MATH7421
Number Theory Part 2:
Introduction to Algebraic Number Theory
2026/27

Dr. Dan Fretwell
(d.fretwell@lancaster.ac.uk)

Lancaster University

Contents

1 Introduction: Solving Integer Problems using Commutative Rings (1 lecture) 1

1.1 Fermat's challenge 1

1.2 Pythagorean Triples 2

1.3 Sums of two squares 3

1.4 Pell's Equation 4

1.5 Fermat's Last Theorem 4

2 Polynomials, Roots and Irreducibility (1 lecture) 5

3 Number Fields and their Rings of Integers (3 lectures) 9

3.1 Number Fields 9

3.2 Conjugates, Norms, Traces and Discriminants. 12

3.3 The Ring of Integers $\mathcal{O}_K$ 16

4 Factorization in Rings of Integers (3 lectures) 20

4.1 Existence of factorisation into irreducibles 20

4.2 Unique Factorisation Domains 24

4.3 Euclidean Domains 28

5 Dirichlet's Unit Theorem (1 lecture) 32

5.1 Pell's equation 34

6 Fermat's Last Theorem (1 lecture) 36

1 Introduction: Solving Integer Problems using Commutative Rings (1 lecture)

Algebraic Number Theory is the branch of Number Theory that uses tools and techniques from Algebra to solve/understand integer problems. To get us going...let's see a few examples!

1.1 Fermat's challenge

A major theme in Number Theory is to find integer solutions to polynomial equations. Such equations are often called **Diophantine equations** and have a rich history. A simple example is the linear equation $36x+47y = 1$, which you already know how to solve (using Euclid's Algorithm).

For centuries, Number Theorists have challenged each other to solve tougher and tougher Diophantine equations. Here's one that Fermat would challenge his Mathematician friends with:

Theorem 1.1. *The only integer solutions to the equation $y^2 = x^3 - 2$ are $(x, y) = (3, \pm 5)$.*

The key idea is to use the ring $\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} \,|\, a, b \in \mathbb{Z}\}$, which behaves similarly to $\mathbb{Z}$:

- It has **units**, $\pm 1 \in \mathbb{Z}[\sqrt{-2}]$.
- It has **irreducibles**, non-units $\alpha \in \mathbb{Z}[\sqrt{-2}]$ that can't be factored into a product of non-units.
- Irreducibles $\alpha \in \mathbb{Z}[\sqrt{-2}]$ satisfy the **prime** property, i.e. $\alpha|\beta\gamma$ implies $\alpha|\beta$ or $\alpha|\gamma$.
- It has **unique factorization** into irreducibles, up to reordering and multiplication by units.

These are not obvious facts...but if we assume them for now then we can solve Fermat's challenge!

Proof. Working over $\mathbb{Z}[\sqrt{-2}]$ allows us to write:

$$y^2 + 2 = (y + \sqrt{-2})(y - \sqrt{-2}) = x^3.$$

Suppose that $\alpha \in \mathbb{Z}[\sqrt{-2}]$ is irreducible and satisfies $\alpha|(y + \sqrt{-2})$ and $\alpha|(y - \sqrt{-2})$. Then taking difference gives $\alpha|2\sqrt{-2} = -\sqrt{-2}^3$, so that $\alpha|\sqrt{-2}$.

We claim that $\sqrt{-2}$ is irreducible in $\mathbb{Z}[\sqrt{-2}]$. Suppose that $\sqrt{-2} = \beta\gamma$ is a factorization into non-units. Then taking absolute values gives $2 = |\sqrt{-2}|^2 = |\beta|^2|\gamma|^2$. But then $|\beta|^2 = 1$ or $|\gamma|^2 = 1$, implying that one of β, γ is a unit (the only integer solutions to $x^2 + 2y^2 = 1$ are $(x, y) = (\pm 1, 0)$).

This shows that $\alpha = \pm\sqrt{-2}$. However, we cannot have that $\pm\sqrt{-2}|(y + \sqrt{-2})$, since taking absolute values gives $2|(y^2 + 2)$, implying that $2|y$. But then $y^2 \equiv 0 \bmod 4$, so that $x^3 \equiv 2 \bmod 4$ (which has no solutions). So α cannot exist, hence $(y + \sqrt{-2})$ and $(y - \sqrt{-2})$ are coprime.

By unique factorization, if two coprime elements of $\mathbb{Z}[\sqrt{-2}]$ multiply to give a cube then each must be a unit multiplied by a cube. Since the units in this ring are ± 1, which are both cubes, we can assume that $(y + \sqrt{-2})$ is a cube, so that there exists $s, t \in \mathbb{Z}$ satisfying:

$$y + \sqrt{-2} = (s + t\sqrt{-2})^3 = s^3 + 3s^2t\sqrt{-2} - 6st^2 - 2t^3\sqrt{-2} = s(s^2 - 6t^2) + t(3s^2 - 2t^2)\sqrt{-2}.$$

We must then have that $1 = t(3s^2 - 2t^2)$, so that $t = \pm 1$ and $s = \pm 1$. This gives that $y = s(s^2 - 6t^2) = \pm 5$, and finally that $x = 3$. Hence the only solutions are $(x, y) = (3, \pm 5)$. □

The equation above is an example of an **Elliptic Curve**. Such curves are highly useful in **Cryptography** and are also the subject of the **Birch Swinnerton-Dyer Conjecture**, one of the most famous unsolved problems of our time (one of the seven Millenium Problems).

1.2 Pythagorean Triples

You've known for some time that there is a right angled triangle with side lengths $(3, 4, 5)$. You might even know that there is one with side lengths $(5, 12, 13)$. But how do we find **all** right angled triangles with integer side lengths?

Definition 1.2. *A **Pythagorean Triple** is a triple $(x, y, z) \in \mathbb{N}_{>0}$ satisfying $x^2 + y^2 = z^2$.*

To find all Pythagorean Triples we can use the ring $\mathbb{Z}[i] = \{a + ib \,|\, a, b, \in \mathbb{Z}\}$, which behaves like $\mathbb{Z}$:

- It has **units**, $\pm 1, \pm i \in \mathbb{Z}[i]$.
- It has **irreducibles**, non-units $\alpha \in \mathbb{Z}[i]$ that can't be factored into a product of non-units.
- Irreducibles $\alpha \in \mathbb{Z}[i]$ satisfy the **prime** property, i.e. $\alpha | \beta\gamma$ implies $\alpha | \beta$ or $\alpha | \gamma$.
- It has **unique factorization** into irreducibles, up to reordering and multiplication by units.

Again, none of this is obvious...but assuming these facts lets us solve the problem!

Theorem 1.3. *Every Pythagorean Triple can be written in the form (up to swapping x and y):*

$$(x, y, z) = (k(s^2 - t^2), \quad 2kst, \quad k(s^2 + t^2)),$$

where k, s, t are positive integers such that $gcd(s, t) = 1$, $s > t > 0$ and $s \not\equiv t \bmod 2$.

Proof. (Sketch) First note that scaling out $\gcd(x, y, z) = k$ reduces to the case $\gcd(x, y, z) = 1$. Working over $\mathbb{Z}[i]$ allows us to write:

$$x^2 + y^2 = (x + iy)(x - iy) = z^2,$$

and a similar argument to the previous section gives that $(x + iy)$ and $(x - iy)$ are coprime (see Exercises).

By unique factorization, if two coprime elements of $\mathbb{Z}[i]$ multiply to give a square then each must be a unit multiplied by a square. Since $-1 = i^2$ and $-i = i^2 = i(i^2)$, we have that:

$$x + iy = (s + it)^2 = (s^2 - t^2) + 2ist \qquad \underline{\text{or}} \qquad x + iy = i(s + it)^2 = 2st + i(s^2 - t^2)$$

for some $s, t \in \mathbb{Z}$.

Swapping x and y if necessary then gives $(x, y, z) = (s^2 - t^2, 2st, s^2 + t^2)$ and the conditions on (s, t) follow from $x > 0$, x being odd and $\gcd(x, y, z) = 1$. □

It is easily checked that all such triples are Pythagorean Triples (see Exercises).

1.3 Sums of two squares

Here's another natural question about right angled triangles. If the two shorter side lengths are integers, then which numbers $\sqrt{n}$ can we see on the hypotenuse? In other words, which positive integers n can be written as $n = x^2 + y^2$?

The key is to note that $n = x^2 + y^2$ if and only if there exists a factorization $n = (x + iy)(x - iy)$ in the ring $\mathbb{Z}[i]$. The case $n = 1 = (\pm 1)^2 + 0^2 = 0^2 + (\pm 1)^2$ corresponds to the unit factorisations $1 = (\pm 1)(\pm 1) = (\pm i)(\mp i)$. So this begs the question...which natural numbers $n \geq 2$ have such a factorisation in $\mathbb{Z}[i]$?

Since n already factors into prime powers in $\mathbb{Z}$, it seems that the primes are a good place to start!

$$\begin{aligned}
2 &= (1+i)(1-i) = 1^2 + 1^2 \\
3 &= \text{no further factorisation} \\
5 &= (1+2i)(1-2i) = 1^2 + 2^2 \\
7 &= \text{no further factorisation} \\
11 &= \text{no further factorisation} \\
13 &= (2+3i)(2-3i) = 2^2 + 3^2 \\
17 &= (1+4i)(1-4i) = 1^2 + 4^2 \\
19 &= \text{no further factorisation} \\
&\vdots
\end{aligned}$$

Theorem 1.4. *(Fermat's Theorem on sums of two squares) A prime p can be written as a sum of two squares if and only if it is reducible in $\mathbb{Z}[i]$. These primes are exactly $p = 2$ or $p \equiv 1 \bmod 4$.*

Proof. Suppose that p is reducible, so that it has a factorisation $p = \alpha\beta$ into non-units $\alpha, \beta \in \mathbb{Z}[i]$. Then $p^2 = |\alpha|^2|\beta|^2$, and the only possibility is $|\alpha|^2 = |\beta|^2 = p$, so that $p = |\alpha|^2 = |a+ib|^2 = a^2 + b^2$ for some $a, b \in \mathbb{Z}$ (i.e. p is a sum of two squares). The converse is true as we saw above.

It now suffices to consider which primes are reducible in $\mathbb{Z}[i]$. We saw above that $2 = (1+i)(1-i)$ is reducible and so we may assume that p is odd. Each prime $p \equiv 3 \bmod 4$ is irreducible in $\mathbb{Z}[i]$, since if $p = x^2 + y^2$ then $x^2 + y^2 \equiv 3 \bmod 4$, which has no solutions (since $x^2, y^2 \equiv 0, 1 \bmod 4$).

We now show that every prime $p \equiv 1 \bmod 4$ is reducible. For such a prime we can find an integer solution to the congruence $x^2 \equiv -1 \bmod p$ (see Exercises). It follows that $p \mid x^2 + 1 = (x+i)(x-i)$. If p is irreducible in $\mathbb{Z}[i]$ then we would have $p \mid (x + i)$ or $p \mid (x - i)$, both of which are false. □

Many more general theorems exist that allow us to understand how prime numbers factor in many other rings like $\mathbb{Z}[i]$.

More generally, we can use the above result to prove the following

Theorem 1.5. *(Sums of two squares) A positive integer n is the sum of two squares if and only if every prime $p \mid n$ such that $p \equiv 3 \bmod 4$ divides n to an even power.*

Proof. See Exercises. □

1.4 Pell's Equation

Ok, so we have a good understanding of equations like $x^2 + y^2 = z^2$ and $x^2 + y^2 = n$. What if we tweak the equation a little and look at say $x^2 - 2y^2 = \pm 1$?

You soon find some integer solutions by playing around:

$$(x, y) = \pm(1, 0), \pm(1, \pm 1), \pm(3, \pm 2), \pm(7, \pm 5), \pm(17, \pm 12), ...$$

but how can we find **all** solutions?

Theorem 1.6. *The integer solutions to $x^2 - 2y^2 = \pm 1$ are given by $(x, y) = \pm(a, \pm b)$, where (a, b) is a term in the recursion $(a_{n+1}, b_{n+1}) = (a_n + 2b_n, a_n + b_n)$ with starting value $(a_0, b_0) = (1, 0)$.*

Proof. (Sketch) We have that $x^2 - 2y^2 = \pm 1$ if and only if $(x + y\sqrt{2})(x - y\sqrt{2}) = \pm 1$, i.e. $x + y\sqrt{2}$ is a unit in the ring $\mathbb{Z}[\sqrt{2}]$. By Dirichlet's Unit Theorem (which we'll see later), the units of this ring are all of the form $\pm(1 + \sqrt{2})^n$ for some $n \in \mathbb{Z}$. The result follows from this (see Exercises). □

Equations such as these are called **Pell equations**, and can be tackled more generally using other (quadratic) rings.

1.5 Fermat's Last Theorem

As we saw earlier, there are infinitely many integer solutions to $x^2 + y^2 = z^2$. One of the most infamous Diophantine equations of all time is the Fermat equation $x^n + y^n = z^n$ (for $n \geq 3$).

Theorem 1.7. *(Fermat's Last Theorem) Let $n \geq 3$. Then the only integer solutions to $x^n + y^n = z^n$ satisfy $xyz = 0$.*

It's fair to say that most of the Algebraic Number Theory we're going to learn about was born out of attempts to solve Fermat's Last Theorem. Indeed, in the 1800's a "proof" of FLT was given by Lamé, using the ring $\mathbb{Z}[\zeta_n]$ to factor as follows (where $\zeta_n = e^{\frac{2\pi i}{n}}$ is an nth root of unity):

$$x^n + y^n = \prod_{0 \leq i \leq n-1} (x + \zeta_n^i y) = z^n.$$

Arguments similar to earlier then finish the proof.

However, Kummer realised that Lamé's proof had a huge flaw...it assumed that $\mathbb{Z}[\zeta_n]$ has unique factorisation into irreducibles (which it doesn't...Kummer gave a counter-example). Much of Algebraic Number Theory was developed in attempts to modify these ideas and make them work. It wouldn't be until the 1990's that Wiles would provide his groundbreaking proof of FLT...ending a 350 year chapter in the history of Mathematics!

Hopefully the above examples convince you that rings like $\mathbb{Z}[\sqrt{-2}], \mathbb{Z}[i], \mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[\zeta_n]$ are highly useful in solving Diophantine equations. However, to unlock these powers it's clear that we need to fully understand such rings, their units and their factorization properties. This is what we will do in this course!

2 Polynomials, Roots and Irreducibility (1 lecture)

Before we get to define fields that behave like $\mathbb{Q}$, we first need to understand a little bit about how polynomial rings over fields.

Recall that a **field** is a commutative ring K such that every element of $K^\times = K\backslash\{0\}$ is a unit (i.e. we can add, subtract, multiply and divide by everything non-zero and the usual rules hold).

