}$$

by Proposition 6.11.

Part 2 follows from Lemma 23.1. □

Now there is a big difference from the imaginary case: the norm is

$$|N_K(\alpha)| = \tau_1(\alpha)\tau_2(\alpha).$$

So it is natural to consider the following region in $\mathbf{R}^2$:

$$M_R = \{(x, y) \in \mathbf{R}^2 \mid |xy| \leq R\}.$$

[63] Here $\lfloor x \rfloor$ denotes the largest integer $\leq x$. For any real number x we have $\lfloor x \rfloor = \lfloor x + \varepsilon \rfloor$ for $\varepsilon > 0$ sufficiently small.

This is the region bounded by two hyperbolas. The good news is that it has infinite volume. The bad news is that it is not convex!

So now the idea is to choose some $S_R \subseteq M_R$ which is convex and centrally symmetric, and ideally with the largest possible volume. We take

$$S_R = \{(x, y) \in \mathbf{R}^2 \mid |x| + |y| \leq 2\sqrt{R}\}.$$

It is easy to convince yourself that $S_R \subseteq M_R$ by drawing a picture, but more algebraically

$$|xy| \leq \left(\frac{|x| + |y|}{2}\right)^2 = R.$$

where the inequality is the AM-GM inequality.[64]

Moroever S_R is a square with side length $2\sqrt{2R}$, and hence $\mathrm{vol}(S_R) = 8R$. Hence if $\mathrm{vol}(S_R) = 8R > 2^2\mathrm{covol}(I) = 4\sqrt{|D_K|}\|I\|$, i.e. $R > \frac{\sqrt{D_K}}{2}\|I\|$, Minkowski's theorem applies, and produces $0 \neq \alpha \in I \cap S_R$, and hence $|N_K(\alpha)| \leq R$. Taking R slightly larger than $\frac{\sqrt{D_K}}{2}\|I\|$ as before we have proved:

Theorem 23.7 (Minkowski's bound for RQF)**.** *For any nonzero ideal $I \subseteq \mathcal{O}_d$, $d > 1$ squarefree, there exists $0 \neq \alpha \in I$ with*

$$N_K(\alpha) \leq \frac{\sqrt{D_K}}{2}\|I\|.$$

23.3 Minkowski in general

Finally we consider the case of a general number field K. How should we view $\mathcal{O}_K$ as a lattice?

Let $\tau_1, \ldots, \tau_r : K \to \mathbf{R}$ be the real embeddings and let $\sigma_1, \overline{\sigma}_1, \ldots, \sigma_s, \overline{\sigma}_s$ be the complex embeddings, so that $r + 2s = n = [K : \mathbf{Q}]$. We define:

$$\begin{aligned} \iota : K &\to \mathbf{R}^n \\ \alpha &\mapsto (\tau_1(\alpha), \ldots, \tau_r(\alpha), \mathrm{Re}(\sigma_1(\alpha)), \mathrm{Im}(\sigma_1(\alpha)), \ldots, \mathrm{Re}(\sigma_s(\alpha)), \mathrm{Im}(\sigma_s(\alpha))) \end{aligned}$$

Note that this exactly generalizes how we viewed either a real or imaginary quadratic field as living inside $\mathbf{R}^2$.

With this we can determine the covolume of an ideal:

Lemma 23.8. *1. $\iota(\mathcal{O}_K) \subset \mathbf{R}^n$ is a lattice with covolume $\frac{\sqrt{|D_K|}}{2^s}$.*

2. For any nonzero ideal $I \subset \mathcal{O}_d$, $\iota(I) \subseteq \mathbf{R}^2$ is a lattice with covolume $\frac{\sqrt{|D_K|}}{2^s}\|I\|$.

Proof. The second part follows from the first as before by Lemma 23.1. To prove the first, pick an integral basis $\alpha_1, \ldots, \alpha_n$ of $\mathcal{O}_K$. We also note the formulas, for a complex number $z \in \mathbf{C}$:

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2}, \qquad \mathrm{Im}(z) = \frac{z - \overline{z}}{2i}$$

[64] Easily proved as follows: $0 \leq (x - y)^2 = x^2 + y^2 - 2xy$, and hence $|xy| \leq \frac{x^2+y^2}{2}$, of $\sqrt{|xy|} \leq \frac{|x|+|y|}{2}$.

Hence we have

$$\begin{pmatrix} \tau_1(\alpha_1) & & \tau_1(\alpha_n) \\ \vdots & & \vdots \\ \tau_r(\alpha_1) & & \tau_r(\alpha_n) \\ \mathrm{Re}(\sigma_1(\alpha_1)) & \cdots & \mathrm{Re}(\sigma_1(\alpha_n)) \\ \mathrm{Im}(\sigma_1(\alpha_1)) & & \mathrm{Im}(\sigma_1(\alpha_n)) \\ \vdots & & \vdots \\ \mathrm{Re}(\sigma_s(\alpha_1)) & & \mathrm{Re}(\sigma_s(\alpha_n)) \\ \mathrm{Im}(\sigma_s(\alpha_1)) & & \mathrm{Im}(\sigma_s(\alpha_n)) \end{pmatrix} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \frac{1}{2} & \frac{1}{2} & & \\ & & & \frac{1}{2i} & \frac{-1}{2i} & & \\ & & & & & \ddots & \\ & & & & & \frac{1}{2} & \frac{1}{2} \\ & & & & & \frac{1}{2i} & \frac{-1}{2i} \end{pmatrix} \begin{pmatrix} \tau_1(\alpha_1) & & \tau_1(\alpha_n) \\ \vdots & & \vdots \\ \tau_r(\alpha_1) & & \tau_r(\alpha_n) \\ \sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha_n) \\ \overline{\sigma}_1(\alpha_1) & & \overline{\sigma}_1(\alpha_n) \\ \vdots & & \vdots \\ \sigma_s(\alpha_1) & & \sigma_s(\alpha_n) \\ \overline{\sigma}_s(\alpha_1) & & \overline{\sigma}_s(\alpha_n) \end{pmatrix}$$

Taking absolute values of determinants and noting that

$$\det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2i} & \frac{-1}{2i} \end{pmatrix} = \frac{i}{2}$$

we obtain $\mathrm{covol}(\iota(\mathcal{O}_K)) = \frac{\sqrt{|D_K|}}{2^s}$. □

Exercise 23.9. By inspecting the above proof, show that if K is a number field and $\alpha_1, \ldots, \alpha_n$ is an integral basis and $\tau_1, \ldots, \tau_n \to \mathbf{C}$ are the embeddings, then $i^s \det(\tau_i(\alpha_j)) \in \mathbf{R}$. Deduce that $(-1)^s D_K > 0$, that is, the sign of the discriminant $D_K \in \mathbf{Z}$ is simply determined by the number of pairs of complex embeddings.

For example for $K = \mathbf{Q}(\sqrt{d})$ quadratic, $D_K > 0$ if $d > 0$ and $D_K < 0$ if $d < 0$. For $K = \mathbf{Q}[x]/(f)$ for $f \in \mathbf{Q}[x]$ cubic irreducible, $D_K > 0$ if f has three real roots, while $D_K < 0$ if f has one real root.

We have

$$N_K(\alpha) = \tau_1(\alpha) \cdots \tau_r(\alpha) \sigma_1(\alpha) \overline{\sigma}_1(\alpha) \cdots$$

and hence

$$|N_K(\alpha)| = |\tau_1(\alpha) \cdots \tau_r(\alpha)| (\mathrm{Re}(\sigma_1(\alpha))^2 + \mathrm{Im}(\sigma_1(\alpha))^2) \cdots (\mathrm{Re}(\sigma_1(\alpha))^2 + \mathrm{Im}(\sigma_1(\alpha))^2)$$

Motivated by this, we define the following region in $\mathbf{R}^n$:

$$M_R = \{(x_1, \ldots, x_r, y_1, z_1, \ldots, y_s, z_s \in \mathbf{R}^n \mid |x_1 \cdots x_r| (y_1^2 + z_1^2) \cdots (y_s^2 + z_s^2) \leq R\}$$

so that $|N_K(\alpha)| \leq R$ if and only if $\iota(\alpha) \in M_R$.

Example 23.10. We can try to visualize this when $n = 3$.

When $r = 3$, $s = 0$ we are looking at the region $|xyz| \leq R$ in $\mathbf{R}^3$. This region has "asymptotes" along each coordinate hyperplane.

When $r = 1$, $s = 1$ we are looking at the region $|x|(y^2 + z^2) \leq R$ in $\mathbf{R}^3$. This region is rotationally symmetric about the x axis, and has asymptotes along the x axis and the y, z plane.

As in the totally real case this region is not convex (unless K is imaginary quadratic.). So we should try to pick a convex centrally symmetric region $S_R \subseteq M_R$ with as large volume as possible. The thing to take is

$$S_R = \{|x_1| + \cdots + |x_r| + 2\sqrt{y_1^2 + z_1^2} + \cdots + 2\sqrt{y_s^2 + z_s^2} \leq nR^{1/n}\}$$

We claim the following things hold:

- S_R is convex and centrally symmetric.
- $S_R \subseteq M_R$.
- $\mathrm{vol}(S_R) = \frac{n^n}{n!} 2^r \left(\frac{\pi}{2}\right)^s R$.

admitting these for the moment, we can apply Minkowski's convex body theorem to S_R when

$$\mathrm{vol}(S_R) > 2^n \mathrm{covol}(I)$$

i.e. when

$$R > \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|D_K|} \|I\| = C_K \|I\|$$

to conclude that there exists $0 \neq \alpha \in I$ with $|N_K(\alpha)| \leq R$. Taking $R > C_K\|I\|$ with the same floor as $C_K\|I\|$ we have proved:

Theorem 23.11. *Let K be a number field and let*

$$C_K = \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|D_K|}.$$

Then for any nonzero ideal $I \subseteq \mathcal{O}_K$ there exists $0 \neq \alpha \in I$ with $|N_K(\alpha)| \leq C_K\|I\|$. Consequently any ideal class in $\mathrm{Cl}(\mathcal{O}_K)$ is represented by an ideal I with $\|I\| \leq C_K$.

It remains to verify the facts we claimed about S_R:

Exercise 23.12.