Basic examples of fields are $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, but there are many more fields out there!

Definition 2.1. *Let K be a field.*

- *The **polynomial ring over** K in n variables is:*

$$K[x_1, x_2, ..., x_n] = \left\{ \sum_{i_1, i_2, ... i_n \in \mathbb{N}_{\geq 0}} a_{i_1, i_2, ..., i_n} x_1^{i_1} x_2^{i_2} ... x_n^{i_n} \;\middle|\; \begin{array}{c} a_{i_1, i_2, ..., i_n} \in K \text{ and is non-zero for} \\ \text{finitely many tuples } (i_1, i_2, ..., i_n) \end{array} \right\}.$$

- *The **degree** of $f(x_1, x_2, ..., x_n) \in K[x_1, x_2, ..., x_n]\backslash\{0\}$ is*

$$deg(f(x_1, x_2, ..., x_n)) = max \left\{ \sum_{t=1}^{n} i_t \;\middle|\; a_{i_1, i_2, ..., i_n} \neq 0 \right\}.$$

- *If $f(x_1, x_2, ..., x_n) \in K[x_1, x_2, ..., x_n]$ then a **root of** f is a tuple $(a_1, a_2, ..., a_n) \in K^n$ such that $f(a_1, a_2, ..., a_n) = 0$.*

- *If $f(x_1, x_2, ..., x_n) \in K[x_1, x_2, ..., x_n]$ then f **is reducible** if $deg(f) \geq 1$ and there exists a factorisation:*

$$f(x_1, x_2, ..., x_n) = g(x_1, x_2, ..., x_n)\, h(x_1, x_2, ..., x_n)$$

with $g(x_1, x_2, ..., x_n), h(x_1, x_2, ..., x_n) \in F[x_1, x_2, ..., x_n]$ satisfying $deg(g), deg(h) < deg(f)$.

*We say that f **is irreducible** if $deg(f) \geq 1$ and there is no such factorisation.*

Remark 2.2. *If L is a field containing K then we can automatically view any polynomial in $K[x_1, x_2, ..., x_n]$ as a polynomial in $L[x_1, x_2, ..., x_n]$. In this way we can talk about a polynomial having a **root over** L or being **reducible over** L.*

Remark 2.3. *We can also consider polynomial rings $R[x_1, x_2, ..., x_n]$ where R is an arbitrary commutative ring (e.g. $\mathbb{Z}[x]$ is the ring of polynomials in one variable with integer coefficients). However, a lot of results that are true for polynomial rings over fields are simply not true for polynomial rings over commutative rings!*

For example, the polynomial $2x + 4 \in \mathbb{Q}[x]$ is irreducible...but as an element of $\mathbb{Z}[x]$ it is reducible (it has a non-trivial factorisation $2(x + 2)$ into non-units).

For most of this course we will consider polynomial rings in one variable, i.e. $K[x]$ and $R[x]$. These are much easier to study than polynomial rings in more variables.

The natural question now is how to tell whether a polynomial in $K[x]$ is irreducible. A couple of easy tests work in general. The first is to check if the polynomial is irreducible over a bigger field.

Lemma 2.4. *Let $f(x) \in K[x]$ with $\deg(f) \geq 1$ and L be a field containing K. If f is irreducible over L then f is irreducible over K.*

Proof. We prove the contrapositive. Suppose that f is not irreducible over K. Since $\deg(f) \geq 1$ this implies that f is reducible over K, so that there exists a non-trivial factorisation $f = gh$ over K. This factorisation would then provide a non-trivial factorisation over L too, so that f would be reducible over L (i.e. not irreducible over L). □

Example 2.5. *We immediately know that the polynomial $x^2+1 \in \mathbb{Q}[x]$ is irreducible over $\mathbb{Q}$, since it is irreducible over $\mathbb{R}$ (any non-trivial factorisation over $\mathbb{R}$ would require a linear factor, implying that x^2+1 has a real root...which it doesn't!).*

The second test for irreducibility is to check whether the polynomial has a root. Let's unpack this.

The polynomial ring $K[x]$ is an example of a **Euclidean Domain** with respect to the degree map, i.e. if $f(x), g(x) \in K[x]$ and $g(x) \neq 0$ then there exist unique polynomials $q(x), r(x) \in K[x]$ satisfying:

$$f(x) = q(x)g(x) + r(x),$$

with $r = 0$ or $0 \leq \deg(r(x)) < \deg(f(x))$

Remark 2.6. *Be careful, there is no such algorithm for polynomial rings over arbitrary commutative rings! For example, consider the polynomials $f(x) = 3x + 1$ and $g(x) = 2x + 1$ in $\mathbb{Z}[x]$. If we attempt to divide $f(x)$ by $g(x)$ we find that:*

$$3x + 1 = \frac{3}{2}(2x + 1) - \frac{1}{2},$$

so that $q(x) = \frac{3}{2}$ and $r(x) = -\frac{1}{2}$, which are not in $\mathbb{Z}[x]$. The issue is that 2 is not a unit in $\mathbb{Z}$, i.e. $2 \notin \mathbb{Z}^\times = \{\pm 1\}$.

The fact that we have Euclid's Algorithm for polynomial rings over fields allows us to link roots with reducibility.

Lemma 2.7. *Let K be a field, $f(x) \in K[x]$ and $a \in K$. Then $f(x) = (x - a)g(x)$ for some $g(x) \in K[x]$ if and only if $f(a) = 0$.*

Proof. It's clear that if $f(x) = (x - a)g(x)$ then $f(a) = 0$. To prove the converse, use Euclid's Algorithm to write $f(x) = q(x)(x - a) + r(x)$, with $\deg(r(x)) = 0$ (i.e. $r(x) \in K$). The claim then follows since $f(a) = 0 = r(a)$. □

Example 2.8. *The polynomial $f(x) = 2x^2 - \frac{9}{2} \in \mathbb{Q}[x]$ satisfies $f\left(\frac{3}{2}\right) = 0$, so $f(x) = \left(x - \frac{3}{2}\right) g(x)$ for some $g(x) \in \mathbb{Q}[x]$. Indeed $f(x) = \left(x - \frac{3}{2}\right)(2x + 3)$.*

From now on we'll look at the special case of testing irreducibility over $\mathbb{Q}$. First of all, notice that $f(x) \in \mathbb{Q}[x]$ is irreducible if and only if $af(x)$ is irreducible for any $a \in \mathbb{Q}\backslash\{0\}$. Because of this, we can assume that $f(x) \in \mathbb{Z}[x]$ (scaling by the biggest coefficient denominator to make it integral).

There's a very quick method for testing whether an integer polynomial has a rational root.

Lemma 2.9. *(Rational root test) Let $f(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n \in \mathbb{Z}[x]$ have $deg(f(x)) \geq 1$ and $a_0 \neq 0$. If $f(a) = 0$ for some $a = \frac{b}{c} \in \mathbb{Q}$ then $b|a_0$ and $c|a_n$.*

Proof. See Exercises. □

Remark 2.10. *If $a_0 = 0$ then $f(0) = 0$ and dividing by a high enough power of x will produce a polynomial $g(x)$ such that $a_0 \neq 0$. If $g(x)$ is constant then there are no further rational roots, otherwise the rational root test can then be used on $g(x)$ to find any non-zero rational roots.*

Example 2.11. *Let $f(x) = 3x^3 - 5x^2 + 5x - 2 \in \mathbb{Z}[x]$. The rational root test tells us that any rational root must belong to the set $\left\{\pm 1, \pm 2, \pm\frac{1}{3}, \pm\frac{2}{3}\right\}$. Testing all possibilities reveals that $f\left(\frac{2}{3}\right) = 0$, so that $\frac{2}{3}$ is the only rational root of $f(x)$. Indeed $f(x) = (x - \frac{2}{3})(3x^2 - 3x + 3)$.*

What happens if $f(x)$ doesn't have a rational root? It could still be reducible over $\mathbb{Q}$, since there might be a factorisation involving higher degree factors! In order to get to a stronger test, we first need a lemma connecting rational factorisations with integer factorisations.

Lemma 2.12. *(Gauss' Lemma) If $f(x) \in \mathbb{Z}[x]$ is reducible over $\mathbb{Q}$ then there exists a non-trivial factorisation with factors in $\mathbb{Z}[x]$.*

Proof. If $f(x) = g(x)h(x)$ is a non-trivial factorisation over $\mathbb{Q}$ then we can scale both sides by a large enough integer k to cancel denominators, giving a factorisation $kf(x) = s(x)t(x)$ with $s(x), t(z) \in \mathbb{Z}[x]$. If we can show that each prime $p|k$ either divides **all** of the coefficients of $s(x)$ or **all** of the coefficients of $t(x)$, then we can cancel p from both sides to still give a factorisation over $\mathbb{Z}$. Repeating this process would allow us to produce a factorisation of $f(x)$ over $\mathbb{Z}$.

Let $p|k$ be prime and suppose that:

$$s(x) = s_0 + s_1x + s_2x^2 + ... + s_{m_1}x^{m_1}$$
$$t(x) = t_0 + t_1x + t_2x^2 + ... + t_{m_2}x^{m_2}$$

If p doesn't divide all of the coefficients s_i then there must be a **smallest** index $a \geq 0$ such that $p \nmid s_a$. Similarly, if p doesn't divide all of the coefficients t_i then there must be a **smallest** index $b \geq 0$ such that $p \nmid t_b$. We now consider the coefficient of x^{a+b} on both sides of the factorisation:

$$kf_{a+b} = s_0t_{a+b} + s_1t_{a+b-1} + s_2t_{a+b-2} + ...s_at_b + ... + s_{a+b-1}t_1 + s_{a+b}t_0.$$

Note that every term in the sum is divisible by p, except s_at_b. This contradicts the fact that the LHS is divisible by p. It follows that one of $s(x), t(x)$ must have all of its coefficients divisible by p, proving the claim. □

Now we can state our better test!

Theorem 2.13. *(Eisenstein's Criterion) Let $f(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n \in \mathbb{Z}[x]$ and suppose that there is a prime p such that:*

1. $p \nmid a_n$
2. $p|a_i$ *for each* $0 \leq i \leq n-1$
3. $p^2 \nmid a_0$

Then $f(x)$ is irreducible over $\mathbb{Q}$.

Proof. Suppose that $f(x)$ is reducible over $\mathbb{Q}$, so that there exists a non-trivial factorisation $f(x) = g(x)h(x)$ with $g(x), h(x) \in \mathbb{Q}[x]$. By Gauss' Lemma we may assume that the two factors are in $\mathbb{Z}[x]$.

Suppose that the coefficients are as follows:

$$g(x) = b_0 + b_1x + b_2x^2 + ...b_{m_1}x^{m_1}$$
$$h(x) = c_0 + c_1x + c_2x^2 + ... + c_{m_2}x^{m_2}$$

Consider the constant term in the product, i.e. $a_0 = b_0c_0$. Since p is a prime satisfying $p|a_0$ we must have that $p|b_0$ or $p|c_0$ or both. However, $p^2 \nmid a_0$ and so we cannot have both. By swapping $g(x)$ and $h(x)$ we may assume that $p|b_0$ and $p \nmid c_0$.

We claim that $g(x)$ cannot have all of it's coefficients divisible by p. This is clear, since if it did then **all** of the coefficients in the product $g(x)h(x) = f(x)$ would be divisible by p. However, we know that $p \nmid a_n$, so this cannot happen.

Now choose the **smallest** index $j \geq 1$ such that $p \nmid b_j$ and consider the coefficient of x^j in the product:

$$a_j = b_0c_j + b_1c_{j-1} + ... + b_{j-1}c_1 + b_jc_0.$$

Note that all terms on the RHS are divisible by p except b_jc_0 and that $1 \leq j \leq m_1 < m_1 + m_2 = n$. This implies that $p \nmid a_j$, which contradicts the fact that $p|a_i$ for each $0 \leq i \leq n-1$. □

Example 2.14. *The polynomial $f(x) = 3x^4 + 5x^3 - 10x^2 + 55x - 40 \in \mathbb{Z}[x]$ is irreducible over $\mathbb{Q}$ by Eisenstein's Criterion with $p = 5$. We could not have proved this using the previous methods (it has no rational root but does have a real root!) and to do it using ad-hoc methods would have taken quite a bit of time!*

Remark 2.15. *Note that if a polynomial $f(x) \in \mathbb{Z}[x]$ is irreducible over $\mathbb{Q}$ then $f(x + a) \in \mathbb{Z}[x]$ is irreducible over $\mathbb{Q}$ for any $a \in \mathbb{Z}$. Such changes of variables are often useful when trying to use Eisenstein's Criterion (see Exercises).*

3 Number Fields and their Rings of Integers (3 lectures)

We saw earlier that there are useful rings out there that behave like $\mathbb{Z}$, e.g. $\mathbb{Z}[\sqrt{-2}], \mathbb{Z}[i], \mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[\zeta_n]$. In order to study analogues of $\mathbb{Z}$ it's easier to first understand the analogues of the field $\mathbb{Q}$.

3.1 Number Fields

Given a field, one can always build bigger fields by making extensions.

Definition 3.1. *An **extension** of a field K is a field L such that $K \subseteq L$. We denote this by L/K.*

A particular way of making extensions of K is by "adjoining" elements to K, i.e. inventing new elements to add/subtract/multiply/divide that aren't necessarily in K.

Definition 3.2. *The field $K(\alpha_1, \alpha_2, ..., \alpha_n)$ is the **smallest** extension of K that contains $\alpha_1, ..., \alpha_n$.*

Lemma 3.3. *For any $\alpha_1, \alpha_2, ..., \alpha_n$ we have that:*

$$K(\alpha_1, \alpha_2, ..., \alpha_n) = \left\{ \frac{p(\alpha_1, \alpha_2, ..., \alpha_n)}{q(\alpha_1, \alpha_2, ..., \alpha_n)} \;\middle|\; \begin{array}{l} p(x_1, x_2, ..., x_n) \\ q(x_1, x_2, ..., x_n) \end{array} \in K[x_1, x_2, ..., x_n] \quad \text{and} \quad q(\alpha_1, \alpha_2, ..., \alpha_n) \neq 0 \right\}.$$

Proof. Let $L = K(\alpha_1, \alpha_2, ..., \alpha_n)$ for ease of notation. Since L contains K and $\alpha_1, \alpha_2, ..., \alpha_n$ and is closed under addition and multiplication, we see that L contains $p(\alpha_1, \alpha_2, ..., \alpha_n)$ for any $p(x_1, x_2, ..., x_n) \in K[x_1, x_2, ..., x_n]$.

Since every non-zero element of L is a unit, we must have that $\frac{1}{q(\alpha_1, \alpha_2, ..., \alpha_n)} \in L$ for any $q(x_1, x_2, ..., x_n) \in K[x_1, x_2, ..., x_n]$ satisfying $q(\alpha_1, \alpha_2, ..., \alpha_n) \neq 0$.

Closure under multiplication then shows that L contains the RHS. Since the RHS is a field (see Exercises), we must have equality (otherwise we would contradict the minimality of L). □

Example 3.4. *Let's consider the field $\mathbb{R}(i)$. Since $i^2 = -1$, any element of the form $p(i)$ with $p(x) \in \mathbb{R}[x]$ can be written as $a + bi$ for some $a, b \in \mathbb{R}$. Then any expression of the form*

$$\frac{a + bi}{c + di}$$

with $a, b, c, d \in \mathbb{R}$ and $(c, d) \neq (0, 0)$ can be written as $x + iy$ for some $x, y \in \mathbb{R}$ by multiplying numerator and denominator by the complex conjugate. Since $x + iy \in \mathbb{R}(i)$ for all $x, y \in \mathbb{R}$ we must have that $\mathbb{R}(i) = \{x + iy \,|\, x, y \in \mathbb{R}\} = \mathbb{C}$.

The fields that we will be most interested in throughout this course will be of the form $\mathbb{Q}(\alpha)$ with α an algebraic number.

Definition 3.5. *An **algebraic number** is a complex root of a non-zero polynomial $f(x) \in \mathbb{Q}[x]$.*

Numbers like $\sqrt{2}, \sqrt[3]{5}$ and $\zeta_7 = e^{\frac{2\pi i}{7}}$ are algebraic numbers (since they are roots of the rational polynomials $x^2 - 2, x^3 - 5$ and $x^7 - 1$ respectively). However, numbers like π and e are not algebraic (although this is not obvious!). Note that they are algebraic numbers "over $\mathbb{R}$" though (since they are the roots of the **real** polynomials $x - \pi$ and $x - e$).

Definition 3.6. *A **number field** is a field of the form $\mathbb{Q}(\alpha_1, \alpha_2, ..., \alpha_n)$ where each α_i is an algebraic number.*

Naturally we wonder what number fields look like in practice. Since algebraic numbers satisfy nice polynomial relations, it turns out that a lot can be said.

Example 3.7. *1. Let's consider the number field $\mathbb{Q}(\sqrt{2})$. Since $\alpha = \sqrt{2}$ is "quadratic", i.e. satisfies $\alpha^2 = 2$, any element of the form $p(\sqrt{2})$ for $p(x) \in \mathbb{Q}[x]$ can be written as $a + b\sqrt{2}$ for some $a, b \in \mathbb{Q}$. Then any expression of the form*

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}}$$

with $a, b, c, d \in \mathbb{Q}$ and $(c, d) \neq (0, 0)$ can be written as $x + y\sqrt{2}$ for some $x, y \in \mathbb{Q}$ by rationalising the denominator. Since $x + y\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ for all $x, y \in \mathbb{Q}$ we must have that $\mathbb{Q}(\sqrt{2}) = \{x + y\sqrt{2} \,|\, x, y \in \mathbb{Q}\}$.

2. Let's consider the number field $\mathbb{Q}(\sqrt[3]{5})$. Since $\alpha = \sqrt[3]{5}$ is "cubic", i.e. satisfies $\alpha^3 = 5$, any element of the form $p(\sqrt[3]{5})$ can be written as $a + b\sqrt[3]{5} + c\sqrt[3]{5}^2$ for some $a, b, c \in \mathbb{Q}$. We now have to deal with any expression of the form:

$$\frac{a + b\sqrt[3]{5} + c\sqrt[3]{5}^2}{d + e\sqrt[3]{5} + f\sqrt[3]{5}^2}$$

with $a, b, c, d, e, f \in \mathbb{Q}$ and $(d, e, f) \neq (0, 0, 0)$. But how do we rationalise the denominator now?!

Multiplying top and bottom by $(d^2 - 5ef) + (5f^2 - de)\sqrt[3]{5} + (e^2 - df)\sqrt[3]{5}^2$ gives (see Exercises):

$$\begin{aligned}
&\frac{ad^2 + 5be^2 + 25cf^2 - 5aef - 5bdf - 5cde}{d^2 + 5e^2 + 25f^2 - 15def} \\
&+ \frac{5af^2 + bd^2 + 5ce^2 - ade - 5bef - 5cdf}{d^2 + 5e^2 + 25f^2 - 15def}\sqrt[3]{5} \\
&+ \frac{ae^2 + 5bf^2 + cd^2 - adf - bde - 5cef}{d^2 + 5e^2 + 25f^2 - 15def}\sqrt[3]{5}^2.
\end{aligned}$$

This is certainly an element of the form $x + y\sqrt[3]{5} + z\sqrt{5}^2$ with $x, y, z \in \mathbb{Q}$. Since all such elements are in $\mathbb{Q}(\sqrt[3]{5})$, we must have that $\mathbb{Q}(\sqrt[3]{5}) = \{x + y\sqrt[3]{5} + z\sqrt[3]{5}^2 \,|\, x, y, z \in \mathbb{Q}\}$.

In this course, we'll mainly focus on number fields of the form $\mathbb{Q}(\alpha)$, i.e. generated by one algebraic number. This is not a restriction, since it can be shown that **every** number field can be written in the above form (see Exercises). In this case, the evidence above suggests the following conjecture.

Conjecture 3.8. *If α is an algebraic number of "degree d" then:*

$$\mathbb{Q}(\alpha) = \{a_0 + a_1\alpha + a_2\alpha^2 + ... + a_{d-1}\alpha^{d-1} \,|\, a_0, a_1, ..., a_{d-1} \in \mathbb{Q}\}$$

But, what do we mean by "degree d"? We need to clarify what this means! It cannot simply be "the degree of the polynomial satisfied by α", since this is not unique. For example $\alpha = \sqrt{2}$ satisfies many polynomials of varying degrees, for example

$$x^2-2 \qquad 7x^2-14 = 7(x^2-2) \qquad x^4-4 = (x^2-2)(x^2+2) \qquad x^7-2x^5+x^2-2 = (x^2-2)(x^5+1)...$$

Definition 3.9. *The **minimal polynomial** of an algebraic number α is the **monic** non-zero polynomial $m_\alpha(x) \in \mathbb{Q}[x]$ of **smallest degree** that has α as a root.*

We should prove that this polynomial exists and is unique before doing anything else!

Lemma 3.10. *The minimal polynomial of an algebraic number α exists and is unique. It is also irreducible over $\mathbb{Q}$.*

Proof. Let S_α be the set of monic non-zero polynomials that have α as a root. We know that S_α is non-empty since α is algebraic (each non-zero polynomial having α as a root can be scaled to give a monic one). In particular, there must exist an element $m_\alpha(x) \in S_\alpha$ of smallest degree.

Suppose that there are two polynomials $m_\alpha(x), m'_\alpha(x) \in S_\alpha$ of smallest degree d. Then $m''_\alpha(x) = m_\alpha(x) - m'_\alpha(x)$ is a polynomial that also has α as a root and has degree $d' < d$. Scaling this gives a monic polynomial in S_α of degree $d' < d$. By minimality of d this can only happen if $m''_\alpha(x) = 0$, i.e. $m_\alpha(x) = m'_\alpha(x)$, proving uniqueness.

To prove irreducibility of $m_\alpha(x)$, suppose that there is a non-trivial factorisation $m_\alpha(x) = f(x)g(x)$ with $f(x), g(x) \in \mathbb{Q}[x]$ having degrees greater than or equal to 1. Then $0 = m_\alpha(\alpha) = f(\alpha)g(\alpha)$ is an equality in $\mathbb{C}$. This implies that $f(\alpha) = 0$ or $g(\alpha) = 0$, so that we would have found a smaller degree polynomial having α as a root, contradicting the minimality of $m_\alpha(x)$. □

Now we can prove our conjecture!

Theorem 3.11. *Let α be an algebraic number and let $d = deg(m_\alpha(x))$. Then:*

$$\mathbb{Q}(\alpha) = \{a_0 + a_1\alpha + a_2\alpha^2 + ... + a_{d-1}\alpha^{d-1} \mid a_0, a_1, ..., a_{d-1} \in \mathbb{Q}\}$$

Proof. Since $m_\alpha(\alpha) = 0$ it's clear that every expression $p(\alpha)$ for $p(x) \in \mathbb{Q}[x]$ can be written as $a_0+a_1\alpha+a_2\alpha^2+...+a_{d-1}\alpha^{d-1}$ for some $a_0, a_1, a_2, ..., a_{d-1} \in \mathbb{Q}$. We now have to deal with expressions of the form:

$$\frac{a_0 + a_1\alpha + a_2\alpha^2 + ... + a_{d-1}\alpha^{d-1}}{b_0 + b_1\alpha + b_2\alpha^2 + ... + b_{d-1}\alpha^{d-1}}$$

with the $a_i, b_i \in \mathbb{Q}$ and the b_i not all zero.

[Our problem is that we again don't know how to rationalise the denominator! It's actually possible to do this (as we'll see later), but it turns out that there's a much slicker way to proceed. We simply need to show that the denominator is invertible... since then each such expression will belong to the RHS.]

Consider the polynomial $b(x) = b_0+b_1x+b_2x^2+...+b_{d-1}x^{d-1} \in \mathbb{Q}[x]$. Using Euclid's Algorithm we can find polynomials $s(x), t(x) \in \mathbb{Q}[x]$ such that $s(x)m_\alpha(x)+t(x)b(x) = 1$ (since $m_\alpha(x)$ irreducible

and $\deg(b(x)) < \deg(m_\alpha(x))$ imply that $\gcd(m_\alpha(x), b(x)) = 1$). Plugging in α gives that $t(\alpha)b(\alpha) = 1$, so that $t(\alpha) = b(\alpha)^{-1}$ exists.

We have now shown that $\mathbb{Q}(\alpha)$ contains the RHS of the theorem. Since the RHS is clearly a subset of $\mathbb{Q}(\alpha)$, we must have equality. □

Remark 3.12. *The above should remind you of the method of getting inverses in $(\mathbb{Z}/N\mathbb{Z})^\times$, i.e. use Euclid's Algorithm to find $s, t \in \mathbb{Z}$ such that $sN + tb = 1$ and then reduce mod N to get $tb \equiv 1 \bmod N$.*

Indeed, "plugging in α" and using the fact that $m_\alpha(\alpha) = 0$ is the analogous step to "reducing mod N" (i.e. declaring that "$N = 0$").

Definition 3.13. *We call $d = deg(m_\alpha(x))$ the **degree** of the extension $K/\mathbb{Q}$, often denoted $d = [K : \mathbb{Q}]$.*

Note that if L/K is a field extension then L is a vector space over K (since $(L, +)$ is an abelian group under addition and we are able to scale by elements of K compatibly, since $K \subseteq L$).

The above theorem tells us that if $\mathbb{Q}(\alpha)$ is a number field then $\{1, \alpha, \alpha^2, ..., \alpha^{d-1}\}$ is a basis for the underlying vector space over $\mathbb{Q}$ (see Exercises). Hence the degree $d = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ is the **dimension** of this vector space.

Example 3.14.

1. *The number field $\mathbb{Q}(i) = \{x + iy \mid x, y \in \mathbb{Q}\}$ has degree $[\mathbb{Q}(i) : \mathbb{Q}] = 2$, since $m_i(x) = x^2 + 1$ (this is irreducible over $\mathbb{R}$, hence is irreducible over $\mathbb{Q}$). The elements $\{1, i\}$ form a basis for this field as a vector space over $\mathbb{Q}$.*
2. *The number field $\mathbb{Q}(\sqrt[3]{5}) = \{x + y\sqrt[3]{5} + z\sqrt[3]{5}^2 \mid x, y, z \in \mathbb{Q}\}$ has degree $[\mathbb{Q}(\sqrt[3]{5}) : \mathbb{Q}] = 3$, since $m_{\sqrt[3]{5}}(x) = x^3 - 5$ (irreducible over $\mathbb{Q}$ by Eisenstein's Criterion with $p = 5$). The elements $\{1, \sqrt[3]{5}, \sqrt[3]{5}^2\}$ form a basis for this field as a vector space over $\mathbb{Q}$.*

3.2 Conjugates, Norms, Traces and Discriminants.

Earlier we defined an algebraic number to be a root of a polynomial in $\mathbb{Q}[x]$. We then showed that such numbers have uniquely defined minimal polynomials. However, there a little bit of ambiguity here since a minimal polynomial can have multiple roots...and all are equally valid from the point of view of algebra!

For example, the polynomial $x^2 - 2$ has **two** roots, i.e. $\pm\sqrt{2}$, and there is no obvious way to distinguish these from an algebraic point of view. For this reason, it will be helpful to keep a record **all** roots of a given polynomial.

Definition 3.15. *Let α be an algebraic number. The **conjugates** of α are the complex roots of the minimal polynomial $m_\alpha(x)$.*

Example 3.16. *The conjugates of $\alpha = i$ are $\pm i$ and the conjugates of $\alpha = \sqrt{11}$ are $\pm\sqrt{11}$.*

Remark 3.17. *Note that the conjugates of an algebraic number α can lie outside of the number field $\mathbb{Q}(\alpha)$. For example, the conjugates of $\alpha = \sqrt[3]{5}$ are $\sqrt[3]{5}, \zeta_3\sqrt[3]{5}$ and $\zeta_3^2\sqrt[3]{5}$, where $\zeta_3 = e^{\frac{2\pi i}{3}}$ is a primitive cube root of unity. The last two of these do not lie in $\mathbb{Q}(\sqrt[3]{5})$ (as they aren't even in $\mathbb{R}$).*

From an algebraic point of view, all conjugates of an algebraic number are considered to be "the same" (i.e. they are all roots of the same minimal polynomial). However, as complex numbers they are distinct. This means that a number field $\mathbb{Q}(\alpha)$ can actually be viewed in many different ways as a subfield of $\mathbb{C}$ (depending on which conjugate of α you choose to represent a root of the minimal polynomial).

Lemma 3.18. *Suppose that α is algebraic of degree $d = deg(m_\alpha(x))$, with conjugates $\beta_1, \beta_2, ..., \beta_d$. Then there are exactly d injective field homomorphisms:*

$$\sigma_i : \mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$$
$$a_0 + a_1\alpha + a_2\alpha^2 + ... + a_{d-1}\alpha^{d-1} \longmapsto a_0 + a_1\beta_i + a_2\beta_i^2 + ... + a_{d-1}\beta_i^{d-1}$$

Proof. First note that a field homomorphism f must preserve the ring structure and so must be a ring homomorphism. This means that it must satisfy $f(0) = 0$ and $f(1) = 1$, and from here it is easy to show that $f(a) = a$ for any $a \in \mathbb{Q}$ (See Exercises).