1. Show that S_R is convex. (Hint: this boils down to: for $v, w \in \mathbf{R}^2$, $t \in [0,1]$, $\|tv + (1-t)w\| \leq t\|v\| + (1-t)\|w\|$ which is a consequence of the triangle inequality. Here $\|(x,y)\| = \sqrt{x^2+y^2}$ is the usual Euclidean norm, not the ideal norm!)
2. Show that $S_R \subseteq M_R$ (Hint: use the AM-GM inequality: $\frac{a_1+\cdots+a_n}{n} \geq (a_1 \cdots a_n)^{1/n}$.)
3. Define

$$V_{r,s}(t) = \mathrm{vol}(\{|x_1| + \cdots + |x_r| + 2\sqrt{y_1^2+z_1^2} + \cdots + 2\sqrt{y_s^2+z_s^2} \leq t\}).$$

Prove the formula:

$$V_{r,s}(t) = \frac{1}{(r+2s)!} 2^r \left(\frac{\pi}{2}\right)^s t^{r+2s}$$

and hence deduce that

$$\mathrm{vol}(S_R) = V_{r,s}(nR^{1/n}) = \frac{n^n}{n!} 2^r \left(\frac{\pi}{2}\right)^s R.$$

To prove this formula, use induction on r and s and reduce to a problem of single variable calculus: first by scaling we have $V_{r,s}(t) = t^{r+2s} V_{r,s}(1)$. Then we have formulas:

$$V_{r,s}(1) = 2 \int_0^1 V_{r-1,s}(1-x)dx$$

and

$$V_{r,s}(1) = \int_{2\sqrt{y^2+z^2}<1} V_{r,s-1}(1 - 2\sqrt{y^2+z^2})dydz.$$

24 Lecture 24: Computing class groups

Before we get to using the Minkowski bound to compute class groups, we remark an amazing thing about Minkowski's bound is that $\frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s$ actually shrinks quite quickly, giving a lower bound on the discriminant of a number field:

Corollary 24.1. *Let K be a number field, then*

$$\sqrt{|D_K|} \geq \frac{n^n}{n!}\left(\frac{\pi}{4}\right)^s.$$

In particular if $K \neq \mathbf{Q}$, $D_K \neq \pm 1$, and hence some prime p ramifies in K by Theorem 18.2.[65]

Proof. Indeed, we claim that the Minkowski constant C_K satisfies $C_K \geq 1$. Indeed, applying the Minkowski bound to the trivial ideal $\mathcal{O}_K$, we conclude that there exists $0 \neq \alpha \in \mathcal{O}_K$ with $|N_K(\alpha)| \leq C_K$. But $N_K(\alpha) \in \mathbf{Z}$, $N_K(\alpha) \neq 0$, hence $|N_K(\alpha)| \geq 1$. It follows that $C_K \geq 1$. □

Exercise 24.2. Prove that if $n > 1$ and $0 \leq s \leq \frac{n}{2}$ then $\frac{n^n}{n!}\left(\frac{\pi}{4}\right)^s > 1$, thus justifying the "in particular" part of the corollary. (Hint: this inequality is pretty weak. It will be enough to start by showing $\frac{n^n}{n!} \geq 2^{n-1}$.)[66]

We now give many examples of computations of class groups. In general if you want to compute the class group of of a number field K what you should do is:

1. Compute the Minkowski constant C_K of K. In practice, in order to do this you need to be able to compute the discriminant D_K. At least for quadratic fields this is no issue![67]

2. Factor all primes $p \leq C_K$. We are very good at this for quadratic fields. For higher degree fields, if $\mathcal{O}_K = \mathbf{Z}[\alpha]$ then you can use Dedekind's criteria, but in general you might suffer for primes where Dedekind's criteria doesn't apply!

3. Try to determine which of the ideals you've found are or are not principal, by using norms: if $I = (\alpha)$ then $\|I\| = |N_K(\alpha)|$, and so you can try to find a solution to $N_K(\alpha) = \pm\|I\|$ (don't forget that norms can be negative!!!!). In practice this step is the hard part: you can try to just "spot" solutions to $N_K(\alpha) = \pm\|I\|$, or prove that no solutions exist using congruences.

For the purpose of this class (and in particular the exam for this class!) we are mostly interested in the quadratic case, but we will also consider some higher degree examples just for fun, to illustrate the power of Minkowski's theorem.

[65] We didn't prove this direction of Theorem 18.2.

[66] Minkowski's bound shows that the discriminant grows exponentially in the degree of the number field. Stronger bounds are obtained using techniques of analytic number theory: for instance Odlyzko showed that $\liminf_K |D_K|^{1/[K:\mathbf{Q}]} \geq 22$ while assuming the Generalized Riemann Hypothesis this can be improved to $8\pi e^\gamma = 44.76\ldots$ (here $\gamma = \lim_{n\to\infty} \sum_{k=1}^n \frac{1}{k} - \log(n) = .577\ldots$ is the Euler-Mascheroni. I'm mostly telling you this because it is amazing that this number with the three constants π, e, γ all appearing occurs naturally!) On the other hand discriminants do not grow faster than exponential in the degree: the best known upper bound is $\liminf_K |D_K|^{1/[K:\mathbf{Q}]} \leq 92.4$, due to J. Martinet. Closing the gap between 22 and 92.4 is a really interesting problem!

[67] In general, you might get away with just bounding D_K and hence C_K above using Proposition 6.9, but then you might have trouble factoring small primes, and you will anyways have to work much harder due to your worse bound on C_K!

Example 24.3. We compute the class group of $\mathcal{O}_{-14} = \mathbf{Z}[\sqrt{-14}]$. The Minkowski constant is $C_K = \frac{2}{\pi}\sqrt{4 \cdot 14} < 5$. Hence we should factor $2, 3$:

$$(2) = (2, \sqrt{-14})^2 = \mathfrak{p}_2^2$$
$$(3) = (3, 1 + \sqrt{-14})(3, 1 - \sqrt{-14}) = \mathfrak{p}_3\overline{\mathfrak{p}}_3$$

Now the Minkowski bound tells us that any ideal class must be one of $1, [\mathfrak{p}_2], [\mathfrak{p}_3], [\overline{\mathfrak{p}}_3]$. Are any of these ideals principal? No because[68]

$$N(a + b\sqrt{-14}) = a^2 + 14b^2 = 2, 3$$

clearly has no solutions, since if $b > 0$ the norm is ≥ 14 and otherwise the norm is a square. But we still need to check that none of the ideal classes $[\mathfrak{p}_2], [\mathfrak{p}_3], [\overline{\mathfrak{p}}_3]$ are the same. If $[\mathfrak{p}_2]$ were $[\mathfrak{p}_3]$, we would have $[\mathfrak{p}_2]^{-1}[\mathfrak{p}_3] = [\mathfrak{p}_2\mathfrak{p}_3]$ is principal. Is that so? We use norms again: $\|\mathfrak{p}_2\mathfrak{p}_3\| = \|\mathfrak{p}_2\|\|\mathfrak{p}_3\| = 6$. Is 6 a norm? No, $a^2 + 14b^2 = 6$ also has no solutions. Hence $[\mathfrak{p}_2] \neq [\mathfrak{p}_3]$, and exactly the same argument shows $[\mathfrak{p}_2] \neq [\overline{\mathfrak{p}}_3]$. Finally we should still check $[\mathfrak{p}_3] \neq [\overline{\mathfrak{p}}_3]$. If they were equal, $[\mathfrak{p}_3][\overline{\mathfrak{p}}_3]^{-1} = [\mathfrak{p}_3^2]$ would be principal. Is it? $\|\mathfrak{p}_3^2\| = 9$. The only solution to $a^2 + 14b^2 = 9$ is $(\pm 3, 0)$. But $(3) = \mathfrak{p}_3\overline{\mathfrak{p}}_3 \neq \mathfrak{p}_3^2$. Hence $\mathfrak{p}_3^2$ is not principal.

So we have shown $\mathrm{Cl}(\mathcal{O}_{-14}) = \{1, \mathfrak{p}_2, \mathfrak{p}_3, \overline{\mathfrak{p}}_3\}$. What is the group structure? Well a group of order 4 is either $\mathbf{Z}/4\mathbf{Z}$ or $(\mathbf{Z}/2\mathbf{Z})^2$. But $\mathfrak{p}_3$ doesn't have order 2 as we saw $\mathfrak{p}_3^2$ isn't principal. Hence $\mathrm{Cl}(\mathcal{O}_{-14}) \simeq \mathbf{Z}/4\mathbf{Z}$ with generator $[\mathfrak{p}_3]$, and so we have $[\mathfrak{p}_2] = [\mathfrak{p}_3]^2$, $[\overline{\mathfrak{p}}_3] = [\mathfrak{p}_3]^3$, $[\mathfrak{p}_3]^4 = 1$.

Our determination of the class group allows us to predict that various ideals are principal. For example $\mathfrak{p}_3^4$ should be a principal ideal. Can we find a generator? We should look for elements of norm $3^4 = 81$, so we want solutions to $a^2 + 14b^2 = 81$. We find solutions $(\pm 9, 0), (\pm 5, \pm 2)$. Of course $(9) = \mathfrak{p}_3^2\overline{\mathfrak{p}}_3^2$ is not what we want, so $\mathfrak{p}_3^4$ must be one of $(5 + 2\sqrt{-14})$, $(5 - 2\sqrt{-14})$. How can we tell which? Well we can see which of these elements is contained in $\mathfrak{p}_3$:

$$5 + 2\sqrt{-14} = 3 + 2(1 + \sqrt{-14}) \in \mathfrak{p}_3$$

but

$$5 - 2\sqrt{-14} = 1 + 6 - 2(1 + \sqrt{-14}) \notin \mathfrak{p}_3$$

so $\mathfrak{p}_3^4 = (5 + 2\sqrt{-14})$.

Remark 24.4. For imaginary quadratic fields it is always easy to check if a number is a norm. For example if I asked you, is 79 a norm in $\mathbf{Z}[\sqrt{-14}]$ you could easily check $x^2 + 14y^2 = 79$ has no solutions: 79, $79 - 14 = 65$, $79 - 14 \cdot 4 = 23$ aren't squares, and $14 \cdot 9^2 > 79$.

But what if you tried to use congruences? $79 \equiv 2 \pmod 7$ and $3^2 \equiv 2 \pmod 7$. What about mod 4 or mod 8? $79 \equiv -1 \pmod 8$ and $14 \equiv -2 \pmod 8$, and $1^2 - 2 \cdot 1^2 \equiv -1 \pmod 8$. It turns out that the congruence $x^2 + 14y^2 \equiv 79 \pmod n$ has solutions for all integers n[69] and so we cannot show 79 is not a norm just using congruences!

This would be more of a problem if we were considering an equation like $x^2 - dy^2 = m$ for $d > 0$!

[68] Remember we should really consider elements of norm $\pm 2, \pm 3$ but the norm is always positive in the imaginary quadratic case!

[69] Try to show this if you like, as an exercise in elementary number theory

Example 24.5. We compute the class group of $K = \mathbf{Q}(\sqrt{33})$. $33 \equiv 1 \pmod 4$ so $D_K = 33$ and $C_K = \frac{\sqrt{33}}{2} < 3$. Thus we only need to factor 2. As $33 \equiv 1 \pmod 8$ we have

$$(2) = \left(2, \frac{1+\sqrt{33}}{2}\right)\left(2, \frac{1-\sqrt{33}}{2}\right) = \mathfrak{p}_2\overline{\mathfrak{p}}_2$$

Are these ideals principal? We consider norms:

$$N(a + b\frac{1+\sqrt{33}}{2}) = a^2 + ab - 8b^2.$$

Luckily we quickly spot $a = 2$, $b = 1$, so

$$\left(\frac{5 \pm \sqrt{33}}{2}\right)$$

are ideals of norm 2 and they must be $\mathfrak{p}_2, \overline{\mathfrak{p}}_2$. Clearly $\frac{5+\sqrt{33}}{2} \in \mathfrak{p}_2$ so in fact

$$\mathfrak{p}_2 = \left(\frac{5+\sqrt{33}}{2}\right).$$

Hence we have shown that the class group group is trivial and so $\mathbf{Z}[\frac{1+\sqrt{33}}{2}]$ is a PID.

In general, class groups are quite mysterious, even for quadratic fields! Here is a famous theorem and conjecture:

The following was conjectured by Gauss. (See problem 4 on example sheet 4 for some evidence that might have lead him to believe this.)

Theorem 24.6 (Heegner-Baker-Stark)**.** *If $\mathcal{O}_d$ is a PID for $d < 0$ squarefree then*[70]

$$d = -1, -2, -3, -7, -11, -19, -43, -67, -163.$$

By contrast it seems much more common for $\mathcal{O}_d$ to be a PID for $d > 0$. Nonetheless the following is still an open problem:

Conjecture 24.7. There are infinitely many $d > 1$ squarefree for which $\mathcal{O}_d$ is a PID.