It follows that:

$$\begin{aligned} f(a_0 + a_1\alpha + a_2\alpha^2 + ... + a_{d-1}\alpha^{d-1}) &= f(a_0) + f(a_1)f(\alpha) + f(a_2)f(\alpha)^2 + ... + f(a_{d-1})f(\alpha)^{d-1} \\ &= a_0 + a_1f(\alpha) + a_2f(\alpha)^2 + ... + a_{d-1}f(\alpha)^{d-1}. \end{aligned}$$

so that f is completely determined by what $f(\alpha)$ is. However, we also know that $m_\alpha(\alpha) = 0$ and applying f to both sides of this gives $f(m_\alpha(\alpha)) = m_\alpha(f(\alpha)) = f(0) = 0$, and so $f(\alpha)$ must be a root of $m_\alpha(x)$, i.e. a conjugate of α. This shows that $f(\alpha) = \beta_i$ for some i, so that $f = \sigma_i$.

It remains to prove that all of the maps σ_i are injective field homomorphisms (they are clearly different maps since they send α to each the conjugates $\beta_1, \beta_2, ..., \beta_d$, which are distinct). Let $\gamma_1, \gamma_2 \in \mathbb{Q}(\alpha)$ be written as:

$$\begin{aligned} \gamma_1 &= a_0 + a_1\alpha + a_2\alpha^2 + ... + a_{d-1}\alpha^{d-1} \\ \gamma_2 &= b_0 + b_1\alpha + b_2\alpha^2 + ... + b_{d-1}\alpha^{d-1} \end{aligned}$$

Then:

$$\begin{aligned} \sigma_i(\gamma_1 + \gamma_2) &= \sigma_i((a_0 + b_0) + (a_1 + b_1)\alpha + (a_2 + b_2)\alpha^2 + ... + (a_{d-1} + b_{d-1})\alpha^{d-1}) \\ &= (a_0 + b_0) + (a_1 + b_1)\beta_i + (a_2 + b_2)\beta_i^2 + ... + (a_{d-1} + b_{d-1})\beta_i^{d-1} \\ &= (a_0 + a_1\beta_i + a_2\beta_i^2 + ... + a_{d-1}\beta_i^{d-1}) + (b_0 + b_1\beta_i + b_2\beta_i^2 + ... + b_{d-1}\beta_i^{d-1}) \\ &= \sigma_i(\gamma_1) + \sigma_i(\gamma_2) \end{aligned}$$

To prove injectivity, suppose $\sigma_i(\gamma_1) = \sigma_i(\gamma_2)$. Since σ_i is a homomorphism we have $\sigma_i(\gamma_1 - \gamma_2) = 0$, so that:

$$(a_0 - b_0) + (a_1 - b_1)\sigma_i(\alpha) + (a_2 - b_2)\sigma_i(\alpha)^2 + ... + (a_{d-1} - b_{d-1})\sigma_i(\alpha)^{d-1} = 0.$$

We must then have $a_i = b_i$ for all i, otherwise we would have found a polynomial of degree $d - 1$ satisfied by $\sigma_i(\alpha)$ (which by definition has minimal polynomial $m_\alpha(x)$ of degree d). Thus $\gamma_1 = \gamma_2$, so that σ_i is injective. □

Example 3.19. *There are two embeddings of the number field $\mathbb{Q}(\sqrt{7})$ into $\mathbb{C}$:*

$$\sigma_1(a+b\sqrt{7}) = a+b\sqrt{7}$$
$$\sigma_2(a+b\sqrt{7}) = a-b\sqrt{7}$$

The images of these embeddings are both $\mathbb{Q}(\sqrt{7})$, i.e. the same set of numbers, but we are stressing the choice of which "square root of 7" to use on the RHS.

There are three embeddings of the number field $\mathbb{Q}(\sqrt[3]{5})$ into $\mathbb{C}$:

$$\sigma_1(a+b\sqrt[3]{5}+c\sqrt[3]{5}^2) = a+b\sqrt[3]{5}+c\sqrt[3]{5}^2$$
$$\sigma_1(a+b\sqrt[3]{5}+c\sqrt[3]{5}^2) = a+b\zeta_3\sqrt[3]{5}+c\zeta_3^2\sqrt[3]{5}^2$$
$$\sigma_1(a+b\sqrt[3]{5}+c\sqrt[3]{5}^2) = a+b\zeta_3^2\sqrt[3]{5}+c\zeta_3\sqrt[3]{5}^2$$

Note that the images of these embeddings are the number fields $\mathbb{Q}(\sqrt[3]{5})$ and $\mathbb{Q}(\zeta_3\sqrt[3]{5}) = \mathbb{Q}(\zeta_3^2\sqrt[3]{5})$. This time we genuinely get different number fields, depending on the choice of which "cube root of 5" to use on the RHS.

We can now define various interesting maps relating elements of number fields to rational numbers.

Definition 3.20. *Let $\mathbb{Q}(\alpha)$ be a number field of degree d. The **norm** of $\beta \in \mathbb{Q}(\alpha)$ is $N(\beta) = \prod_{i=1}^{d}\sigma_i(\beta)$ and the **trace** is $T(\beta) = \sum_{i=1}^{d}\sigma_i(\beta)$.*

Example 3.21. *If $a+b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ then its norm and trace are given by:*

$$N(a+b\sqrt{2}) = (a+b\sqrt{2})(a-b\sqrt{2}) \qquad = a^2-2b^2$$
$$T(a+b\sqrt{2}) = (a+b\sqrt{2})+(a-b\sqrt{2}) \qquad = 2a$$

We note a few nice properties of these maps.

Lemma 3.22. *Let α be algebraic of degree $d = deg(m_\alpha(x))$. Then for every $\beta, \gamma \in \mathbb{Q}(\alpha)$ and $\lambda \in \mathbb{Q}$ we have:*

1. *$N(\beta\gamma) = N(\beta)N(\gamma)$, i.e. the norm is multiplicative.*
2. *$T(\beta+\gamma) = T(\beta)+T(\gamma)$ and $T(\lambda\beta) = \lambda T(\beta)$, i.e. the trace is linear.*
3. *$N(\lambda) = \lambda^d$ and $T(\lambda) = d\lambda$.*
4. *$N(\beta), T(\beta) \in \mathbb{Q}$.*

Proof. 1. This is a simple consequence of the fact that the σ_i are homomorphisms:

$$N(\beta\gamma) = \prod_{i=1}^{d}\sigma_i(\beta\gamma) = \prod_{i=1}^{d}\sigma_i(\beta)\sigma_i(\gamma) = \left(\prod_{i=1}^{d}\sigma_i(\beta)\right)\left(\prod_{i=1}^{d}\sigma_i(\gamma)\right) = N(\beta)N(\gamma).$$

2. Similarly:

$$T(\beta+\gamma)=\sum_{i=1}^{d}\sigma_i(\beta+\gamma)=\sum_{i=1}^{d}(\sigma_i(\beta)+\sigma_i(\gamma))=\left(\sum_{i=1}^{d}\sigma_i(\beta)\right)+\left(\sum_{i=1}^{d}\sigma_i(\gamma)\right)=T(\beta)+T(\gamma)$$

and

$$T(\lambda\beta)=\sum_{i=1}^{d}\sigma_i(\lambda\beta)=\sum_{i=1}^{d}\sigma_i(\lambda)\sigma_i(\beta)=\sum_{i=1}^{d}\lambda\sigma_i(\beta)=\lambda\sum_{i=1}^{d}\sigma_i(\beta)=\lambda T(\beta).$$

3. These follow since $N(\lambda)=\prod_{i=1}^{d}\sigma_i(\lambda)=\prod_{i=1}^{d}\lambda=\lambda^d$ and $T(\lambda)=\sum_{i=1}^{d}\sigma_i(\lambda)=\sum_{i=1}^{d}\lambda=d\lambda$.

4. We will only provide a sketch proof, based on the following fact:

   FACT: If $\beta\in\mathbb{Q}(\alpha)$ then β is algebraic of degree $d'\mid d$ and the values $\sigma_i(\beta)$ are the conjugates of β, each repeated $\frac{d}{d'}$ times.

   (Note: we already knew that $\sigma_i(\beta)$ had to be a conjugate of β since $\sigma_i(m_\beta(\beta))=m_\beta(\sigma_i(\beta))=\sigma_i(0)=0$. However, the fact that each occurs with the same multiplicity is not obvious).

   The claim is now clear, since if the conjugates of β are $\beta_1,\beta_2,...,\beta_{d'}$ then:

$$N(\beta)=\prod_{i=1}^{d}\sigma_i(\beta)=\left(\prod_{i=1}^{d'}\beta_i\right)^{\frac{d}{d'}}$$

   and the quantity inside the brackets is the constant coefficient of $m_\beta(x)$ (which is in $\mathbb{Q}$). Similarly:

$$T(\beta)=\sum_{i=1}^{d}\sigma_i(\beta)=\frac{d}{d'}\left(\sum_{i=1}^{d'}\beta_i\right)$$

   and the quantity inside the brackets is the $x^{d'-1}$ coefficient of $m_\beta(x)$ (which is in $\mathbb{Q}$).

□

We finish by defining one final interesting quantity associated to a tuple of elements in $\mathbb{Q}(\alpha)$.

Definition 3.23. *Let $\mathbb{Q}(\alpha)$ be a number field of degree d and $\beta_1,\beta_2,...,\beta_d\in\mathbb{Q}(\alpha)$. The **discriminant** is $\Delta(\beta_1,\beta_2,...,\beta_d)=det(M_{\beta_1,\beta_2,...,\beta_d})^2$, where:*

$$M_{\beta_1,\beta_2,...,\beta_d}=\begin{pmatrix}\sigma_1(\beta_1) & \sigma_1(\beta_2) & \dots & \sigma_1(\beta_d)\\ \sigma_2(\beta_1) & \sigma_2(\beta_2) & \dots & \sigma_2(\beta_d)\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_d(\beta_1) & \sigma_d(\beta_2) & \dots & \sigma_d(\beta_d)\end{pmatrix}$$

The facts we will need are as follows:

Lemma 3.24. *Let $\mathbb{Q}(\alpha)$ be a number field of degree d and $\beta_1,\beta_2,...,\beta_d\in\mathbb{Q}(\alpha)$ be a basis.*

1. *$\Delta(\beta_1,\beta_2,...,\beta_d)\in\mathbb{Q}\backslash\{0\}$.*
2. *If $\beta'_j=\sum_{i=1}^{d}c_{i,j}\beta_i$ for some $c_{i,j}\in\mathbb{Q}$ then*

$$\Delta(\beta'_1,\beta'_2,...,\beta'_d)=(det(C))^2\Delta(\beta_1,\beta_2,...,\beta_d)$$

   where C is the matrix having (i,j)-entry $c_{i,j}$.

3.3 The Ring of Integers $\mathcal{O}_K$

Now that we have defined analogues of the field $\mathbb{Q}$ of rational numbers, we can get started on defining the analogues of the subring $\mathbb{Z}$ of integers.

Definition 3.25. *An algebraic number β such that $m_\beta(x) \in \mathbb{Z}[x]$ is called an **algebraic integer**.*

Example 3.26. *The algebraic number $\beta = \sqrt{3}$ is an algebraic integer, since $m_\beta(x) = x^2 - 3 \in \mathbb{Z}[x]$.*

Definition 3.27. *Let K be a number field. The **ring of integers of** K is:*

$$\mathcal{O}_K = \{\beta \in K \,|\, m_\beta(x) \in \mathbb{Z}[x]\}.$$

This looks like a strange definition, but it turns out to be the correct one for many reasons which we will see later.

Example 3.28. *The ring of integers of $\mathbb{Q}$ is $\mathbb{Z}$. Indeed, the minimal polynomial of $\beta = \frac{a}{b}\mathbb{Q}$ is clearly $x - \frac{a}{b}$ and this is in $\mathbb{Z}[x]$ if and only if $b = \pm 1$, so that $\beta \in \mathbb{Z}$.*

Example 3.29. *The ring of integers of $\mathbb{Q}(i)$ is $\mathbb{Z}[i]$. This is because if $\beta = a + bi \in \mathbb{Q}(i)$ then $m_\beta(x) = x^2 - 2ax + (a^2 + b^2)$. For this to be in $\mathbb{Z}[x]$ we would require:*

$$2a \in \mathbb{Z} \qquad \text{and} \qquad (a^2 + b^2) \in \mathbb{Z}.$$

The first inclusion tells us that $a \in \frac{1}{2}\mathbb{Z}$ and the second then tells us that $b^2 \in \frac{1}{4}\mathbb{Z}$, so that $b \in \frac{1}{2}\mathbb{Z}$.

Let $a = \frac{c}{2}$ and $b = \frac{d}{2}$ with $c, d \in \mathbb{Z}$. Then the second inclusion gives $c^2 + d^2 \in 4\mathbb{Z}$, so that $c^2 + d^2 \equiv 0 \bmod 4$. The only solution to this congruence is $c \equiv d \equiv 0 \bmod 2$ (since if exactly one of c, d is odd then $c^2 + d^2$ would be odd and if both c, d are odd then $c^2 + d^2 \equiv 2 \bmod 4$). This means that $a, b \in \mathbb{Z}$ so that $\mathcal{O}_K \subseteq \mathbb{Z}[i]$.

It's clear that every element of $\mathbb{Z}[i]$ satisfies the two inclusions, hence $\mathcal{O}_K = \mathbb{Z}[i]$.

Example 3.30. *It might be tempting at this stage to think that if $K = \mathbb{Q}(\alpha)$ and $\alpha \in \mathcal{O}_K$ then $\mathcal{O}_K = \mathbb{Z}[\alpha]$. However, this is **not** true!*

For example, consider the element $\beta = \frac{1+\sqrt{-3}}{2} \in \mathbb{Q}(\sqrt{-3})$. It's minimal polynomial is $m_\beta(x) = x^2 - x + 1 \in \mathbb{Z}[x]$ and so $\beta \in \mathcal{O}_K$, yet clearly $\beta \notin \mathbb{Z}[\sqrt{-3}]$!

In this case it's possible to show that $\mathcal{O}_K = \left[\frac{1+\sqrt{-3}}{2}\right] = \left\{a + b\left(\frac{1+\sqrt{-3}}{2}\right) \,\middle|\, a, b \in \mathbb{Z}\right\}$ (see Exercises).