Now we consider some higher degree examples, just for fun.

Example 24.8. We consider the field from Example 6.16, $K = \mathbf{Q}(\alpha)$ where α is a root of $f = x^3 + x + 1$. We saw that $\mathrm{disc}(\mathbf{Z}[\alpha]) = -31$, so that $\mathcal{O}_K = \mathbf{Z}[\alpha]$ and $D_K = -31$. Moreover f has one real root (e.g. because the discriminant is negative) so $r = 1, s = 1$. Thus the Minkowski bound is

$$C_K = \frac{3!}{3^3}\left(\frac{4}{\pi}\right)\sqrt{31} = \frac{8\sqrt{31}}{9\pi} < 2$$

So we don't have to factor anything! $\mathbf{Z}[\alpha]$ is a PID.

[70] By a theorem of Siegel, for any $\varepsilon > 0$ there is a constant $C(\varepsilon) > 0$ with $\#\mathrm{Cl}(\mathcal{O}_d) \geq C(\varepsilon)d^{1/2-\varepsilon}$. Unfortunately nobody knows what $C(\varepsilon)$ is (it might be really small!) and so this can't be used to prove this theorem!

Example 24.9. We consider $K = \mathbf{Q}(\sqrt[3]{2})$, which we have considered before on example sheet 2, and examples 17.3, 18.11. We have seen that $\mathcal{O}_K = \mathbf{Z}[\sqrt[3]{2}]$, $D_K = -3^3 \cdot 2^2 = -108$. We have $n = 3, r = 1, s = 1$.

The Minkowski constant is then

$$C_K = \frac{3!}{3^3}\frac{4}{\pi}\sqrt{3^3 \cdot 2^2} = \frac{2^4\sqrt{2}}{3\pi} < 3$$

Thus any non principal ideal in $\mathcal{O}_K$ has norm 2. But we saw that

$$(2) = (\sqrt[3]{2})^3$$

and so $(\sqrt[3]{2})$ is the only ideal of norm 2 and it is principal! Hence the class group is trivial, and so $\mathbf{Z}[\sqrt[3]{2}]$ is a PID!

Example 24.10. We consider $K = \mathbf{Q}(\sqrt[4]{2})$, which you considered on coursework 1. You showed that $\operatorname{disc}(\mathbf{Z}[\sqrt[4]{2}]) = -2^{11} = -2048$. Moreover, the minimal polynomial of $\sqrt[4]{2}$, $x^4 - 2$, is Eisenstein at 2, so $\mathcal{O}_K = \mathbf{Z}[\sqrt[4]{2}]$ by Proposition 18.9, and so $D_K = -2^{11}$. Two of the roots of $x^4 - 2$ are real and two are complex, so $r = 2, s = 1$. The Minkowski constant is thus

$$C_K = \frac{4!}{4^4}\frac{4}{\pi}\sqrt{2^{11}} = \frac{12\sqrt{2}}{\pi} < 6$$

Now we start factoring primes:

$$(2) = (\sqrt[4]{2})^4$$

only provides principal ideals. What about 3 and 5? Well we could factor them, but we are only worried about ideals $\mathfrak{p}$ of $\|\mathfrak{p}\| = 3, 5$, so by Dedekind's criteria we just need to check if $x^4 - 2$ has a root mod 3 or mod 5. But it doesn't: $a^4 \equiv 0, 1 \pmod 3$ and $a^4 \equiv 0, 1 \pmod 5$.

Finally let's see how far we can get towards proving the following theorem, which is beyond the scope of this course:

Theorem 24.11. *Let p be an odd prime. Then $\mathbf{Z}[\zeta_p]$ is a PID if and only if $p \leq 19$.*

Let $K = \mathbf{Q}(\zeta_p)$ be the pth cyclotomic field. We have $n = [K : \mathbf{Q}(\zeta_p)] = p - 1$, and all the embeddings are complex (there are no non trivial pth roots of 1 in $\mathbf{R}$) so $s = \frac{p-1}{2}$. The discriminant $|D_K| = p^{p-2}$ by Theorem 20.1. Thus

$$C_K = \frac{(p-1)!}{(p-1)^{p-1}}\left(\frac{4}{\pi}\right)^{\frac{p-1}{2}} p^{\frac{p-2}{2}}$$

For $p = 3, 5, 7, 11, 13$ we have $C_K < 2, < 2, < 5, < 59, < 307$, while for $p = 17$, $C_K < 1.4 \times 10^4$ and for $p = 19$, $C_K < 1.1 \times 10^5$.

Thus we see immediately that $\mathbf{Z}[\zeta_3]$ and $\mathbf{Z}[\zeta_5]$ are PIDs (we already knew the first one, as $\mathbf{Z}[\zeta_3] = \mathcal{O}_{-3}$) while to go farther we have to work a bit.

We remind ourselves that $(p) = (\zeta_p - 1)^{p-1}$ only provides principal primes, while a prime $q \neq p$ factors in $\mathbf{Z}[\zeta_p]$ as $\mathfrak{q}_1 \cdots \mathfrak{q}_r$ where each factor $\mathfrak{q}_i$ satisfies $\|\mathfrak{q}_i\| = q^f$ where f is the order q in $\mathbf{Z}/p\mathbf{Z}$. In particular $q^f \equiv 1 \pmod p$. So we only have to worry about primes q where $q^f < C_K$.

In particular, $q^f > p$, so for $\mathbf{Z}[\zeta_7]$ we are done immediately. What about $\mathbf{Z}[\zeta_{11}]$? Of the numbers < 59 which are 1 mod 11: $12, 23, 34, 45, 56$, only one is a prime power: 23. We know that $(23) = \mathfrak{q}_1 \cdots \mathfrak{q}_{10}$ where each factor has norm 23. Can you find an element of $\mathbf{Z}[\zeta_{11}]$ with norm ± 23?

25 Lecture 25: Example class 4

1. Prove that for an odd prime p, $\left(\frac{2}{p}\right) = 1$ if and only if $p \equiv \pm 1 \pmod 8$:

   (a) Prove that $\Phi_8(x) = \frac{x^8-1}{x^4-1} = x^4 + 1$ is irreducible and hence is the minimal polynomial of $\zeta_8 = e^{2\pi i/8}$. Show that $\mathcal{O}_{\mathbf{Q}(\zeta_8)} = \mathbf{Z}[\zeta_8]$.

   (b) Show that if $p \equiv 1 \pmod 8$ then $p\mathbf{Z}[\zeta_8]$ is the product of four distinct primes. (Hint: you just need to show that $x^8 - 1$ has 8 roots in $\mathbf{F}_p$).

   (c) Show that $\sqrt{2}, \sqrt{-2}, \sqrt{-1} \in \mathbf{Q}(\zeta_8)$ (Hint: $\zeta_8 = \frac{\sqrt{2}+\sqrt{-2}}{2}$.)

   (d) Deduce from the previous two parts that if $p \equiv 1 \pmod 8$ then p splits in $\mathcal{O}_2$, $\mathcal{O}_{-2}$, and $\mathcal{O}_{-1}$, and in particular $\left(\frac{2}{p}\right) = 1$.

   (e) Show that if $p \equiv \pm 3 \pmod 8$ then p is inert in $\mathcal{O}_2$, and hence $\left(\frac{2}{p}\right) = -1$. (Hint: use that $\mathbf{Z}[\sqrt{2}]$ is a PID and norms to reduce this to showing that $x^2 - 2y^2 = \pm p$ has no solutions, and then show this has no solutions mod 8.)

   (f) The remaining case is $p \equiv -1 \pmod 8$. Do this by writing $\left(\frac{2}{p}\right) = \left(\frac{-2}{p}\right)\left(\frac{-1}{p}\right)$ and show that both of the factors are -1.

2. Let $d \neq 1$ be square free, and let $N = |D_{\mathbf{Q}(\sqrt{d})}|$ (so $N = |d|$ if $d \equiv 1 \pmod 4$ and $N = 4|d|$ otherwise.)

   (a) Write $d = (-1)^a 2^b q_1^* \cdots q_r^*$ where $a, b \in \{0, 1\}$ where $q_1, \ldots, q_r$ are odd primes. Show that $N = 2^c q_1 \cdots q_r$ where c is determined by a, b.

   (b) Use quadratic reciprocity and part a to show that show that $\left(\frac{d}{p}\right) = \left(\frac{d}{p'}\right)$ if $p \equiv p' \pmod N$.

   (c) Use part a to show that $\sqrt{d} \in \mathbf{Q}(\zeta_N)$.

3. Compute the class group of $\mathcal{O}_{-6}, \mathcal{O}_{-23}, \mathcal{O}_{14}$.

4. The goal of this exercise is to prove the following result: if $\mathcal{O}_d$ is a PID for $d < 0$ then either $d = -1, -2, -7$ or $d \equiv 5 \pmod 8$ and $-d$ is prime.

   (a) Show that $\mathcal{O}_d$ has no elements of norm 2 unless $d = -1, -2, -7$ (you might have already done this on example sheet 2) and deduce that otherwise 2 must be inert in $\mathcal{O}_d$ and hence $d \equiv 5 \pmod 8$.

   (b) Now assuming $d \equiv 1 \pmod 4$, show that if a prime p is a norm, then $p \geq \frac{1-d}{4}$. Deduce that any prime $p < \frac{1-d}{4}$ must be inert in $\mathcal{O}_d$.

   (c) Deduce from the previous part that $-d$ must be prime. (Hint: any prime factor of d ramifies in $\mathcal{O}_d$).

   (d) Just for fun, if you wanted to look for more $d < 0$ for which $\mathcal{O}_d$ is a PID (there aren't any!) you could consider the congruence conditions imposed on d for $3, 5, 7$ to be inert. Show that if $-d > 19$ then

   $$d \equiv -43, -67, -163, -403, -547, -667 \pmod{8 \cdot 3 \cdot 5 \cdot 7}$$

   and check that $\mathcal{O}_{-403}, \mathcal{O}_{-547}, \mathcal{O}_{-667}$ aren't PIDs. Keep going if you'd like!

$$-d = 883, 907, 1723, 1747, 2083, 2347, 2683, 3067, 3187, \ldots$$

(e) Euler discovered the rather remarkable fact n^2-n+41 is prime for $1 \leq n \leq 40$. Prove this using that $\mathcal{O}_{-163}$ is a PID: start by showing that $N(n+\frac{-1+\sqrt{-163}}{2}) = n^2-n+41$. Now show that if $n+\frac{-1+\sqrt{-163}}{2}$ is not prime, it has a prime factor with norm < 41, and use part b.

5. Prove Fermat's theorem that if $p \equiv 1 \pmod 4$ is prime then $p = x^2+y^2$ using Minkowski's convex body theorem: consider the lattice in $\mathbf{R}^2$ spanned by $(p,0)$ and $(a,1)$ where $a^2 \equiv -1 \pmod p$ and the circle $S = \{(x,y) \mid x^2+y^2 < 2p\} \subset \mathbf{R}^2$.

26 Lecture 26: Applications of class groups I

26.1 Primes of the form x^2+ny^2

Now that we know how to compute class groups of quadratic fields, our goal is to give some number theoretic applications. We first consider generalizations of the following theorem of Fermat:

Theorem 26.1 (Fermat)**.** *Let p be a prime. Then we have:*

1. *$p = x^2+y^2$ for $x, y \in \mathbf{Z}$ if and only if $p = 2$ or $p \equiv 1 \pmod 4$.*
2. *$p = x^2+2y^2$ for $x, y \in \mathbf{Z}$ if and only if $p = 2$ or $p \equiv 1, 3 \pmod 8$.*

Proof. We have $p = x^2+y^2$ if and only if $p = (x+iy)(x-iy) \in \mathbf{Z}[i]$. Since $\mathbf{Z}[i]$ is a PID, this is equivalent to the existence of an ideal $\mathfrak{p} \subset \mathbf{Z}[i]$ with $\|\mathfrak{p}\| = p$. This is equivalent to p ramifying or splitting in $\mathbf{Z}[i]$, which is equivalent to $p = 2$ or $\left(\frac{-1}{p}\right) = 1$, which is finally equivalent to $p = 2$ or $p \equiv 1 \pmod 4$.

The argument for x^2+2y^2 is exactly the same, again using that $\mathbf{Z}[\sqrt{-2}]$ is a PID. We now conclude that $p = x^2+2y^2$ if and only if $p = 2$ or $\left(\frac{-2}{p}\right) = 1$. By writing $\left(\frac{-2}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{-1}{p}\right)$ and going through the possibilities for p mod 8 we see that this is equivalent to $p = 2$ or $p \equiv 1, 3 \pmod 8$. □

Exercise 26.2. 1. Show that $p = x^2+xy+y^2$ for $x, y \in \mathbf{Z}$ if and only if $p = 3$ or $p \equiv 1 \pmod 3$. (Hint: recall that $N(x+y\frac{1+\sqrt{-3}}{2}) = x^2+xy+y^2$).

2. Show more generally that if $d < 0$ is squarefree and $d \equiv 1 \pmod 4$ and $\mathcal{O}_d$ is a PID (so $d = -3, -7, -11, -19, -43, -67, -163$) then $p = x^2+xy+\frac{-d-1}{4}y^2$ for $x, y \in \mathbf{Z}$ if and only if $\left(\frac{p}{-d}\right) = 0, 1$.

3. Show that $p = x^2+3y^2$ if and only if $p = 3$ or $p \equiv 1 \pmod 3$ (Hint: start with $\alpha \in \mathcal{O}_{-3}$ with $N(\alpha) = 3$ and consider $\zeta_6^i\alpha$.)

4. Show that $p = x^2+7y^2$ if and only if $p = 7$ or $p \neq 2$ and $p \equiv 1, 2, 4 \pmod 7$. (Hint: show that if $p = x^2+xy+2y^2$ and p is odd then in fact y is even.)

5. Show that $p = x^2+4y^2$ if and only if $p \equiv 1 \pmod 4$.

All of the examples above have used that the ring $\mathcal{O}_d$ is a PID. What if it is not? The first example to consider is, when is $p = x^2 + 5y^2$? This example was considered by Euler, and he made the following conjecture[71]:

- $p \equiv x^2 + 5y^2$ if and only if $p = 5$ or $p \equiv 1, 9 \pmod{20}$.
- If $p, q \equiv 3, 7 \pmod{20}$ then $pq = x^2 + 5y^2$.

Using congruences mod 4 and mod 5 we see that $p = 5$ or $p \equiv 1, 9 \pmod{20}$ is a necessary condition for $p = x^2 + 5y^2$ so the hard part is to prove sufficiency. The second part of Euler's conjecture is quite interesting because nothing like this happens in the cases studied by Fermat: if $pq = x^2+y^2$ for primes $p \neq q$ then both p and q are themselves sums of two squares (factor $x + iy \in \mathbf{Z}[i]$ into irreducibles and take their norms).

It turns out we know two things which Euler didn't which allows us to solve this problem! These two things are: quadratic reciprocity and the class group.

We recall that we computed the class group of $\mathcal{O}_{-5}$ in Example 22.4. The Minkowski bound is $\frac{2}{\pi}\sqrt{20} < 3$ so any ideal class is represented by an ideal of norm 1 and 2. But 2 ramifies: $(2) = \mathfrak{p}_2^2$ where $\mathfrak{p}_2 = (2, 1 + \sqrt{-5})$. Moreover $\mathfrak{p}_2$ is not principal because $2 = x^2 + 5y^2$ has no solutions. Thus

$$\mathrm{Cl}(\mathcal{O}_{-5}) = \{1, [\mathfrak{p}_2]\} \simeq \mathbf{Z}/2\mathbf{Z}.$$

If a prime p splits as $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ in $\mathcal{O}_{-5}$, then either $\mathfrak{p}$ is principal or $[\mathfrak{p}] = \mathfrak{p}_2$. Remarkably we can describe which happens as a congruence condition on p:

Theorem 26.3. *Let p be prime.*

- *If $p \equiv 1, 9 \pmod{20}$ then $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ splits in $\mathcal{O}_{-5}$ and $\mathfrak{p}, \overline{\mathfrak{p}}$ are principal. Moreover $p = x^2 + 5y^2$ for $x, y \in \mathbf{Z}$.*
- *If $p \equiv 3, 7 \pmod{20}$ then $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ splits in $\mathcal{O}_{-5}$ and $[\mathfrak{p}] = [\overline{\mathfrak{p}}] = [\mathfrak{p}_2]$. Moreover $2p = x^2 + 5y^2$ for $x, y \in \mathbf{Z}$.*
- *$(2) = \mathfrak{p}_2^2$ and $(5) = (\sqrt{-5})^2$ ramify.*
- *In the remaining cases p is inert in $\mathcal{O}_{-5}$.*

Proof. We know that an odd prime p splits in $\mathcal{O}_{-5}$ if and only if $\left(\frac{-5}{p}\right) = 1$. We first use quadratic reciprocity to describe this in terms of a congruence condition on p mod 20. We have

$$\left(\frac{-5}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{5}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{5}\right)$$

□

Thus to have $\left(\frac{-5}{p}\right) = 1$ there are two possibilities:

- $\left(\frac{-1}{p}\right) = \left(\frac{p}{5}\right) = 1$ i.e. $p \equiv 1 \pmod{4}$ and $p \equiv \pm 1 \pmod{5}$. By the Chinese remainder theorem this is equivalent to $p \equiv 1, 9 \pmod{20}$.

[71] If you made a table of numbers of the form $x^2 + 5y^2$ you would have discovered this conjecture too!

- $\left(\frac{-1}{p}\right) = \left(\frac{p}{5}\right) = -1$ i.e. $p \equiv 3 \pmod 4$ and $p \equiv \pm 2 \pmod 5$. By the Chinese remainder theorem this is equivalent to $p \equiv 3, 7 \pmod{20}$.

Thus if $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ splits if and only if $p \equiv 1, 3, 7, 9 \pmod{20}$, and what remains to do is consider the ideal class of $\mathfrak{p}, \overline{\mathfrak{p}}$. There are two cases, either $[\mathfrak{p}] = 1$ or $[\mathfrak{p}] = [\mathfrak{p}_2]$.

- If $[\mathfrak{p}] = 1$ then $\mathfrak{p}$ is principal. Thus there exists an element of $\mathcal{O}_{-5}$ with norm p, and thus $p = x^2 + 5y^2$. Considering this mod 5 we see that $p \equiv \pm 1 \pmod 5$.

- If $[\mathfrak{p}] = [\mathfrak{p}_2]$ then $\mathfrak{p}\mathfrak{p}_2$ is principal (since $[\mathfrak{p}\mathfrak{p}_2] = [\mathfrak{p}_2]^2 = 1$). Thus there exists an element of $\mathcal{O}_{-5}$ with norm $2p$, and thus $2p = x^2 + 5y^2$. Considering this mod 5 we see that $2p = \pm 1 \pmod 5$ or $p = \pm 2 \pmod 5$.

If $p \equiv 1, 9 \pmod{20}$ then the second possibility cannot occur so the first does, while if $p \equiv 3, 7 \pmod{20}$ then the first possibility does not occur so the second does. This proves the theorem.

Corollary 26.4.
- *We have $p = x^2 + 5y^2$ if and only if $p = 5$ or $p \equiv 1, 9 \pmod{20}$*

- *We have $2p = x^2 + 5y^2$ if and only if $p = 2$ or $p \equiv 3, 7 \pmod{20}$.*

- *If p, q are either 2 or $3, 7 \pmod{20}$ then $pq = x^2 + 5y^2$.*

Proof. We have $p = x^2 + 5y^2$ if and only if $p = N(x + y\sqrt{-5})$ if and only if there exists a principal ideal $\mathfrak{p} = (x + \sqrt{-5})$ with $\|\mathfrak{p}\| = p$. By Theorem 26.3 this happens if and only if $p = 5$ or $p \equiv 1, 9 \pmod{20}$.

Similarly we have $2p = x^2 + 5y^2$ if and only if $2p = N(x + y\sqrt{-5})$ if and only if there exists a principal ideal $I = (x + y\sqrt{-5})$ with $\|I\| = 2p$. Factoring this ideal into primes, we must have $I = \mathfrak{p}_2\mathfrak{p}$ where $\|\mathfrak{p}\| = p$ and $[\mathfrak{p}] = [\mathfrak{p}_2]^{-1} = [\mathfrak{p}_2]$. By the theorem this happens when $p = 2$ or $p \equiv 3, 7 \pmod{20}$.

For the third part, if p, q are either 2 or $3, 7 \pmod{20}$ then by the theorem again there exist ideals $\mathfrak{p}, \mathfrak{q} \subseteq \mathcal{O}_{-5}$ with $[\mathfrak{p}] = [\mathfrak{q}] = [\mathfrak{p}_2]$ and $\|\mathfrak{p}\| = p$, $\|\mathfrak{q}\| = q$. Then $\|\mathfrak{p}\mathfrak{q}\| = pq$ and $[\mathfrak{p}\mathfrak{q}] = [\mathfrak{p}_2]^2 = 1$ so $\mathfrak{p}\mathfrak{q} = (x + y\sqrt{-5})$ is principal, and $pq = N(x + y\sqrt{-5}) = x^2 + 5y^2$. □

26.2 Mordell's equation

Now we consider examples of a Diophantine equation which can be studied using unique factorization, as we motivated in the first lecture. The problem is the following:

Given $d \neq 0$, find all solutions $x, y \in \mathbf{Z}$ to the equation

$$y^2 = x^3 + d.$$

This Diophantine is called Mordell's equation and is an example of an elliptic curves.[72]

Here is a basic strategy:

- Rewrite the equation as $y^2 - d = x^3$ and then factor the left hand side in $\mathbf{Z}[\sqrt{d}]$:

$$(y + \sqrt{d})(y - \sqrt{d}) = x^3.$$

[72]Typically when we study elliptic curves we are also interested in solutions where $x, y \in \mathbf{Q}$, but here we only look for integer solutions.

- Try to use what we know about factorization in $\mathcal{O}_d$ to show that the factors on the left, $y + \sqrt{d}$ and $y - \sqrt{d}$ are cubes, and then use this to solve the equation.

You might compare this to the strategy for finding all solutions to $x^2 + y^2 = z^2$ in the first lecture, or the approach to Fermat's last theorem.

We do an example which you might have seen in elementary number theory.

Example 26.5. We consider the Diophantine equation $y^2 = x^3 - 1$. You might easily spot the solution $y = 0$, $x = 1$, but are there any more?

Following the strategy we write this as $(y + i)(y - i) = x^3$.