For general quadratic number fields we have the following description of $\mathcal{O}_K$:

Proposition 3.31. *Let $d \neq 0, \pm 1$ be square-free. Then the ring of integers of $\mathbb{Q}(\sqrt{d})$ is:*

$$\mathcal{O}_K = \begin{cases} \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right] & \text{if } d \equiv 1 \bmod 4 \\ \mathbb{Z}[\sqrt{d}] & \text{if } d \not\equiv 1 \bmod 4 \end{cases}$$

Proof. See Exercises. □

For higher degree number fields it's not so clear how to work out what $\mathcal{O}_K$ is. However, there exist efficient algorithms to compute it (see any good book on Algebraic Number Theory).

So far, we've been calling $\mathcal{O}_K$ a ring...but we haven't actually proved that it's a ring! We should do that now...but first a small lemma. We defined element of $\mathcal{O}_K$ to be the ones with integer minimal polynomial. It turns out that we didn't need to be so strict!

Lemma 3.32. *Let K be a number field and $\beta \in K$. If β is the root of a monic polynomial $f(x) \in \mathbb{Z}[x]$ then $\beta \in \mathcal{O}_K$.*

Proof. We must show that $m_\beta(x) \in \mathbb{Z}[x]$. First note that $m_\beta(x)|f(x)$, since by Euclid's Algorithm we can write $f(x) = q(x)m_\beta(x) + r(x)$ with $r(x) = 0$ or $\deg(r(x)) < \deg(m_\beta(x))$. Substituting $x = \beta$ gives $0 = f(\beta) = q(\beta)m_\beta(\beta)+r(\beta) = r(\beta)$, so that $r(\beta) = 0$. If $r(x) \neq 0$ we would contradict the minimality of $m_\beta(x)$, so we must have $r(x) = 0$ and $f(x) = q(x)m_\beta(x)$.

Now note that $f(x) \in \mathbb{Z}[x]$ has been factored into a product of two polynomials $q(x), m_\beta(x) \in \mathbb{Q}[x]$. Note that $m_\beta(x)$ is monic and so is $q(x)$. It follows from Exercise 4 of Chapter 2 that both $q(x)$ and $m_\beta(x)$ lie in $\mathbb{Z}[x]$. □

We can now prove that $\mathcal{O}_K$ is a ring.

Theorem 3.33. *Let K be a number field. Then $\mathcal{O}_K$ is a commutative ring.*

Proof. Let $\beta_1, \beta_2 \in \mathcal{O}_K$ have minimal polynomials

$$m_{\beta_1}(x) = x^{d_1} + a_{d_1-1}x^{d_1-1} + ... + a_1x + a_0 \in \mathbb{Z}[x]$$
$$m_{\beta_2}(x) = x^{d_1} + b_{d_2-1}x^{d_2-1} + ... + b_1x + b_0 \in \mathbb{Z}[x]$$

Let $v \in K^{d_1d_2}$ be any vector having the components $\beta_1^i\beta_2^j$ for $0 \leq i < d_1$ and $0 \leq j < d_2$ (i.e. fix an ordering).

Consider the vector $\beta_1 v$. It's entries are of the form $\beta_1^{i+1}\beta_2^j$. If $1 \leq i + 1 < d_1$ then this is simply another entry of the vector v. If $i + 1 = d_1$ then we may use the minimal polynomial of β_1 to write as an **integer** linear combination of the entries of v as follows:

$$\beta_1^{d_1}\beta_2^j = -(a_{d_1-1}\beta_1^{d_1-1} + ... + a_1\beta_1 + a_0)\beta_2^j = -a_{d_1-1}\beta_1^{d_1-1}\beta_2^j - ... - a_1\beta_1\beta_2^j - a_0\beta_2^j.$$

It follows that $\beta_1 v = M_{\beta_1} v$ for some matrix $M_{\beta_1} \in M_{d_1d_2}(\mathbb{Z})$, so that β_1 is an eigenvalue of M_{β_1} with eigenvector v.

An identical argument shows that $\beta_2 v = M_{\beta_2} v$ for some matrix $M_{d_1d_2}(\mathbb{Z})$, so that β_2 is an eigenvalue of M_{β_2} with eigenvector v.

Now consider the matrices $M_{\beta_1} + M_{\beta_2}$ and $M_{\beta_1}M_{\beta_2}$. These matrices have $\beta_1 + \beta_2$ and $\beta_1\beta_2$ as eigenvalues respectively, since

$$(M_{\beta_1} + M_{\beta_2})v = M_{\beta_1}v + M_{\beta_2}v = \beta_1 v + \beta_2 v = (\beta_1 + \beta_2)v$$
$$(M_{\beta_1}M_{\beta_2})v = M_{\beta_1}(\beta_2 v) = \beta_2(M_{\beta_1}v) = (\beta_1\beta_2)v$$

This means that $\beta_1 + \beta_2$ and $\beta_1\beta_2$ are roots of the characteristic polynomials of $M_{\beta_1} + M_{\beta_2}$ and $M_{\beta_1}M_{\beta_2}$ respectively. Since these are integer matrices, their characteristic polynomials must be monic integer polynomials.

Thus we have found two monic integer polynomials having $\beta_1 + \beta_2$ and $\beta_1\beta_2$ as roots. By Lemma 3.32, we find that $\beta_1 + \beta_2 \in \mathcal{O}_K$ and $\beta_1\beta_2 \in \mathcal{O}_K$. This proves closure of $\mathcal{O}_K$ under sums and products. The other ring axioms follow from the fact that K is a field and $\mathcal{O}_K \subset K$.

□

Remark 3.34. *The above proof shows that the sum/product of any two algebraic integers is an algebraic integer. Indeed, if α, β are any two algebraic integers then we can apply the theorem to the number field $\mathbb{Q}(\alpha, \beta)$ (since $\alpha, \beta \in \mathcal{O}_K$ for this field and the theorem proves that $\mathcal{O}_K$ is a ring.*

Earlier we saw that elements of number fields have rational norm and trace. When we take the norm and trace of an element of $\mathcal{O}_K$ something even nicer happens!

Lemma 3.35. *Let K be a number field and $\beta \in \mathcal{O}_K$. Then $N(\beta), T(\beta) \in \mathbb{Z}$.*

Proof. Going back to the proof of Lemma 3.22, we saw that if $[K : \mathbb{Q}] = d$ and $\beta_1, \beta_2, ..., \beta_{d'}$ are the conjugates of β then:

$$N(\beta) = \left(\prod_{i=1}^{d'} \beta_i\right)^{\frac{d}{d'}}$$
$$T(\beta) = \frac{d}{d'}\left(\sum_{i=1}^{d'} \beta_i\right)$$

The quantities inside the brackets are the constant coefficient and the $x^{d'-1}$ coefficient of $m_\beta(x)$ (which are in $\mathbb{Z}$ since $m_\beta(x) \in \mathbb{Z}[x]$). □

Recall that every number field $K = \mathbb{Q}(\alpha)$ is a $\mathbb{Q}$-vector space and so has a basis over $\mathbb{Q}$ (of size $d = \deg(m_\alpha(x))$). We finish this section by proving that a similar claim holds for the ring $\mathcal{O}_K \subset K$.

Theorem 3.36. *Let K be a number field of degree d. Then there exist $\beta_1, \beta_2, ..., \beta_d \in \mathcal{O}_K$ such that:*

$$\mathcal{O}_K = \{a_1\beta_1 + a_2\beta_2 + ... + a_d\beta_d \,|\, a_1, a_2, ..., a_d \in \mathbb{Z}\}.$$

Proof. We first write $K = \mathbb{Q}(\alpha)$ for some $\alpha \in K$. Note that we can assume that $\alpha \in \mathcal{O}_K$ (see Exercises).

Pick elements $\beta_1, \beta_2, ..., \beta_d \in \mathcal{O}_K$ that are a basis for K (such elements exist, for example $1, \alpha, \alpha^2, ..., \alpha^{d-1}$). If these elements are a $\mathbb{Z}$-basis for $\mathcal{O}_K$ then we are done. Otherwise, there exists $\alpha \in \mathcal{O}_K$ such that:

$$\alpha = a_1\beta_1 + a_2\beta_2 + ... + a_d\beta_d,$$

with at least one $a_i \in \mathbb{Q}\backslash\mathbb{Z}$. We can reorder the β_i so that $a_1 \in \mathbb{Q}\backslash\mathbb{Z}$.

Now consider $\beta_1' = \alpha - \lfloor a_1 \rfloor \beta_1$. Then since $\lfloor a_1 \rfloor \in \mathbb{Z}$ and $\mathcal{O}_K$ is a ring, we have that $\beta_1' \in \mathcal{O}_K$. The set $\beta_1', \beta_2, ..., \beta_d$ is **still** a basis for K.

Now consider the following fact about discriminants of algebraic integers:

Fact: If $\gamma_1, \gamma_2, ..., \gamma_d \in \mathcal{O}_K$ are a basis for K then $\Delta(\gamma_1, \gamma_2, ..., \gamma_d) \in \mathbb{Z}\backslash\{0\}$.

It follows that $D = |\Delta(\beta_1, \beta_2, ..., \beta_d)| \geq 1$ and $D' = |\Delta(\beta_1', \beta_2, ..., \beta_d)| \geq 1$. If we can show that $D > D' \geq 1$ then we would be done, since we may repeat the same argument with the new basis $\{\beta_1', \beta_2, ..., \beta_d\}$. This is either a $\mathbb{Z}$-basis for $\mathcal{O}_K$ or we would produce a third basis which has absolute value of discriminant D'' such that $D > D' > D'' \geq 1$. Repeating the argument cannot keep producing new bases with smaller and smaller positive integer discriminants forever...at some point the process would have to terminate, hence giving a $\mathbb{Z}$-basis for $\mathcal{O}_K$.

Let's show the claim that $D > D' \geq 1$. The change of basis matrix from $\{\beta_1, \beta_2, ..., \beta_d\}$ to $\{\beta_1', \beta_2, .., \beta_d\}$ is:

$$M = \begin{pmatrix} a_1 - \lfloor a_1 \rfloor & a_2 & \cdots & a_d \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

and so (via change of basis properties of the discriminant):

$$1 \leq D' = |\det(M)|^2 D = (a_1 - \lfloor a_1 \rfloor)^2 D < D$$

as required. □

Example 3.37. *We already saw examples of this theorem previously. For example, if $K = \mathbb{Q}(i)$ then*

$$\mathcal{O}_K = \mathbb{Z}[i] = \{a + bi \,|\, a, b \in \mathbb{Z}\}$$

has a $\mathbb{Z}$-basis $\{1, i\}$ of size $2 = [\mathbb{Q}(i) : \mathbb{Q}]$.

If $K = \mathbb{Q}(\sqrt{-3})$ then

$$\mathcal{O}_K = \mathbb{Z}\left[\frac{1+\sqrt{-3}}{2}\right] = \left\{a + b\left(\frac{1+\sqrt{-3}}{2}\right) \,\middle|\, a, b \in \mathbb{Z}\right\}$$

has a $\mathbb{Z}$-basis $\left\{1, \frac{1+\sqrt{-3}}{2}\right\}$ of size $2 = [\mathbb{Q}(\sqrt{-3}) : \mathbb{Q}]$.

4 Factorization in Rings of Integers (3 lectures)

As we saw in the introduction, if we're going to use rings like $\mathcal{O}_K$ to solve problems in Number Theory then we need to understand how factorisation works in these rings.

4.1 Existence of factorisation into irreducibles

Let's begin by clarifying exactly what the factorisations of interest are and what kinds of rings we want to work with. First, it sounds weird but we shouldn't be working in rings where zero has a factorisation into non-zero elements (that would be weird and tricky to deal with!).

Definition 4.1. *An **integral domain** is a commutative ring R such that if $\alpha\beta = 0$ for $\alpha, \beta \in R$ then $\alpha = 0$ or $\beta = 0$.*

Example 4.2. *The following are examples of integral domains (see Examples):*

- *Every field K.*
- *Every polynomial ring $R[x_1, x_2, ..., x_n]$ over an integral domain R.*
- *The ring of integers $\mathcal{O}_K$ of every number field K (e.g. $\mathbb{Z}$).*

Definition 4.3. *Let R be a commutative ring. An element $\alpha \in R$ is a **unit** if has a multiplicative inverse, i.e. there exists $\beta \in R$ such that $\alpha\beta = 1$. The set of units in R is denoted $R^\times$.*

Definition 4.4. *Let R be an integral domain. An element $\alpha \in R\backslash(R^\times \cup \{0\})$:*

- *is **reducible** if we can write $\alpha = \beta\gamma$ for two non-units $\beta, \gamma \in R$. Otherwise it is **irreducible**.*
- ***admits a factorisation into irreducibles** if $\alpha = \pi_1\pi_2...\pi_n$ for a (finite number of) irreducibles $\pi_i \in R$.*

*We say that R **admits factorisation into irreducibles** if every $\alpha \in R\backslash(R^\times \cup \{0\})$ admits a factorisation into irreducibles.*

Example 4.5. *When $R = \mathbb{Z}$ we recover the usual notion of factorisation:*

- *The units are the elements $\mathbb{Z}^\times = \{\pm 1\}$.*
- *The irreducible elements are $\pm p$ where p is prime, since by definition a prime number is a number $p \in \mathbb{Z}_{\geq 2}$ that can only factor as $p = (\pm 1) \cdot (\pm p)$, all of which involve units. The reducible elements are $\pm n$ where n is composite.*
- *You've known for a while that every $n \in \mathbb{Z}\backslash\{0, \pm 1\}$ admits a factorisation into irreducibles, since $n = \pm p_1 p_2 ... p_n$ for a (finite number of) primes.*

Lemma 4.6. *Let R be a commutative ring. The set $R^\times$ is an abelian group under multiplication.*

Proof. See Exercises. □

Our aim is to prove that the ring of integers $\mathcal{O}_K$ of a number field K always admits factorisation into irreducibles. To get there, let's remind ourself of how we prove that such factorisations exist in $\mathbb{Z}$.

Lemma 4.7. *The ring $\mathbb{Z}$ admits factorisation into irreducibles.*

Proof. Suppose that $\mathbb{Z}$ does not admit factorisation into irreducibles and pick an integer $n \in \mathbb{Z}\backslash\{0, \pm 1\}$ that does not have such a factorisation and has smallest absolute value.

We know that n must be reducible (otherwise it would itself be a factorisation into irreducibles!) and so $n = ab$ for two non-units $a, b \in \mathbb{Z}$. However, we have that $1 < |a|, |b| < |n|$ and by minimality of $|n|$ there must exist factorisations:

$$a = \pm p_1 p_2 ... p_m \qquad b = \pm q_1 q_2 ... q_n$$

for some primes p_i, q_i.