We first claim that $y + i$ and $y - i$ are coprime in $\mathbf{Z}[i]$. Indeed if they had a common prime factor π, we would have $\pi|(y + i) - (y - i) = 2i$, which would imply that $\pi|2$ or $\pi = 1 + i$ up to unit. But then this implies that $\pi \mid x$ or in other words $2 \mid x$ so x is even. But then considering the original Diophantine equation mod 8 we get $y^2 \equiv -1 \pmod 8$ which has no solutions.

Once we know that $y + i$ and $y - i$ are coprime we apply unique factorization in $\mathbf{Z}[i]$. We claim $y + i = u\alpha^3$ for $u \in \mathbf{Z}[i]^\times$ and $\alpha \in \mathbf{Z}[i]$. Indeed for any prime factor π of $y + i$, π divides x^3 a multiple of three times, while it cannot divide $y - i$, so it must divide $y + i$ a multiple of three times also. We furthermore note that $\mathbf{Z}[i]^\times = \{\pm 1, \pm i\}$ and each element is also a cube: either use that 3 is coprime to the order 4 of this group, or just observe explicitly: $(\pm 1)^3 = \pm 1$, $i^3 = -i$, $(-i)^3 = i$. It follows that $y + i = (a + ib)^3$ is a cube in $\mathbf{Z}[i]$.

Expanding this out we find:

$$y + i = (a^3 - 3ab^2) + (3a^2b - b^3)i$$

Considering the coefficient of i we get the new Diophantine equation:

$$3a^2b - b^3 = 1$$

This is much easier to solve than the original one! We factor $b(3a^2 - b^2) = 1$. Thus either

- $b = 3a^2 - b^2 = 1$. But this gives $3a^2 = 2$ which has no solutions.
- $b = 3a^2 - b^2 = -1$. This gives $3a^2 = 0$, so we get the solution $a = 0$, $b = -1$.

To find the solution to the original equation, we have then $y = a^3 - 3ab^2 = 0$, and finally $x = 1$. Thus $y = 0$, $x = 1$ is the only solution.

Here are some more examples of using unique factorization to solve Diophantine equations. In the next lecture we will discuss what happens when we don't have unique factorization.

Exercise 26.6. 1. Show that the only solutions to $y^2 = x^3 - 2$ are $y = \pm 5$, $x = 3$ using unique factorization in $\mathbf{Z}[\sqrt{-2}]$.

2. Show that the only solutions to $y^2 = x^3 + 1$ are $y = 0$, $x = -1$ and $y = \pm 3$, $x = 2$, using unique factorization in $\mathbf{Z}$.[73]

[73]The famous Catalan conjecture, formulated in 1844, asserts that the only solution to $x^a - y^b = 1$ for $a, b > 1$, $x, y > 0$ is $3^2 - 2^3 = 1$. This was proved by Mihăilescu in 2002.

27 Lecture 27: Applications of class groups II

Now we consider cases of Mordell's equation where we do not have unique factorization. We begin with an example

Example 27.1. We consider the equation $y^2 = x^3 - 5$. We rewrite this as

$$(y + \sqrt{-5})(y - \sqrt{-5}) = x^3.$$

We first claim that the factors $y + \sqrt{-5}, y - \sqrt{-5} \in \mathcal{O}_{-5}$ are coprime.[74] We need to show there is no prime ideal $\mathfrak{p} \subseteq \mathcal{O}_{-5}$ with $y + \sqrt{-5}, y - \sqrt{-5} \in \mathfrak{p}$. Suppose there were. Then we would also have $(y + \sqrt{-5}) - (y - \sqrt{-5}) = 2\sqrt{-5} \in \mathfrak{p}$. Hence $\|\mathfrak{p}\| \mid N(2\sqrt{-5}) = 20$. It follows that $\mathfrak{p}$ lies over 2 or 5. Now $x^3 = (y + \sqrt{-5})(y - \sqrt{-5}) \in \mathfrak{p}$ and hence $x \in \mathfrak{p}$. Since $x \in \mathbf{Z}$ it follows that $2 \mid x$ or $5 \mid x$.

But if $2 \mid x$ then $y^2 \equiv -1 \pmod 4$ which is not possible. Similarly if $5|x$ then $5 \mid y^2 = x^3 - 5$ and so $5 \mid y$. But then $25 \mid y^2 - x^3 = -5$, hence $5 \mid x$ is also not possible. Thus we have proven that $y + \sqrt{-5}, y - \sqrt{-5}$ are coprime.

Now we use unique factorization into prime ideals. The equation

$$(y + \sqrt{-5})(y - \sqrt{-5}) = (x)^3$$

shows that we must have $(y + \sqrt{-5}) = I^3$ for some ideal I. Indeed any prime ideal occurring in the factorization of $(y + \sqrt{-5})$ cannot occur in the factorization of $y - \sqrt{-5}$ by what we have just proved, and it occurs in $(x)^3$ a multiple of three times. Thus it must occur in $y + \sqrt{-5}$ a multiple of three times.

What can we say about the ideal I? In the class group $\mathrm{Cl}(\mathcal{O}_{-5})$ we have

$$[I]^3 = [(y + \sqrt{-5})] = 1.$$

But we have computed $\mathrm{Cl}(\mathcal{O}_{-5}) = \{1, [\mathfrak{p}_2]\} \simeq \mathbf{Z}/2\mathbf{Z}$. In particular as $[\mathfrak{p}_2]^3 = [\mathfrak{p}_2]$, the formula $[I]^3 = 1$ implies $[I] = 1$, i.e. I is principal. Even though we don't have unique factorization in $\mathcal{O}_{-5}$, using the class group we still concluded what we wanted!

Writing $I = (\alpha)$ for $\alpha \in \mathcal{O}_{-5}$ we have $(y+\sqrt{-5}) = I^3 = (\alpha^3)$, and hence $y+\sqrt{-5} = u\alpha^3$ for $u \in \mathcal{O}_{-5}^\times$ a unit. But $\mathcal{O}_{-5}^\times = \{\pm 1\}$ and so $u = u^3$ and we have $u\alpha^3 = (u\alpha)^3$. Thus writing $u\alpha = a + b\sqrt{-5}$ we have

$$y + \sqrt{-5} = (a + b\sqrt{-5})^3 = (a^3 - 15ab^2) + (3a^2b - 5b^3)\sqrt{-5}$$

or equating coefficients:

$$1 = b(3a^2 - 5b^2).$$

Hence either

- $b = 3a^2 - 5b^2 = 1$ and hence $3a^2 = 6$ or $a^2 = 2$, which has no solutions.
- or $b = 3a^2 - 5b^2 = -1$ and hence $3a^2 = 4$ which has no solutions.

Thus we have shown that the original diophantine equation

$$y^2 = x^3 - 5$$

has no solutions $x, y \in \mathbf{Z}$.

[74] I.e. $(y + \sqrt{-5}, y - \sqrt{-5}) = \mathcal{O}_{-5}$, or equivalently if we consider the factorization of $(y + \sqrt{-5})$ and $(y - \sqrt{-5})$ into prime ideals, there are no common prime factors.

Let's try to generalize this. We consider $d < 0$ squarefree and $d \equiv 2, 3 \pmod 4$,[75] and try to solve the equation

$$y^2 = x^3 + d.$$

As usual we rewrite this as $(y + \sqrt{d})(y - \sqrt{d}) = x^3$, and begin by showing that $y + \sqrt{d}$ and $y - \sqrt{d}$ are coprime.

If $\mathfrak{p} \subset \mathcal{O}_d$ is a prime ideal and $y + \sqrt{d}, y - \sqrt{d} \in \mathfrak{p}$ then $(y + \sqrt{d}) - (y - \sqrt{d}) = 2\sqrt{d} \in \mathfrak{p}$. As $N(2\sqrt{d}) = -4d \in \mathfrak{p}$, we conclude that either $\mathfrak{p}$ lies over 2 or $\mathfrak{p}$ lies over one of the prime factors of d. Then as $x^3 \in \mathfrak{p}$ and hence $x \in \mathfrak{p}$ and $x \in \mathbf{Z}$, we conclude that either $2 \mid x$ or $\gcd(d, x) \neq 1$.

If $2 \mid x$ then we obtain $y^2 \equiv 0 + d \pmod 4$ which has no solutions as $d \equiv 2, 3 \pmod 4$. Otherwise if $p \mid d$ and $p \mid x$ then $p \mid y^2 = x^3 + d$ so $p \mid y$. But then $p^2 \mid d = y^2 - x^3$ which is impossible as d is squarefree.

Thus we have proved that $y + \sqrt{d}$ and $y - \sqrt{d}$ are coprime. Now using unique factorization into prime ideals as before, we conclude from this and

$$(y + \sqrt{d})(y - \sqrt{d}) = (x)^3$$

that

$$(y + \sqrt{d}) = I^3.$$

for some ideal $I \subseteq \mathcal{O}_d$. Hence $[I]^3 = 1$ in $\mathrm{Cl}(\mathcal{O}_d)$. At this point we are stuck if we don't make an additional assumption. So lets assume that $3 \nmid \#\mathrm{Cl}(\mathcal{O}_d)$. This implies that $\mathrm{Cl}(\mathcal{O}_d)$ has no elements of order 3, so $[I]^3 = 1$ implies $[I] = 1$. Thus $I = (\alpha)$ is principal, and so $(x + \sqrt{d}) = (\alpha^3)$ and hence $y + \sqrt{d} = u\alpha^3$ for $\alpha \in \mathcal{O}_d^\times$. By our assumptions on d, either $\mathcal{O}_d^\times = \{\pm 1\}$ if $d \neq -1$ or $\mathcal{O}_d^\times = \{\pm 1, \pm i\}$ if $d = 1$. Either way $u = u'^3$ for some unit $u' \in \mathcal{O}_d^\times$, and so $y + \sqrt{d} = (u'\alpha)^3$. Writing $u'\alpha = a + b\sqrt{d}$ we obtain the equation:

$$y + \sqrt{d} = (a + b\sqrt{d})^3 = (a^3 + 3dab^2) + (3a^2b + db^3)\sqrt{d}$$

equating coefficients this gives the equation

$$1 = b(3a^2 + db^2)$$

which has a solution if and only if $3a^2 + d = \pm 1$ has a solution for $a \in \mathbf{Z}$ (note that if a is a solution then so is $-a$, and that this equation can only have a solution for $+1$ or -1 but not both.)

If this equation has a solution, we can solve for the original x, y in terms of a:

$$y = \pm a(a^2 + 3d)$$

and

$$x^3 = (y + \sqrt{d})(y - \sqrt{d}) = (a + b\sqrt{d})^3(a - b\sqrt{d})^3 = (a^2 - d)^3$$

and hence

$$x = a^2 - d$$

To summarize what we have proved:

[75] You can also try to handle $d \equiv 5 \pmod 8$ similarly, with things becoming a bit more annoying since $\mathcal{O}_d \neq \mathbf{Z}[\sqrt{d}]$. The case $d \equiv 1 \pmod 8$ presents a new difficulty because you can't rule only solutions to $y^2 = x^3 + d$ with x even. You will run in to similar complications if you don't assume d is squarefree. We will discuss what happens for $d > 0$ later.

Theorem 27.2. *Suppose $d < 0$ squarefree with $d \equiv 2, 3 \pmod 4$. Then if there exists $a \in \mathbf{Z}$ with $3a^2 + d = \pm 1$, then*

$$(x, y) = (a^2 - d, \pm a(a^2 + 3d))$$

is a solution to the Diophantine equation

$$y^2 = x^3 + d.$$

If moreover $3 \nmid \#\mathrm{Cl}(\mathcal{O}_d)$ then these are the only solutions.