We now see that:

$$n = \pm(p_1 p_2 ... p_m)(q_1 q_2 ... q_n)$$

has a factorisation into primes, hence we have constructed a factorisation of n into irreducibles. This is a contradiction, hence $\mathbb{Z}$ admits factorisation into irreducibles. □

The above proof relied heavily on the ability to choose a smallest element that doesn't have a factorisation. However, for a general integral domain what does "small" mean (e.g. what is a "small" polynomial?). It turns out the best way to proceed is to use the language of ideals.

Definition 4.8. *Let R be a commutative ring. An **ideal** is a subset $I \subset R$ that is a subgroup under addition satisfying $rI \subset I$ for every $r \in R$.*

Example 4.9. *If $R = \mathbb{Z}$ and $n \in \mathbb{Z}$ then $I = n\mathbb{Z}$ is an ideal, since it's a subgroup under addition and any **integer** multiple of a multiple of n is another multiple of n. As we'll see in a later section, every ideal of $\mathbb{Z}$ is of this form (due to the existence of Euclid's Algorithm).*

Taking inspiration from the above example, we can always create ideals of a ring by looking at the set of multiples of an element.

Definition 4.10. *Let R be a commutative ring and $\beta \in R$. The **principal ideal** generated by β is $\langle\beta\rangle = \{\gamma\beta \,|\, \gamma \in R\}$.*

Here are some nice properties of ideals (see Exercises for proofs):

Lemma 4.11. *Let R be a commutative ring and $x, y \in R$. Then:*

- $\langle x \rangle$ *is an ideal of* R.
- $\langle x \rangle \subseteq \langle y \rangle$ *if and only if* $y|x$.
- *If further R is an integral domain then $\langle x \rangle = \langle y \rangle$ if and only $x = uy$ for some unit $u \in R^\times$.*

Definition 4.12. *An integral domain R is* ***Noetherian*** *if every non-empty set S of ideals of R has a maximal element with respect to inclusion, i.e. there exists $I \in S$ that is not contained in any other $J \in S$.*

We're now in a position to prove that Noetherian integral domains have nice factorisation properties. Compare the following proof with what happened over $\mathbb{Z}$.

Theorem 4.13. *If R is a Noetherian integral domain then R admits factorisation into irreducibles.*

Proof. Suppose that R does not admit factorisation into irreducibles and consider the set of ideals:

$$S = \{\langle \beta \rangle \,|\, \beta \in R \backslash (R^\times \cup \{0\}) \text{ and } \beta \text{ does not admit factorisation into irreducibles } \}.$$

We know that S is non-empty and that R is Noetherian, and so there exists a maximal element $\langle \gamma \rangle \in S$.

We know that γ is reducible (otherwise γ would itself be a factorisation into irreducibles!), so that $\gamma = \gamma_1 \gamma_2$ for two non-units $\gamma_1, \gamma_2 \in R$. It follows that $\langle \gamma \rangle \subset \langle \gamma_1 \rangle$ and $\langle \gamma \rangle \subset \langle \gamma_2 \rangle$. Since $\langle \gamma \rangle \in S$ is maximal, we must have that $\langle \gamma_1 \rangle \notin S$ and $\langle \gamma_2 \rangle \notin S$. This means that $\gamma_1 = \pi_1 \pi_2 ... \pi_m$ and $\gamma_2 = \mu_1 \mu_2 ... \mu_n$ for some irreducibles $\pi_i, \mu_i \in \mathcal{O}_K$. Taking the product shows that:

$$\gamma = (\pi_1 \pi_2 ... \pi_m)(\mu_1 \mu_2 ... \mu_n),$$

admits a factorisation into irreducibles, which is a contradiction. Hence R must admit factorisation into irreducibles. □

The question now is how to tell if a given integral domain is Noetherian. There are a couple of alternative criteria that let you do this.

Definition 4.14. *Let R be a commutative ring and $\beta_1, \beta_2, ..., \beta_n \in R$. The* ***ideal generated by*** *$\beta_1, \beta_2, ..., \beta_n$ is:*

$$\langle \beta_1, \beta_2, ..., \beta_n \rangle = \{a_1 \beta_1 + a_2 \beta_2 + ... + a_n \beta_n \,|\, a_1, a_2, ..., a_n \in R\}.$$

If an ideal of R has this form then it is called ***finitely generated****.*

One easily checks that this is an ideal (see Exercises).

Proposition 4.15. *Let R be an integral domain. Then the following conditions are equivalent:*

- *R is Noetherian.*
- *Every ideal of R is **finitely generated**.*
- *Every infinite sequence of ideals $I_1, I_2, I_3...$ of R such that:*

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq ...$$

eventually stabilises, i.e. at some point we have $I_n = I_{n+1} = I_{n+2} = ...$

Proof. See Exercises. □

Remark 4.16. *Note that the entire ring $\mathbb{A}$ of algebraic integers does not admit factorisation into irreducibles. For example we can continue to factor $2 \in \mathbb{A}$ into reducible elements indefinitely as follows:*

$$2 = \sqrt{2} \cdot \sqrt{2} = (\sqrt[4]{2} \cdot \sqrt[4]{2}) \cdot (\sqrt[4]{2} \cdot \sqrt[4]{2}) = (\sqrt[8]{2} \cdot \sqrt[8]{2}) \cdot (\sqrt[8]{2} \cdot \sqrt[8]{2}) \cdot (\sqrt[8]{2} \cdot \sqrt[8]{2}) \cdot (\sqrt[8]{2} \cdot \sqrt[8]{2}) = ...$$

The issue is that $\mathbb{A}$ is not Noetherian, this process produces an infinite sequence of ideals:

$$\langle 2 \rangle \subset \langle \sqrt{2} \rangle \subset \langle \sqrt[4]{2} \rangle \subset \langle \sqrt[8]{2} \rangle \subset ...$$

that cannot stabilise.

However, the rings $\mathcal{O}_K$ are much more well behaved then $\mathbb{A}$.

Theorem 4.17. *Let K be a number field. Then the ring of integers $\mathcal{O}_K$ is a Noetherian integral domain, hence admits factorisation into irreducibles.*

Proof. (Sketch) Let $I \subseteq \mathcal{O}_K$ be an arbitrary non-zero ideal. It is possible to show that the quotient ring $\mathcal{O}_K/I$ is finite, of size say $m \geq 1$ (see Exercises).

Pick a $\mathbb{Z}$-basis $\beta_1, \beta_2, .., \beta_d \in \mathcal{O}_K$ for $\mathcal{O}_K$ (which we know exists by Theorem 3.36). Scaling gives elements $m\beta_i \in I$ that are still a basis for K. The same proof as Theorem 3.36 then shows that I has a $\mathbb{Z}$-basis, i.e. there exists $\gamma_1, \gamma_2, ..., \gamma_d \in I$ such that $I = \{a_1\gamma_1 + a_2\gamma_2 + ... + a_d\gamma_d \,|\, a_1, a_2, ..., a_d \in \mathbb{Z}\}$.

We claim that $I = \langle \gamma_1, \gamma_2, ..., \gamma_d \rangle$, so that I is finitely generated. This would imply that $\mathcal{O}_K$ is a Noetherian intergral domain, and so admits unique factorisation into irreducibles.

The claim is clear since by the ideal properties:

$$I = \{a_1\gamma_1 + a_2\gamma_2 + ... + a_d\gamma_d \,|\, a_1, a_2, ..., a_d \in \mathbb{Z}\} \subseteq \{a_1\gamma_1 + a_2\gamma_2 + ... + a_d\gamma_d \,|\, a_1, a_2, ..., a_d \in \mathcal{O}_K\} \subseteq I$$

□

4.2 Unique Factorisation Domains

We have just proved that $\mathcal{O}_K$ admits factorisation into irreducibles. We wonder how **unique** such a factorisation is, since this is vital when trying to use such rings to solve Number Theoretic problems.

Let's once again go back to the integers and discuss this.

Example 4.18. *You've accepted that $\mathbb{Z}$ has unique factorisation for many years, but what does this really mean?*

If $n \in \mathbb{Z}\backslash\{0, \pm 1\}$ then we know that there exists a factorisation into irreducibles $n = p_1 p_2 ... p_n$. However, this is far from "unique" since we can always reorder the p_i to get a new factorisation into irreducibles (e.g. $15 = 3 \cdot 5 = 5 \cdot 3$).

Even if we fix an ordering of the p_i we can change some of the signs (i.e. multiply by units in $\mathbb{Z}^\times = \{\pm 1\}$) to give a new factorisation into irreducibles (e.g. $15 = 3 \cdot 5 = (-3) \cdot (-5)$.

This is true more generally. If R is an integral domain and $\alpha \in R\backslash(R^\times \cup \{0\})$ admits a factorisation into irreducibles $\alpha = \pi_1 \pi_2 ... \pi_n$ then we can:

- Reorder the π_i to get a new factorisation into irreducibles.
- Multiply any of the π_i by a unit $u \in R^\times$, as long as we multiply another π_j by u^{-1}.

Both of these moves are regarded as "cheating" and we are interested in whether two genuinely different factorisations into irreducibles can occur.

Definition 4.19. *Let R be an integral domain and $\alpha \in R\backslash(R^\times \cup \{0\})$. We say that α has a **unique factorisation into irreducibles** if for every two factorisations into irreducibles:*

$$\alpha = \pi_1 \pi_2 ... \pi_m = \mu_1 \mu_2 ... \mu_n,$$

we have $m = n$, a permuation $\sigma \in S_m$ and units $u_i \in R^\times$ such that $\pi_i = u_i \mu_{\sigma(i)}$.

Definition 4.20. *Let R be an integral domain that admits factorisation into irreducibles. We say that R is a **Unique Factorisation Domain** (UFD) if every $\alpha \in R\backslash(R^\times \cup \{0\})$ has a unique factorisation into irreducibles.*

Naturally, we would like to know which integral domains are UFD's. Let's once again return to the case $R = \mathbb{Z}$ to get a feel for this.

Theorem 4.21. *The commutative ring $\mathbb{Z}$ is a UFD.*

Proof. We know that $\mathbb{Z}$ is an integral domain that admits factorisation into irreducibles. It then suffices to show that every element $n \in \mathbb{Z}\backslash\{0, \pm 1\}$ has unique factorisation into irreducibles.

Suppose $n = p_1 p_2 ... p_m = q_1 q_2 ... q_n$ are two factorisations into irreducibles, with $n \geq m$. The crucial ingredient is the following fact that you take for granted (see Exercises):

FACT: If $p \in \mathbb{Z}$ is irreducible then it satisfies the property that $p|ab$ implies $p|a$ or $p|b$.

Then $p_1|n = q_1q_2...q_n$, and since p_1 is prime we have that $p_1|q_i$ for some i. But q_i is irreducible and so $q_i = u_1p_1$ for some unit $u_1 \in \mathbb{Z}^\times = \{\pm 1\}$.

We can now cancel p_1 from both sides of the equation to get:

$$n' = p_2p_3...p_m = u_1q_1q_2...q_{i-1}q_{i+1}...q_n$$

We continue to do the same with p_2, since this must now divide the RHS. Once again, we find that $q_j = u_2p_2$ for some $j \neq i$ and some unit $u_2 \in \mathbb{Z}^\times = \{\pm 1\}$. Cancelling removes another irreducible from both the LHS and RHS:

$$n'' = p_3p_4...p_m = (u_1u_2)q_1q_2...q_{j-1}q_{j+1}...q_{i-1}q_{i+1}...q_n$$

If we continue cancelling irreducibles in this fashion we will end up with a factorisation:

$$1 = uq_{i_1}q_{i_2}...q_{i_{n-m}}$$

for some unit $u \in \mathbb{Z}^\times = \{\pm 1\}$. If $n - m \neq 0$ then we would have that $q_{i_1}|1$, which implies that $q_{i_1} \in \mathbb{Z}^\times = \{\pm 1\}$, This is a contradiction, so that $n = m$ and we have shown that the two factorisations are equivalent. □

Since $\mathbb{Z}$ is a UFD, we wonder whether $\mathcal{O}_K$ is a UFD in general. Unfortunately, the answer is no!

Example 4.22. *Let $K = \mathbb{Q}(\sqrt{-5})$, which has ring of integers $\mathcal{O}_K = \mathbb{Z}[\sqrt{-5}]$. Consider the factorisations:*

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

The units of this ring are $\mathbb{Z}[\sqrt{-5}]^\times = \{\pm 1\}$ (we'll see why in a later section). These are the elements satisfying $N(\alpha) = N(a + b\sqrt{-5}) = a^2 + 5b^2 = 1$.

We can easily check that each factor in these factorisations is irreducible as follows:

- *2 is irreducible since if $2 = \alpha\beta$ is a non-trivial factorisation then $4 = N(2) = N(\alpha\beta) = N(\alpha)N(\beta)$. But $N(\alpha), N(\beta) \in \mathbb{Z}\backslash\{1\}$ and so $N(\alpha) = N(\beta) = 2$. However, there are no such elements since there are no integer solutions to $a^2 + 5b^2 = 2$.*
- *3 is irreducible since if $3 = \alpha\beta$ is a non-trivial factorisation then $9 = N(\alpha)N(\beta)$. But $N(\alpha), N(\beta) \in \mathbb{Z}\backslash\{1\}$ and so $N(\alpha) = N(\beta) = 3$. However, there are no such elements since there are no integer solutions to $a^2 + 5b^2 = 3$.*
- *$(1 \pm \sqrt{-5})$ are irreducible since if $(1 \pm \sqrt{-5}) = \alpha\beta$ then $6 = N(\alpha)N(\beta)$. But $N(\alpha), N(\beta) \in \mathbb{Z}\backslash\{1\}$ and so $N(\alpha) = 2$ and $N(\beta) = 3$ or vice verse. However, we just showed that no such elements exist.*

*It follows that $\mathbb{Z}[\sqrt{-5}]$ is **not** a UFD. The underlying issue is that we have an irreducible 2 that doesn't satisfy the key property that we used in the proof above. We have that $2|6 = (1+\sqrt{-5})(1-\sqrt{-5})$ but $2 \nmid (1 \pm \sqrt{-5})$, i.e. we are unable to pair 2 with an irreducible on the RHS.*

Because of the issue just discussed, we need to be careful to distinguish the property of $\pi \in R$ being irreducible from the property that $\pi|\alpha\beta$ implies that $\pi|\alpha$ or $\pi|\beta$.

Definition 4.23. *Let R be a commutative ring. We say that $\pi \in R\backslash(R^\times \cup \{0\})$ is **prime** if whenever $\pi|\alpha\beta$ for some $\alpha, \beta \in R$ then $\pi|\alpha$ or $\pi|\beta$.*

In $\mathbb{Z}$ the irreducibles and the primes are the same thing...and this is why $\mathbb{Z}$ is a UFD. In general, this correspondence only goes one way, so that UFD can fail.

Lemma 4.24. *Let R be an integral domain. Then every prime $\pi \in R$ is irreducible.*

Proof. See Exercises. □

We can now state our main theorem on UFD's.

Proposition 4.25. *Let R be an integral domain that admits unique factorisation into irreducibles. Then R is a UFD if and only if every irreducible is prime.*

Proof. If every irreducible is prime then one can easily adapt the argument that $\mathbb{Z}$ is a UFD.

Suppose $\alpha \in R\backslash(R^\times \cup \{0\})$ has two factorisations $\alpha = \pi_1\pi_2...\pi_m = \mu_1\mu_2...\mu_n$ into irreducibles, with $n \geq m$. Then $\pi_1|\alpha = \mu_1\mu_2...\mu_n$, and since π_1 is prime we have that $\pi_1|\mu_i$ for some i. But μ_i is irreducible and so $\mu_i = u_1\pi_1$ for some unit $u_1 \in R^\times$.

We can now cancel π_1 from both sides of the equation to get:

$$\alpha' = \pi_2\pi_3...\pi_m = u_1\mu_1\mu_2...\mu_{i-1}\mu_{i+1}...\mu_n$$

Similarly to earlier, we can continue to do this until we reach a factorisation

$$1 = u\mu_{i_1}\mu_{i_2}...\mu_{i_{n-m}}$$

for some unit $u \in \mathbb{Z}^\times = \{\pm 1\}$. If $n - m \neq 0$ then we would have that $\mu_{i_1}|1$, which implies that $\mu_{i_1} \in R^\times$, This is a contradiction, so that $n = m$ and we have shown that the two factorisations are equivalent.

Conversely, suppose that there is an irreducible $\pi \in R$ that is not prime. Then there exist $\alpha, \beta \in R$ such that $\pi|\alpha\beta$ but $\pi \nmid \alpha$ and $\pi \nmid \beta$. It follows that $\alpha\beta = \pi\gamma$ for some $\gamma \in R$.

We must have that $\alpha \neq 0$ and $\beta \neq 0$, since otherwise we would have $\pi|\alpha$ or $\pi|\beta$ (since $\pi|0$ is true). We cannot have that $\gamma = 0$, since this would imply that $\alpha = 0$ or $\beta = 0$ (since R is an integral domain).

We also cannot have that $\alpha \in R^\times$ or $\beta \in R^\times$ since then $\pi|\alpha\beta$ would imply that $\pi|\alpha$ or $\pi|\beta$. We cannot have that $\gamma \in R^\times$ since this would imply that one of $\alpha, \beta \in R^\times$ (since π is irreducible).

It follows that we have two factorisations $\alpha\beta = \pi\gamma$ with $\alpha, \beta, \pi, \gamma \in R\backslash(R^\times \cup \{0\})$. Since R admits unique factorisation, we can factor α, β, γ into irreducibles as follows:

$$\alpha = \mu_1\mu_2...\mu_n$$
$$\beta = \eta_1\eta_2...\eta_m$$
$$\gamma = \psi_1\psi_2...\psi_t$$

Now we have the two factorisations into irreducibles:

$$(\mu_1\mu_2...\mu_n)(\eta_1\eta_2...\eta_m) = \pi(\psi_1\psi_2...\psi_t).$$

These are inequivalent factorisations, since if $\pi = u\mu_i$ for some $u \in R^\times$ then $\pi|\alpha$ and if $\pi = \psi_i$ for some $u \in R^\times$ then $\pi|\beta$. This contradicts the above assumption on π, α, β. Hence R is not a UFD. □

Now that we have a much better understanding of when $\mathcal{O}_K$ is a UFD, one wonders how often this is the case! The following is a long standing conjecture that is widely believed.

Conjecture 4.26. *There are infinitely many number fields K such that $\mathcal{O}_K$ is a UFD.*

Such conjectures have been studied for a long time, and we have an awful lot of heuristics, but so far exact proofs have resisted us!

The special case of **imaginary quadratic fields** was originally studied by Gauss in the 1800's. He conjectured the following, which is now a theorem due to Stark and Heegner (using sophisticated methods way beyond this course!).

Theorem 4.27. *If $d \geq 1$ is square-free and $K = \mathbb{Q}(\sqrt{-d})$ then $\mathcal{O}_K$ is a UFD if and only if*

$$d = 1, 2, 3, 7, 11, 19, 43, 67, 163.$$

People have wondered for a long time whether we may find infinitely many UFD's among the **real quadratic fields**.

Conjecture 4.28. *There are infinitely many square-free integers $d > 1$ such that the ring of integers $\mathcal{O}_K$ of $K = \mathbb{Q}(\sqrt{d})$ is a UFD.*

4.3 Euclidean Domains

So far we've seen that $\mathcal{O}_K$ admits unique factorisation into irreducibles, and that **sometimes** it's unique, i.e. if and only if all irreducibles are prime. However, this seems like a bit of a cheat...since finding an explicit irreducible that isn't prime is essentially equivalent to finding an instance of non-unique factorisation.

Question: Are there alternative ways to prove that $\mathcal{O}_K$ is a UFD?

It turns out that the answer is yes...if you have an analogue of Euclid's Algorithm!

Let's once again return to $\mathbb{Z}$ and see why this case works smoothly.

Example 4.29. *In $\mathbb{Z}$ we can prove that every irreducible is prime quite easily. Suppose that $p \in \mathbb{Z}\backslash\{0, \pm 1\}$ is irreducible and that $p|ab$ for some $a, b \in \mathbb{Z}$.*

If $p|a$ then p is prime. If $p \nmid a$ then we must have that $gcd(p, a) = 1$, since the only divisors of p are ± 1 and $\pm p$ (i.e. the units are $\mathbb{Z}^\times = \{\pm 1\}$) and $\pm p \nmid a$.

Euclid's Algorithm lets us find $s, t \in \mathbb{Z}$ such that $sa + tp = 1$. multiplying through by b gives $sab + tpb = b$. Since the LHS is a multiple of p, so must the RHS, i.e. $p|b$. Hence p is prime.

The key steps in this proof come from the fact that there is a well defined notion of gcd in $\mathbb{Z}$ and that Euclid's Algorithm can be used to write $sa + tp = \gcd(p, a) = 1$.

There are more general notions of gcd for other rings, but these are a little messy. It turns out that the better way to proceed is to instead use ideals, since the fact that there exists $sa + tp = \gcd(p, a) = 1$ is equivalent to the fact that the ideal of $\mathbb{Z}$ given by $\langle a, p\rangle = a\mathbb{Z} + p\mathbb{Z} = \mathbb{Z} = \langle 1\rangle$ is a principal ideal.

Definition 4.30. *Let R be an integral domain. Then R is a **Euclidean Domain** (ED) if there exists a function $\delta : R\backslash\{0\} \to \mathbb{N}_{\geq 0}$ with the property that whenever $\alpha, \beta \in R$ and $\beta \neq 0$ then there exist $q, r \in R$ satisfying $\alpha = q\beta + r$ with $r = 0$ or $0 \leq \delta(r) < \delta(\beta)$.*

Example 4.31. *We've seen two examples of Euclidean Domains before:*

- *Firstly, $\mathbb{Z}$ is a Euclidean Domain with $\delta(n) = |n|$ (the remainder when we divide one non-zero integer into another integer is either 0 or is smaller in absolute value).*
- *Secondly, for any field K we have that $K[x]$ is a Euclidean Domain with $\delta(f(x)) = deg(f(x))$ (the remainder when we divide one non-zero polynomial in $K[x]$ into another polynomial in $K[x]$ is either 0 or is smaller in degree).*

Definition 4.32. *Let R be an integral domain. Then R is a **Principal Ideal Domain** (PID) if every ideal $I \subseteq R$ is principal, i.e. $I = \langle x\rangle$ for some $x \in R$.*

Step one of proving that Euclidean Domains are UFD's is the following:

Proposition 4.33. *Every ED is a PID*

Proof. Suppose that R is a ED and that $I \subseteq R$ is an ideal. If $I = \{0\}$ then $I = \langle 0 \rangle$ is principal. If $I \neq \{0\}$ then we can pick $x \in I \backslash \{0\}$ such that $f(x)$ is minimal (since f takes values in $\mathbb{N}_{\geq 0}$ such a minimum must exist). We claim that $I = \langle x \rangle$ is a principal ideal.

To prove the claim, let $y \in I$ be arbitrary. Since R is a ED we can find $q, r \in R$ such that $y = qx + r$ with $r = 0$ or $0 \leq \delta(r) < \delta(x)$. The second case cannot occur since then we would have found an element $r = y - qx \in I$ such that $\delta(r) < \delta(x)$, contradicting the minimality of $\delta(x)$. Thus $r = 0$ and so $y = qx$, i.e. $y \in \langle x \rangle$. It follows that $I \subseteq \langle x \rangle$. The reverse inclusion $\langle x \rangle \subseteq I$ follows since $x \in I$ and I is an ideal. Hence $I = \langle x \rangle$. □

Example 4.34. *Earlier I promised that every ideal of $\mathbb{Z}$ is of the form $m\mathbb{Z}$, i.e. is principal. This is because $\mathbb{Z}$ is a ED, hence is a PID.*

We're now ready for step two of our main result.

Proposition 4.35. *Every PID is a UFD.*

Proof. Let R be a PID. Since every ideal is principal, every ideal is finitely generated (it has one generator). Hence R is Noetherian, so that R admits factorisation into irreducibles.

It suffices to prove that every irreducible is prime. To this end, suppose that $\pi \in R \backslash (R^\times \cup \{0\})$ is irreducible and that $\pi | \alpha\beta$ for $\alpha, \beta \in R$.

If $\pi | \alpha$ then π is prime. If $\pi \nmid \alpha$ then the ideal $\langle \alpha \rangle$ of R is not contained in the ideal $\langle \pi \rangle$ of R. Consider the ideal $\langle \pi, \alpha \rangle$ of R. First note that $\langle \pi, \alpha \rangle \supset \langle \pi \rangle$. Second note that $\langle \pi, \alpha \rangle = \langle \gamma \rangle$ for some $\beta \in R$ (since R is a PID).

It follows that $\langle \gamma \rangle \supset \langle \pi \rangle$, so that $\gamma | \pi$ but $\gamma \neq u\pi$ for any unit $u \in R^\times$. Since π is irreducible the only other possibility is that $\gamma \in R^\times$ (otherwise we would have found a factorisation of π into non-units).

We now find that $\langle \pi, \alpha \rangle = \langle \gamma \rangle = \langle 1 \rangle$, since $\beta \in R^\times$. This means that there exist $s, t \in R$ such that $s\alpha + t\pi = 1$. Multiplying by β gives $s\alpha\beta + t\pi = \beta$ and the LHS being divisible by π implies that the RHS is. Hence $\pi | \beta$ showing that π is prime. □

Remark 4.36. *It might not seem like it at first glance but this proof mimics the proof we saw for $\mathbb{Z}$. Indeed the ideal $\langle \pi, \alpha \rangle$ of R is the direct analogue of the ideal $\langle p, a \rangle$ of $\mathbb{Z}$. The assumption that this ideal is principal is akin to the fact that $\langle p, a \rangle = \langle gcd(p, a) \rangle$ over $\mathbb{Z}$ (a fact provided by being a ED).*

Example 4.37. *We've just seen that if an integral domain R is a PID then it is a UFD. The converse of this statement is generally not true. For example $R = \mathbb{Q}[x, y]$ is a UFD but it is not a PID (e.g. the ideal $I = \langle x, y \rangle$ is not principal).*

However, it is possible to show that for a number field K the ring of integers $\mathcal{O}_K$ is a UFD if and only if it is a PID. The proof of this is beyond the course, but it highlights another way of showing thata specific ring of integers $\mathcal{O}_K$ is not a UFD...find an ideal that is not principal!

For example, the ideal $\langle 2, 1+\sqrt{-5}\rangle \subset \mathbb{Z}[\sqrt{-5}]$ is not principal. If it were, say $\langle 2, 1+\sqrt{-5}\rangle = \langle \alpha\rangle$ for some $\alpha \in \mathbb{Z}[\sqrt{-5}]$ then we would have that $\langle 2\rangle \subseteq \langle \alpha\rangle$ and $\langle 1+\sqrt{-5}\rangle \subseteq \langle \alpha\rangle$. This would imply that $\alpha|2$ and $\alpha|(1+\sqrt{-5})$, and taking norms would show that $N(\alpha)|N(2) = 4$ and $N(\alpha)|N(1+\sqrt{-5}) = 6$, so that $N(\alpha) = 1$ or 2. In the first case we would have $\alpha = \pm 1 \in \mathbb{Z}[\sqrt{-5}]^\times$, which isn't possible (since $\langle 2, 1+\sqrt{-5}\rangle \neq \langle 1\rangle = \mathbb{Z}[\sqrt{-5}]$. The second case is also impossible since there are no integer solutions to $x^2 + 5y^2 = 2$ (as we saw earlier).

The fact that this ideal is not principal shows that $\mathbb{Z}[\sqrt{-5}]$ is not a UFD. Compare this with the specific non-unique factorisation that we considered before:

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

The fact that this ideal is not principal is essentially equivalent to the fact that the factors 2 and $(1 + \sqrt{-5})$ have no well defined gcd (hence that there is no cancellation from either side that can "correct" the factorisation).

Corollary 4.38. *Every ED is a UFD.*

Proof. This is clear from the previous two propositions. □

Let's now make good on one of our promises in the introduction.

Proposition 4.39. *We have that $\mathbb{Z}[i]$ is a UFD.*

Proof. We do this by showing that $\mathbb{Z}[i]$ is a ED with function $\delta(\alpha) = N(\alpha) \in \mathbb{N}_{\geq 0}$.

Let $\alpha, \beta \in \mathbb{Z}[i]$ and $\beta \neq 0$. We must produce $q, r \in \mathbb{Z}[i]$ such that $\alpha = q\beta + r$ and either $r = 0$ or $0 \leq N(r) < N(\beta)$.

Consider $\frac{\alpha}{\beta} = a + bi \in \mathbb{Q}(i)$. We can round off the coefficients to produce $q = \lfloor a \rceil + \lfloor b \rceil i \in \mathbb{Z}[i]$ and consider the remainder $r = \alpha - q\beta \in \mathbb{Z}[i]$. We then certainly have that $\alpha = q\beta + r$ and so it suffices to show that $r = 0$ or $0 \leq N(r) < N(\beta)$.