This theorem covers the cases $d = -1, -2$ which you might have seen in elementary number theory, as well as the case $d = -5$ we considered above. Here are some other examples:

Example 27.3. We consider $d = -13$. In this case $a = \pm 2$ works since $3(\pm 2)^2 - 13 = -1$. We find the solution

$$(x, y) = (4 + 13, \pm 2(2^2 - 3 \cdot 13)) = (17, \pm 70)$$

to the Diophantine equation

$$y^2 = x^3 - 13.$$

Is this the only solution? Well, we need to determine $\mathrm{Cl}(\mathcal{O}_{-13})$.

We compute $C_K = \frac{2}{\pi}\sqrt{4 \cdot 13} < 5$. $-13 \equiv 3 \pmod 4$ so $2 = \mathfrak{p}_2^2$ ramifies. $-13 \equiv 2 \pmod 3$ so 3 is inert. Moreover $\mathfrak{p}_2$ cannot be principal because $2 = x^2 + 13y^2$ has no solutions. Thus $\mathrm{Cl}(\mathcal{O}_{-13}) = \{1, [\mathfrak{p}_2]\} = \mathbf{Z}/2\mathbf{Z}$. In particular it has no element of order 3, so the solution we found above is unique.

Example 27.4. We consider $d = -26$. In this case $a = \pm 3$ works since $3(\pm 3)^2 - 26 = 1$. We find the solution

$$(x, y) = (35, \pm 207)$$

to the Diophantine equation

$$y^2 = x^3 - 26.$$

But we also spot another solution:

$$(x, y) = (3, \pm 1).$$

Hence we must have $3 \mid \#\mathrm{Cl}(\mathcal{O}_{-26})$.

Exercise 27.5. Show that $\mathrm{Cl}(\mathcal{O}_{-26})$ has order 6.

You might wonder what happens if we try to study the Diophantine equation $y^2 = x^3 + d$ for $d > 0$ in the same way. We might get to the point of showing that $(y + \sqrt{d}) = (\alpha)^3$ and hence that $y + \sqrt{d} = u\alpha^3$ for $u \in \mathcal{O}_d^\times$. But now we run in to a problem: unlike the case $d < 0$, when $d > 0$, $\mathcal{O}_d^\times$ is infinite! Understanding the unit group $\mathcal{O}_d^\times$, and more generally $\mathcal{O}_K^\times$ for K a number field, will be the last topic in this course.

We finally mention the following theorem, whose proof is a bit beyond the scope of this course:

Theorem 27.6. *For $d \neq 0$, the Diophantine equation $y^2 = x^3 + d$ has only finitely many solutions for $x, y \in \mathbf{Z}$.*

Remark 27.7. By contrast, if you consider $x, y \in \mathbf{Q}$ then very often there will be infinitely many solutions. Take the course on elliptic curves to learn more!

28 Lecture 28: Units in real quadratic fields

Our basic problem is: given a number field K, what can we say about the group of units $\mathcal{O}_K^\times$?

We know something about this, going back to the first example sheet:

Example 28.1. If $d < 0$ is squarefree then we have computed $\mathcal{O}_d^\times$ on example sheet 1. What we found is that $\mathcal{O}_d^\times = \{\pm 1\}$ unless $d = -1, -3$, while $\mathcal{O}_{-1}^\times = \{\pm 1, \pm i\}$ and $\mathcal{O}_{-3}^\times = \{\zeta_6^k \mid 0 \leq k < 6\}$. In particular, the unit group is finite.

By contrast we have seen that the situation is totally different for $d > 1$ squarefree:

Example 28.2. We see that $1+\sqrt{2} \in \mathcal{O}_2^\times$ as $N(1+\sqrt{2}) = -1$, and hence $(1+\sqrt{2})(-1+\sqrt{2}) = 1$. As $1 + \sqrt{2} > 1$, we have $\{(1 + \sqrt{2})^n \mid n \in \mathbf{Z}\}$ is an infinite set of units in $\mathcal{O}_2$. In fact we will soon prove that

$$\mathcal{O}_2^\times = \{\pm(1 + \sqrt{2})^n \mid n \in \mathbf{Z}\}.$$

Theorem 28.3. *Let $d > 1$ be squarefree. There exists a unique unit $\varepsilon \in \mathcal{O}_d^\times$ with $\varepsilon > 1$ and with the property that*

$$\mathcal{O}_d^\times = \{\pm\varepsilon^n \mid n \in \mathbf{Z}\} \simeq \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}.$$

The unit ε is called the *fundamental unit* of $\mathcal{O}_d$. Here are some examples:

- $1 + \sqrt{2} \in \mathcal{O}_2^\times$.
- $2 + \sqrt{3} \in \mathcal{O}_3^\times$.
- $\frac{1+\sqrt{5}}{2} \in \mathcal{O}_5^\times$.
- $170 + 39\sqrt{19} \in \mathcal{O}_{19}^\times$.
- $2143295 + 221064\sqrt{94} \in \mathcal{O}_{94}^\times$.

The examples above show that the fundamental unit can be quite large!

Remark 28.4. In elementary number theory you may have studied the fundamental units in $\mathbf{Z}[\sqrt{d}]^\times$, which are not the same as the fundamental unit of $\mathcal{O}_d^\times$ when $d \equiv 1 \pmod 4$. For example when $d = 5$, the fundamental unit of $\mathcal{O}_5^\times$ is $\varepsilon = \frac{1+\sqrt{5}}{2}$, while the fundamental unit of $\mathbf{Z}[\sqrt{5}]^\times$ is $\varepsilon^3 = 2 + \sqrt{5}$.

When $d \not\equiv 1 \pmod 4$, the units are just solutions to the Diophantine equation

$$x^2 - dy^2 = \pm 1$$

Which is known as Pell's equation. The theorem asserts that non trivial solutions (i.e. other than $(\pm 1, 0)$) to this equation always exist.

We note that we might have $N(\varepsilon) = 1$ or -1. In the former case, all elements of $\mathcal{O}_d^\times$ have norm 1 so the equation $x^2 - dy^2 = -1$ has no solutions. Sometimes this can be see for reasons of congruences: if $p \mid d$ with $p \equiv 3 \pmod 4$ then this equation has no solutions mod p.

The proof of the theorem will proceed in three steps. The first step is to show that there exists a unit $u \in \mathcal{O}_d^\times$ with $u \neq \pm 1$. In some sense this is the most important step. We will use Minkowski's convex body theorem again. Rather than use it to directly produce units, we will first prove the following statement:

- There exists a constant C so that there are infinitely many $\alpha \in \mathcal{O}_d$ with $|N(\alpha)| < C$.

To see why this is implies the existence of a non trivial unit, we consider the ideals (α). We have $\|(\alpha)\| = |N(\alpha)| < C$. But now by Lemma 21.8 there are only finitely many ideals $I \subset \mathcal{O}_d$ with $\|I\| < C$. Hence by the pigeonhole principal, there must be at least three (actually even infinitely many!) different elements $\alpha, \beta, \gamma \in \mathcal{O}_d$ with $(\alpha) = (\beta) = (\gamma)$. Then at least one of β, γ must not be $-\alpha$, say β. Then $\beta \neq \pm\alpha$, but $\beta = u\alpha$ for $u \in \mathcal{O}_d^\times$, and so $u \neq \pm 1$ is the sought after unit!

Now we explain how to produce the infinitely many elements α. We recall the map

$$\begin{aligned} \iota : K &\to \mathbf{R}^2 \\ \alpha &\mapsto (\alpha, \overline{\alpha}) \end{aligned}$$

If we define the function $N : \mathbf{R}^2 \to \mathbf{R}$ by $N((x,y)) = |xy|$ then we have for $\alpha \in \mathcal{O}_d$,

$$N(\alpha) = \alpha\overline{\alpha} = N(\iota(\alpha)).$$

We recall that the region $N(x,y) = |xy| < C$ in $\mathbf{R}^2$ is bounded by two hyperbolas. It has infinite volume but it isn't convex! We also recall that $\iota(\mathcal{O}_d) \subseteq \mathbf{R}^2$ is a lattice with covolume $\sqrt{D}$ (see Lemma 23.6). We pick any $C > \sqrt{D}$ (unlike in the proof of the Minkowski bound, choosing the optimal value won't be important here!)

We consider the rectangle

$$S_r = \{(x,y) \in \mathbf{R}^2 \mid |x| < r, |y| < \frac{C}{r}\}$$

It has volume $(2r)(2\frac{C}{r}) = 4C$. Hence $\mathrm{vol}(S_r) > 2^2\mathrm{covol}(\iota(\mathcal{O}_d))$ and so Minkowski's theorem applies: we conclude that for any $r > 0$ there exists $0 \neq \alpha \in \mathcal{O}_d$ with $\iota(\alpha) \in S_r$. In other words, $|\alpha| < r$ and $|\overline{\alpha}| < \frac{C}{r}$ and hence $|N(\alpha)| = |\alpha\overline{\alpha}| < C$.

Now we pick $0 \neq \alpha_1 \in \mathcal{O}_d \cap S_1$, $0 \neq \alpha_2 \in \mathcal{O}_d \cap S_{|\alpha_1|}$, $0 \neq \alpha_3 \in \mathcal{O}_d \cap S_{|\alpha_2|}$, and so in. Proceeding in this way we construct a sequence $\alpha_1, \alpha_2, \ldots$ of nonzero elements of $\mathcal{O}_d$ satisfying:

$$|\alpha_1| > |\alpha_2| > |\alpha_3| > \cdots$$

and $|N(\alpha_i)| < C$. We have thus found infinitely many elements $\alpha \in \mathcal{O}_d$ with $|N(\alpha)| < C$.

The second step is to prove that there exists a smallest $\varepsilon \in \mathcal{O}_d^\times$ with $\varepsilon > 1$. By what we have proved, there exists a $u \in \mathcal{O}_d^\times$ with $u > 1$. Indeed starting with $u \in \mathcal{O}_d^\times$, $u \neq 1$, one of $u, u^{-1}, -u, -u^{-1}$ will be > 1. But we still need to prove that a smallest such unit exists (e.g. we need to rule out the possibility that there is a sequence of units converging to 1 from above. We have to use something about $\mathcal{O}_d^\times$ to prove this, as this could happen in $\mathcal{O}_K^\times$ for K a cubic field!)

We will prove the following lemma, which will also allow us to verify that we have really found the fundamental unit in examples:

Lemma 28.5. *1. Let $u = a + b\sqrt{d} \in \mathcal{O}_d^\times$. Then $u > 1$ if and only if $a, b > 0$.*

2. Let $u = a + b\sqrt{d}, u' = a' + b'\sqrt{d} \in \mathcal{O}_d^\times$ with $u, u' > 1$. Then $u \geq u'$ if and only if $a \geq a'$ and $b \geq b'$. In fact $u \geq u'$ if and only if $b \geq b'$ with one exception: $d = 5$, $u = \frac{1+\sqrt{5}}{2}$ and $u' = u^2 = \frac{3+\sqrt{5}}{2}$.

Proof. If $a, b > 0$ then $a, b \geq \frac{1}{2}$ so $u = a + b\sqrt{d} \geq \frac{1+\sqrt{d}}{2} > 1$. For the converse, we note that $u \in \mathcal{O}_d^\times$ implies $N(u) = u\overline{u} = \pm 1$. Thus $u > 1$ implies $|\overline{u}| < 1$. Now we have formulas

$$u + \overline{u} = a + b\sqrt{d} + a - b\sqrt{d} = 2a$$
$$u - \overline{u} = a + b\sqrt{d} - (a - b\sqrt{d}) = 2\sqrt{d}b$$

Now $u > 1$ and $|\overline{u}| < 1$ implies that $u + \overline{u}, u - \overline{u} > 0$ hence $a, b > 0$.

For the second part of the lemma, using the above formula for $u - \overline{u}$ and the same formula for $u' - \overline{u}'$ we find

$$u - u' = 2\sqrt{d}(b - b') + \overline{u} - \overline{u'}.$$

If $b \neq b'$ then: if $d \not\equiv 1 \pmod 4$ $|b - b'| \geq 1$ and so $2\sqrt{d}|b - b'| > 2$ and if $d \equiv 1 \pmod 4$ then $|b - b'| \geq \frac{1}{2}$ and so $2\sqrt{d}|b - b'| \geq \sqrt{d} > 2$ since then $d \geq 5$. Either way we conclude if $b \neq b'$ then $2\sqrt{d}|b - b'| > 2$. Since $|\overline{u} - \overline{u'}| < 2$ this implies that if $b > b'$ then $u > u'$ while if $b < b'$ then $u < u'$.

It remains to consider what happens when $b = b'$ of course then $u \geq u'$ if and only if $a \geq a'$, but in fact we show that this implies that $u = u'$ except in the one exception in the statement of the lemma. Since $a^2 - db^2 = \pm 1$ and $(a')^2 - db^2 = \pm 1$ and $a, a' > 0$, after possibly switching u, u' we must have $a^2 - db^2 = 1$ and $(a')^2 - db^2 = -1$, or $a^2 - (a')^2 = 2$. If $b \in \mathbf{Z}$ then $a, a' \in \mathbf{Z}$ and this is impossible: $(n + k)^2 - n^2 = 2nk + k^2 > 2$ for $n, k \geq 1$. If $b \in \frac{1}{2} + \mathbf{Z}$ then $a, a' \in \frac{1}{2}\mathbf{Z}$ and we see that $\frac{(2n+2k+1)^2-(2n+1)^2}{4} = (2n + 1)k + k^2 = 2$ has one solution with $n \geq 0$, $k \geq 1$ namely $n = 0$ $k = 1$. So we must have $a = \frac{1}{2}$, $b = \frac{3}{2}$. But then $\frac{1}{4} - db^2 = -1$ implies $= 5$ and $b = \frac{1}{2}$. □

Using the lemma, we see that we can just take $\varepsilon = a + b\sqrt{d} \in \mathcal{O}_d^\times$ where $a, b > 0$ and b is as small as possible. This achieves the second step.

Now in the third step we will show that

$$\mathcal{O}_d^\times = \{\pm\varepsilon^n \mid n \in \mathbf{Z}\}.$$

To prove this, start with a unit $u \in \mathcal{O}_d^\times$. We can consider $u' = \pm u$ so that $u' > 0$. Then we have $\varepsilon^k < u'$ for $k \ll 0$ and $\varepsilon^k > u'$ for $k \gg 0$ so there is a largest k with $\varepsilon^k \leq u'$ and hence $u' < \varepsilon^{k+1}$. Hence we have

$$1 \leq \frac{u'}{\varepsilon^k} < \varepsilon$$

But by the definition of ε, we cannot have $\frac{u'}{\varepsilon^k} > 1$ so we must have $\frac{u'}{\varepsilon^k} = 1$. In other words $u = \pm\varepsilon^k$.

We give an example of how to provably find the fundamental unit.

Example 28.6. We find the fundamental unit of $\mathcal{O}_{11}$. By the lemma, we just need to find the solution to $a^2 - db^2$ with the smallest value of b. We try:

$$a^2 - 11 \cdot 1 = \pm 1$$
$$a^2 - 11 \cdot 4 = \pm 1$$
$$a^2 - 11 \cdot 9 = \pm 1$$

and the first two have no solutions, but then $a^2 = 99 + 1 = 100$ has a solution. So $\varepsilon = 10 + 3\sqrt{11}$ is the fundamental unit.

We find the fundamental unit of $\mathcal{O}_{13}$. Now $13 \equiv 1 \pmod 4$ so we look for solutions to

$$N(\frac{a+b\sqrt{13}}{2}) = \pm 1$$

or

$$a^2 - 13b^2 = \pm 4$$

with b as small as possible. In this case there is a solution with $b = 1$: $3^2 - 13 \cdot 1^2 = -4$ so $\frac{3+\sqrt{13}}{2}$ is the fundamental unit. If we looked for solutions to

$$a^2 - 13b^2 = \pm 1$$

instead, we would have found $18 + 5\sqrt{13} = \left(\frac{3+\sqrt{13}}{2}\right)^3$.

29 Lecture 29: More on units

When we try to compute $\mathrm{Cl}(\mathcal{O}_d)$ we need to be able to determine if $N(\alpha) = n$ has any solutions for $\alpha \in \mathcal{O}_d$. In other words we need to solve the Diophantine equation

$$a^2 - db^2 = n$$

for $a, b \in \mathbf{Z}$ or also $a, b \in \frac{1}{2} + \mathbf{Z}$ when $d \equiv 1 \pmod 4$. In the imaginary case $d < 0$, this is easy as we can bound a, b: for instance, assuming $n \geq 0$ as otherwise there are no solutions,

$$-db^2 \leq n$$

and so $|b| \leq \sqrt{\frac{n}{-d}}$.

This breaks down in the real case $d > 0$, as there may exist solutions to

$$a^2 - db^2 = n$$

with $|a|, |b|$ arbitrarily large. If this equation does have a solution, then of course we could search for it and eventually find it. But if it does not have a solution, then for the moment we have no way to prove this aside from using congruences, and congruences will not always suffice.

Instead we use our knowledge of the units in $\mathcal{O}_d$. Let $\varepsilon \in \mathcal{O}_d$ be the fundamental unit. Write $\eta = \varepsilon$ of $N(\varepsilon) = 1$ and $\eta = \varepsilon^2$ if $N(\varepsilon) = -1$. Then $\eta \in \mathcal{O}_d^\times$ is the smallest unit with $\eta > 1$ and $N(\eta) = 1$.

Now the key observation is that if $N(\alpha) = n$, then $N(\pm\eta^k\alpha) = n$ for any $k \in \mathbf{Z}$. Using this, we replace α with an element of the for $\pm\eta^k\alpha$ in order to ensure that

$$\sqrt{|n|}\eta^{-1/2} \leq \alpha \leq \sqrt{|n|}\eta^{1/2}.$$

With this choice we have

$$|\overline{\alpha}| = \frac{|n|}{\alpha} \leq \sqrt{|n|}\eta^{1/2}$$

and hence

$$|\alpha - \overline{\alpha}| \leq 2\sqrt{|n|}\eta^{1/2}.$$

On the other hand, if $\alpha = a + b\sqrt{d}$ then

$$\alpha - \overline{\alpha} = a + b\sqrt{d} - (a - b\sqrt{d}) = 2b\sqrt{d}$$

and hence obtain

$$|b| \leq \sqrt{\frac{|n|\eta}{d}}.$$

To summarize we have proved:

Proposition 29.1. *Any solution to $N(\alpha) = n$ for $\alpha \in \mathcal{O}_d$ has the form $\alpha = \pm\eta^k(a+b\sqrt{d})$ where $a + b\sqrt{d} \in \mathcal{O}_d$, $N(a + b\sqrt{d}) = n$ and $|b| \leq \sqrt{\frac{|n|\eta}{d}}$.*

Using this, we can now prove that no solution to $N(\alpha) = n$ exists when congruences don't work.

Remark 29.2. Some improvements to the proposition are possible: if $n > 0$ then $\overline{\alpha} = \frac{n}{\alpha} > 0$ and so $|\alpha - \overline{\alpha}| \leq \sqrt{|n|\eta^{1/2}}$ and we obtain the bound $|b| \leq \frac{1}{2}\sqrt{\frac{|n|\eta}{d}}$. Also if we are interested in solving the equation $N(\alpha) = \pm n$, as we often are when computing class groups, and $N(\varepsilon) = -1$, we may use ε rather than η and obtain the bound $|b| \leq \sqrt{\frac{|n|\varepsilon}{d}}$.

We note however that when ε is very large, this upper bound on $|b|$ is also very large and thus not very practical.

Example 29.3. We compute the class group of $\mathcal{O}_{79}$. The Minkowski bound is $C_K = \frac{\sqrt{4\cdot 79}}{2} < 9$.

2 ramifies as $79 \equiv 3 \pmod 4$. We also compute $79 \equiv 1 \pmod 3$, $79 \equiv 4 \pmod 5$, $79 \equiv 2 \pmod 7$ are all squares, so $3, 5, 7$ all split. We feel very unlucky!

We start searching for norms. We spot the following, using the squares that are closest to 79.

$$\begin{aligned} 9^2 - 79 &= 2 \\ 10^2 - 79 &= 21 \\ 8^2 - 79 &= -15 \end{aligned}$$

Suddenly we feel very lucky! The first of these tells us that the ideal $\mathfrak{p}_2$ lying above 2 is principal. The second tells us there is an ideal of norm 21, which must factor as an ideal of norm 3 and an ideal of norm 7, so we get a relation in the class group $[\mathfrak{p}_3]^{\pm 1}[\mathfrak{p}_7]^{\pm 1} = 1$, and hence $[\mathfrak{p}_7] = [\mathfrak{p}_3]^{\pm 1}$. Similarly the fact that -15 is a norm tells us that $[\mathfrak{p}_5] = [\mathfrak{p}_3]^{\pm 1}$. Hence the class group is generated by $[\mathfrak{p}_3]$. With a bit more work we also spot with some more work[76]:

$$17^2 - 2^2 \cdot 79 = -27.$$

Hence $(17 + 2\sqrt{79}) = \mathfrak{p}_3^3$ or $\overline{\mathfrak{p}}_3^3$ (it cannot be say $\mathfrak{p}_3^2\overline{\mathfrak{p}}_3$ because then we would have $17 + 2\sqrt{79} \in \mathfrak{p}_3\overline{\mathfrak{p}}_3 = (3)$ which is false!).

At this point the only question left is: is $\mathfrak{p}_3$ principal? In other words, can we solve $a^2 - 79b^2 = \pm 3$? We try congruences first. Working mod 79, we have $\left(\frac{3}{79}\right) = -\left(\frac{79}{3}\right) = -1$, and $\left(\frac{-3}{79}\right) = 1$. Hence we conclude that $a^2 - 79b^2 = 3$ has no solutions mod 79, but $a^2 - 79b^2 = -3$ does. We also try a power of 2, e.g. 8: but $-79 \equiv 1 \pmod 8$ and we have $1^2 + 2^2 = 5 \equiv -3 \pmod 8$. In fact the equation $a^2 - 79b^2 = -3$ has solutions mod N for all $N > 0$.

We must use our new method for showing that $a^2 - 79b^2 = -3$ has no solutions. We first find the fundamental unit, which in this case is rather painful: we find $\varepsilon = 80 + 9\sqrt{79}$,

[76] At this point the Minkowski bound tells us that $[\mathfrak{p}_3]$ has order ≤ 3 so we should be looking for elements of norm $3, 9, 27$

and $N(\varepsilon) = 1$. Hence we conclude that if $a^2 - 79b^2 = -3$ has a solution, it has a solution with

$$|b| \leq \sqrt{\frac{3\varepsilon}{79}} < 3$$

Thus we only have to check that $-3 + 79 = 76$ and $-3 + 4 \cdot 79 = 313$ are not squares to see that $a^2 - 79b^2 = -3$ has no solutions.

We conclude that $\mathfrak{p}_3$ is not principal, and hence $\mathrm{Cl}(\mathcal{O}_{79}) = \mathbf{Z}/3\mathbf{Z}$, generated by $[\mathfrak{p}_3]$.

29.1 Dirichlet's Unit theorem

We finally say something about units in more general number fields. We state but don't prove the following theorem:

Theorem 29.4 (Dirichlet's Unit theorem)**.** *Let K be a number field. Suppose K has r real embeddings and s pairs of conjugate complex embeddings. Let $\mu(K) = \{x \in K \mid x^n = 1, \text{ for some } n > 0\}$ be the set of roots of unity in K. Then*

$$\mathcal{O}_K^\times \simeq \mu(K) \times \mathbf{Z}^{r+s-1}.$$

In other words there exists units $\varepsilon_1, \ldots, \varepsilon_{r+s-1} \in \mathcal{O}_K^\times$ such that any unit $u \in \mathcal{O}_K^\times$ can be written uniquely in the form $u = \zeta\varepsilon_1^{n_1} \cdots \varepsilon_{r+s-1}^{n_{r+s-1}}$ for $\zeta \in \mu(K)$ a root of unity and $n_1, \ldots, n_{r+s-1} \in \mathbf{Z}$.

Remark 29.5. Usually the group of roots of unity $\mu(K)$ in a number field K is just $\{\pm 1\}$. For instance if K has a real embedding $K \to \mathbf{R}$, then since the only roots of unity in $\mathbf{R}$ are $\{\pm 1\}$ the same will be true of K.

Regardless $\mu(K)$ is always a finite cyclic group.

Example 29.6. If $K = \mathbf{Q}$, $r = 1$ and $s = 0$ so $r + s - 1 = 0$ and we have $\mathbf{Z}^\times = \{\pm 1\} = \mu(\mathbf{Q})$.

If K is an imaginary quadratic field $r = 0$, $s = 1$ so $r + s - 1 = 0$ and we have $\mathcal{O}_K^\times = \mu(K)$, which is $\{\pm 1\}$ unless $K = \mathbf{Q}(i)$ or $K = \mathbf{Q}(\zeta_3)$.

If K is a real quadratic field $r = 2$, $s = 0$ so $r + s - 1 = 1$ and Dirichlet's theorem says that $\mathcal{O}_K^\times = \{\pm\varepsilon^n\}$ confirming what we proved last time.

If K is a cubic field, then either $r = 1$, $s = 1$ and $r + s - 1 = 1$ and again we have $\mathcal{O}_K^\times = \{\pm\varepsilon^n\}$ for some unit ε, or $r = 3$ and $s = 0$ and $r + s - 1 = 2$ and in this case we have $\mathcal{O}_K^\times = \{\pm\varepsilon_1^{n_1}\varepsilon_2^{n_2}\}$.

There is one more case where $r + s - 1 = 1$, namely when K has degree 4 and $r = 0$, $s = 2$.

In general, it is an extremely difficult problem to compute the group of units in a given number field K. In fact it can be extremely difficult to even find a single non trivial unit! Like our proof of the existence of a non trivial unit in the real quadratic case, the proof of Dirichlet's theorem is based on the pigeonhole principal and Minkowski's convex body theorem, and doesn't really help us find units directly!

Some typical strategies are:

1. Get lucky and "spot" some units directly.
2. Go searching for elements $\alpha \in K$ of small norm (this can already be hard!) and try to combine them to find a unit. E.g. if you found two elements $\alpha, \beta \in K$ with $N_K(\alpha) = N_K(\beta) = 3$ you might hope that $(\alpha) = (\beta)$ so that α/β is a unit.

3. Take advantage of any subfields of K to find some units/elements of small norm!

Even once you've found $r+s-1$ "independent" units, it is still extremely difficult in general to prove that you've found them all.[77]

Example 29.7. We consider $K = \mathbf{Q}(\sqrt[3]{2})$. This is a cubic field with $r = 1$ $s = 1$, so we expect the unit group to have rank 1 by Dirichlet's theorem. In this case it is not too hard to write out: $N_K(a + b\sqrt[3]{2} + c\sqrt[3]{4}) = a^3 + 2b^3 + 4d^3 - 6abc$ and we can just "spot" a solution to $N_K(\alpha) = \pm 1$: $N_K(1 + \sqrt[3]{2} + \sqrt[3]{4}) = 1$. It turns out that

$$\mathbf{Z}[\sqrt[3]{2}]^\times = \{\pm(1 + \sqrt[3]{2} + \sqrt[3]{4})^n\}.$$

Example 29.8. We consider $K = \mathbf{Q}(\zeta_5)$. In this case $r = 0$ $s = 2$ so we expect the unit group to have rank $r + s - 1 = 1$. But $\mathbf{Q}(\zeta_5)$ contains the real quadratic field $\mathbf{Q}(\sqrt{5})$, and we know

$$\mathcal{O}_5^\times = \left\{ \pm \left(\frac{1+\sqrt{5}}{2} \right)^n \right\} \subseteq \mathbf{Z}[\zeta_5]^\times.$$

By Dirichlet's theorem, this must in fact be a finite index subgroup, so we have already found "most" of the units! In fact the full unit group is:

$$\mathbf{Z}[\zeta_5]^\times = \left\{ \zeta_{10}^k \left(\frac{1+\sqrt{5}}{2} \right)^n \right\}.$$

We won't prove Dirichlet's theorem in this course. The proof is a fairly straightforward generalization of the strategy of Theorem 28.3, with the new difficulties being mostly a matter of bookkeeping! Instead we just give a hint of where the number $r+s-1$ comes from.

The idea is to consider all the embeddings $\tau_1, \ldots, \tau_n : K \to \mathbf{C}$ and think about what the relations are between the absolute values $|\tau_i(\alpha)|$ for $\alpha \in \mathcal{O}_K^\times$ a unit.[78] If τ_i is a complex embedding then we have the relation $|\tau_i(\alpha)| = |\overline{\tau}_i(\alpha)|$. Also because α is a unit, we have

$$1 = |N_K(\alpha)| = |\tau_1(\alpha)| \cdots |\tau_n(\alpha)|$$

Thus there are s relations coming from complex conjugations, and 1 relation coming from the norm, for a total of $s+1$ relations. Since there are $n = r + 2s$ embeddings total, we can expect $r + 2s - (s+1) = r + s - 1$ total independent units.

Exercise 29.9. Let K be a cubic field with three real embeddings $\tau_1, \tau_2, \tau_3 : K \to \mathbf{R}$. Using Minkowski's convex body theorem, show that there exists a constant C and a sequence $\alpha_1, \alpha_2, \ldots$ of nonzero elements of $\mathcal{O}_K$ with

$$|\tau_2(\alpha_1)| > |\tau_2(\alpha_2)| > |\tau_2(\alpha_3)| > \cdots$$

and

$$|\tau_3(\alpha_1)| > |\tau_3(\alpha_2)| > |\tau_3(\alpha_3)| > \cdots$$

Deduce that there exists a unit $u_1 \in \mathcal{O}_K^\times$ with $|\tau_2(u_1)|, |\tau_3(u_1)| < 1$.

Make the same argument to find a unit $u_2 \in \mathcal{O}_K^\times$ with $|\tau_1(u_2)|, |\tau_3(u_2)| < 1$. Show that u_1, u_2 are independent units in the senes that $u_1^{n_1} \neq u_2^{n_2}$ unless $n_1 = n_2 = 0$.

[77] One tool here, which is slightly beyond the scope of our course, is the "class number formula".

[78] We know from a problem on example sheet one that if $|\tau_i(\alpha)| = 1$ for all i then α is a root of unity.

30 Lecture 30: Example class 5

1. In this problem you will investigate which primes are of the form $x^2 + 13y^2$ for $x, y \in \mathbf{Z}$.

   (a) Show that $\mathrm{Cl}(\mathcal{O}_{-13}) = \{1, [\mathfrak{p}_2]\}$. (We did this in lecture but do it again!)

   (b) Show that p splits in $\mathcal{O}_{-13}$ if and only if $p \equiv 1 \pmod 4$ and $\left(\frac{p}{13}\right) = 1$ or $p \equiv 3 \pmod 4$ and $\left(\frac{p}{13}\right) = -1$.

   (c) Show that in the first case, $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ where $\mathfrak{p}, \overline{\mathfrak{p}}$ are principal, while in the second case, $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ where $[\mathfrak{p}] = [\overline{\mathfrak{p}}] = [\mathfrak{p}_2]$.

   (d) Show that $p = x^2 + 13y^2$ if and only if $p = 13$ or $p \equiv 1 \pmod 4$ and $\left(\frac{p}{13}\right) = 1$.

2. Consider $\mathcal{O}_{-14}$ and the non principal ideal $\mathfrak{p}_2 = (2, \sqrt{-14})$.

   (a) Show that there exists an ideal $\mathfrak{p} \subseteq \mathcal{O}_K$ with $\|\mathfrak{p}\| = p$ and $[\mathfrak{p}] = [\mathfrak{p}_2]$ if and only if $p = 2x^2 + 7y^2$ for some $x, y \in \mathbf{Z}$ (Hint: you might first try to show this is equivalent to $2p = x^2 + 14y^2$.)

   (b) Show that $x^2 + 14y^2$ and $2x^2 + 7y^2$ take the same values mod 8 and mod 7. (If you'd like you can try to show they take the same values mod N for all $N > 0$.) As a consequence we do not see any way to distinguish between these two cases with a congruence condition on p.

3. Let K be a number field and suppose $\alpha\beta = \gamma^n$ for $\alpha, \beta, \gamma \in \mathcal{O}_K$ and $n > 0$. Suppose that α, β are coprime, i.e. $(\alpha, \beta) = \mathcal{O}_K$. Show that if n is coprime to $\#\mathrm{Cl}(\mathcal{O}_K)$ then $\alpha = u\delta^n$ for some $u \in \mathcal{O}_K^\times$.

4. The method we used in class to solve $y^2 = x^3 + d$ didn't work when $d > 0$. In this problem you will use a clever trick to show that $y^2 = x^3 + 7$ has no solutions. The trick is to write the equation as

$$y^2 + 1 = x^3 + 8 = (x + 2)(x^2 - 2x + 4).$$

   (a) Show that x must be odd.

   (b) Show that $x^2 - 2x + 4 \equiv 3 \pmod 4$ and $x^2 - 2x + 4 > 0$.

   (c) Deduce that there is an odd prime $p \equiv 3 \pmod 4$ with $p \mid x^2 - 2x + 4$ and hence $p \mid y^2 + 1$.

   (d) Conclude that the Diophantine equation has no solutions.

5. Show that $(x, y) = (1, 0)$ is the only solution to the Diophantine equation $y^2 = x^5 - 1$ with $x, y \in \mathbf{Z}$.

6. Show that $3 \mid \#\mathrm{Cl}(\mathcal{O}_{-109})$ by spotting a solution to $y^2 = x^3 - 109$.

7. Find the fundamental units of $\mathcal{O}_6, \mathcal{O}_7, \mathcal{O}_{10}$.