If $r = 0$ then we are done. Consider:

$$N\left(\frac{r}{\beta}\right) = N\left(\frac{\alpha - q\beta}{\beta}\right) = N\left(\frac{\alpha}{\beta} - q\right) = N((a - \lfloor a \rceil) + (b - \lfloor b \rceil)i) = (a - \lfloor a \rceil)^2 + (b - \lfloor b \rceil)^2.$$

We have that:

$$(a - \lfloor a \rceil)^2 + (b - \lfloor b \rceil)^2 \leq \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}.$$

It follows that $0 \leq N\left(\frac{r}{\beta}\right) \leq \frac{1}{2}$ and by multiplicativity of the norm we get that $0 \leq \frac{N(r)}{N(\beta)} \leq \frac{1}{2}$, so that $0 \leq N(r) \leq \frac{1}{2}N(\beta) < N(\beta)$, as required. □

Remark 4.40. *Earlier we remarked that $\mathcal{O}_K$ is a UFD if and only if it is a PID. This suggests that the more natural notion of factorisation for $\mathcal{O}_K$ might be in terms of ideals. Indeed, it is possible to multiply ideals and define analogues of irreducibles and primes (maximal ideals and prime ideals).*

One of the foundational results in Algebraic Number Theory is that $\mathcal{O}_K$ has unique factorisation on the level of ideals, i.e. each ideal $I \subseteq \mathcal{O}_K$ has a prime ideal factorisation $I = \mathfrak{p}_1\mathfrak{p}_2...\mathfrak{p}_g$ that is unique upto reordrering of the prime ideals $\mathfrak{p}_i$.

Remark 4.41. *We mentioned earlier that $\mathcal{O}_K$ is a PID if and only if UFD. We just saw that $\mathcal{O}_K$ being a ED implies both. One might wonder whether the converse is true, i.e. does $\mathcal{O}_K$ have unique factorisation **exactly** when it has a Euclidean Algorithm?*

*Unfortunately, the answer is **no**. For example, Motzkin showed in 1939 that the ring of integers $\mathcal{O}_K = \mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ of the number field $K = \mathbb{Q}(\sqrt{-19})$ is a UFD but is not a ED (i.e. there can be **no** Euclid's Algorithm for this ring). The proof of this result is beyond the scope of this course, but it is relatively easy to prove the special case that this ring is not a ED with respect to the norm function (see Exercises).*

In fact, it is quite rare to find that the ring of integers $\mathcal{O}_K$ of a quadratic field K is a ED:

- *For **imaginary quadratic fields**, i.e. $K = \mathbb{Q}(\sqrt{-d})$ with $d \geq 1$ square-free it is known that $\mathcal{O}_K$ is a ED with respect to the norm function if and only if $d = 1, 2, 3, 7, 11$.*
- *For **real quadratic fields**, i.e. $K = \mathbb{Q}(\sqrt{d})$ with $d > 1$ square-free it is known that $\mathcal{O}_K$ is a ED with respect to the absolute value of the norm function if and only if $d = 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73$.*

5 Dirichlet's Unit Theorem (1 lecture)

We saw in the previous section that the units of $\mathcal{O}_K$ play a significant role in understanding how elements of $\mathcal{O}_K$ factor into irreducibles. Naturally one asks the question, what **is** the unit group $\mathcal{O}_K^\times$? How **big** is $\mathcal{O}_K$?

Example 5.1. *If $K = \mathbb{Q}$ then $\mathcal{O}_K = \mathbb{Z}$ and we know that the unit group is $\mathbb{Z}^\times = \{\pm 1\}$.*

Example 5.2. *If $K = \mathbb{Q}[i]$ then $\mathcal{O}_K = \mathbb{Z}[i]$ and $a+ib \in \mathbb{Z}[i]^\times$ if and only if there exists $c+di \in \mathbb{Z}[i]$ such that $(a+bi)(c+di) = 1$. This implies the equations:*

$$ac - bd = 1$$
$$ad + bc = 0$$

If $a \neq 0$ then the second equation gives $d = -\frac{bc}{a}$, and the first then gives $ac + \frac{b^2c}{a} = 1$, i.e. $c(a^2+b^2) = a$. If $b \neq 0$ then $a^2 + b^2 > a^2 \geq |a|$, so we must have $b = 0$. Then we must have that and $a^2 = |a|$, so that $a = \pm 1$ (giving the units $\pm 1 \in \mathbb{Z}[i]^\times$).

If $a = 0$ we must have $-bd = 1$, implying that $b = \pm 1$ (giving the units $\pm i \in \mathbb{Z}[i]^\times$). Thus $\mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$.

It's clear that this kind of brute force argument is going to get very messy very quickly...the equations will become very complicated. Fortunately, there are better tools for understanding units.

Lemma 5.3. *Let K be a number field. Then $\alpha \in \mathcal{O}_K^\times$ if and only if $N(\alpha) = \pm 1$.*

Proof. Suppose that $\alpha \in \mathcal{O}_K^\times$. Then there exists $\beta \in \mathcal{O}_K$ such that $\alpha\beta = 1$. Taking norms gives $N(\alpha)N(\beta) = N(1) = 1$. Since $N(\alpha), N(\beta) \in \mathbb{Z}$ it follows that $N(\alpha) = \pm 1$.

Suppose that $N(\alpha) = \pm 1$ and that $\sigma_1, ..., \sigma_n$ are the embeddings of K into $\mathbb{C}$. Then $N(\alpha) = \prod_i \sigma_i(\alpha) = \pm 1$. We know that α appears in this product and by swapping the σ_i we may assume that $\sigma_1(\alpha) = \alpha$. Letting $x = \sigma_2(\alpha)\sigma_3(\alpha)...\sigma_d(\alpha)$ we see that $x \in \mathcal{O}_K$ (since each $\sigma_i(\alpha)$ is an algebraic integer (it's a root of $m_\alpha(x)$) and so the product x is algebraic integer. Also $x = \frac{N(\alpha)}{\alpha} \in K$). Hence $\alpha \in \mathcal{O}_K^\times$, since it has multiplicative inverse $\pm x \in \mathcal{O}_K$. □

Example 5.4. *This lemma makes the above computation much easier! If $a + ib \in \mathbb{Z}[i]^\times$ then we must have $N(a+ib) = a^2 + b^2 = \pm 1$. The only integer solutions to this equation are $(a, b) = (\pm 1, 0), (0, \pm 1)$, so that the units are $\mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$.*

In fact, we can now find the units of many quadratic rings of integers.

Proposition 5.5. *Let $d \geq 1$ be square-free and $K = \mathbb{Q}(\sqrt{-d})$. Then:*

$$\mathcal{O}_K^\times = \begin{cases} \{\pm 1, \pm i\} & \text{if } d = 1 \\ \left\{\pm 1, \pm\left(\frac{-1+\sqrt{-3}}{2}\right), \pm\left(\frac{-1-\sqrt{-3}}{2}\right)\right\} & \text{if } d = 3 \\ \{\pm 1\} & \text{otherwise} \end{cases}$$

Proof. If $-d \not\equiv 3 \bmod 4$ then $\mathcal{O}_K = \mathbb{Z}[\sqrt{-d}]$ and $a + b\sqrt{-d} \in \mathbb{Z}[\sqrt{-d}]^\times$ implies that

$$N(a + b\sqrt{-d}) = a^2 + db^2 = \pm 1.$$

If $d > 1$ then the only integer solutions are $(a, b) = (\pm 1, 0)$, so that $\mathbb{Z}[\sqrt{-d}]^\times = \{\pm 1\}$. If $d = 1$ then we saw above that $\mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$.

If $-d \equiv 3 \bmod 4$ then $\mathcal{O}_K = \mathbb{Z}\left[\frac{1+\sqrt{-d}}{2}\right]$ and $a + b\left(\frac{1+\sqrt{-d}}{2}\right) \in \mathbb{Z}\left[\frac{1+\sqrt{-d}}{2}\right]^\times$ implies that

$$N\left(a + b\left(\frac{1+\sqrt{-d}}{2}\right)\right) = \frac{(2a+b)^2 + db^2}{4} = \pm 1,$$

i.e. $(2a + b)^2 + db^2 = \pm 4$. If $d > 4$ then we must again have $b = 0$, so that $4a^2 = 4$, i.e. $(a, b) = (\pm 1, 0)$. This implies that $\mathbb{Z}\left[\frac{1+\sqrt{-d}}{2}\right]^\times = \{\pm 1\}$. If $d = -3$ (the only remaining case) then $(2a + b)^2 + 3b^2 = \pm 4$, the integer solutions are then $(a, b) = (\pm 1, 0), (0, 1), (-1, 1), (1, -1), (0, -1)$, giving $\mathbb{Z}\left[\frac{1+\sqrt{-3}}{2}\right]^\times = \left\{\pm 1, \pm\left(\frac{-1+\sqrt{-3}}{2}\right), \pm\left(\frac{-1-\sqrt{-3}}{2}\right)\right\}$. □

Naturally, we wonder what happens when we instead look at the number fields $\mathbb{Q}(\sqrt{d})$ with $d > 0$ square-free. In this case, the unit group is infinite!

Example 5.6. *Consider the number field $K = \mathbb{Q}(\sqrt{2})$, so that $\mathcal{O}_K = \mathbb{Z}[\sqrt{2}]$. There are the usual units $\pm 1 \in \mathbb{Z}[\sqrt{2}]^\times$, but there's also the unit $u = 1 + \sqrt{2}$, which has norm $N(1 + \sqrt{2}) = -1$. Since $u > 1$, the powers u^m are all distinct and give units. This tells us that $\mathbb{Z}[\sqrt{2}]^\times \supseteq \{\pm u^m \,|\, m \in \mathbb{Z}\}$. In fact, this is the entire unit group, as we will see soon!*

Theorem 5.7. *Let $d > 1$ be square-free and $K = \mathbb{Q}(\sqrt{d})$. If there exists a unit $u \in \mathcal{O}_K^\times \backslash \{\pm 1\}$ then $\mathcal{O}_K^\times = \{\pm v^m \,|\, m \in \mathbb{Z}\}$ for some unit $v \in \mathcal{O}_K^\times$.*

Proof. Note that any such unit $u \in \mathcal{O}_K^\times \backslash \{\pm 1\}$ gives rise to the four units $\{u, -u, u^{-1}, -u^{-1}\}$, and exactly one of these is greater than 1. Thus, it suffices to show that every unit greater than 1 can be written as v^m for some fixed unit $v > 1$ and $m \geq 1$.

We let v is the **smallest** unit satisfying $v > 1$ (this exists, see Exercises). Then any unit $u > 1$ must satisfy $v^m \leq u < v^{m+1}$ for some $m \geq 1$. Dividing by v^m (which is allowed, since v is a unit!) gives $1 \leq \frac{u}{v^m} < v$. However, $\frac{u}{v^m}$ is a unit (since $N\left(\frac{u}{v^m}\right) = N(u)N(v)^m = \pm 1$) and we cannot have that $1 < \frac{u}{v^m} < v$, otherwise we would contradict minimality of v. It must be that $\frac{u}{v^m} = 1$, i.e. $u = v^m$, as required. □

Definition 5.8. *The unit v in the above theorem is called a **fundamental unit**.*

Example 5.9. *Returning to the previous example, we claim that $v = 1 + \sqrt{2}$ is a fundamental unit for $\mathbb{Z}[\sqrt{2}]$, so that $\mathbb{Z}[\sqrt{2}]^\times = \{\pm(1 + \sqrt{2})^m | m \in \mathbb{Z}\}$. As per the proof given above, the claim will follow if we can show that $1 + \sqrt{2}$ is the smallest unit in $\mathbb{Z}[\sqrt{2}]^\times$ that is bigger than 1.*

Suppose that $w = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]^\times$ satisfies $1 < w < 1 + \sqrt{2}$ and so is a smaller such unit. Since $N(a + b\sqrt{2}) = (a + b\sqrt{2})(a - b\sqrt{2}) = \pm 1$ we must have that $-1 < a - b\sqrt{2} < 1$. Then the trace satisfies $0 < T(w) = 2a < 2 + \sqrt{2} < 4$, so that $a = 1$ (since $a \in \mathbb{Z}$). We then must have $N(w) = a^2 - 2b^2 = 1 - 2b^2 = \pm 1$, so that $b = 0, \pm 1$. Hence $w = 1$ or $w = 1 \pm \sqrt{2}$, which are all invalid. It follows that w cannot exist, hence that $1 + \sqrt{2}$ is a fundamental unit for $\mathbb{Z}[\sqrt{2}]^\times$.

Let's take stock of where we are. The above theorem has proved that when $K = \mathbb{Q}(\sqrt{d})$ with $d > 1$ square-free then the unit group is either $\{\pm 1\}$ or is $\{\pm v^m | m \in \mathbb{Z}\}$ for some fundamental unit.

Proposition 5.10. *If $K = \mathbb{Q}(\sqrt{d})$ with $d > 1$ square-free then there exists a fundamental unit $v \in \mathcal{O}_K^\times$, so that $\mathcal{O}_K^\times = \{\pm v^m | m \in \mathbb{Z}\}$.*

Proof. Omitted. The easiest way to prove this is to use continued fractions to construct a fundamental unit (see any good book on Algebraic Number Theory). □

The above proposition is a special case of a more general result that tells us about the unit group of **any** ring of integers.

Theorem 5.11. *(Dirichlet's Unit Theorem) Let $K = \mathbb{Q}(\alpha)$ be a number field and suppose that α has r real conjugates and $2s$ complex conjugates. Then there exist fundemantal units $v_1, v_2, ..., v_{r+s-1} \in \mathcal{O}_K^\times$ such that*

$$\mathcal{O}_K^\times = \{\zeta v_1^{m_1} v_2^{m_2} ... v_{r+s-1}^{m_{r+s-1}} \,|\, \zeta \text{ is a root of unity in } K \text{ and } m_i \in \mathbb{Z}\}.$$

The proof of this result is beyond the scope of this course but can be found in any good book on Algebraic Number Theory.

Example 5.12. *The number field $K = \mathbb{Q}(\sqrt[3]{2})$ has $r = 1$ real conjugate and $2s = 2$ complex conjugates. Since $K \subset \mathbb{R}$, the roots of unity in K are $\{\pm 1\}$.*

Dirichlet's Unit Theorem tells us that there is $r + s - 1 = 1$ fundamental unit in $\mathcal{O}_K^\times = \mathbb{Z}[\sqrt[3]{2}]^\times$ (see Exercises for a proof of the fact that this is the ring of integers). In fact, it is possible to show that $v = 1 + \sqrt[3]{2} + \sqrt[3]{2}^2$ is a fundamental unit, so that:

$$\mathbb{Z}[\sqrt[3]{2}]^\times = \{\pm \left(1 + \sqrt[3]{2} + \sqrt[3]{2}^2\right)^m \,\Big|\, m \in \mathbb{Z}\}.$$

5.1 Pell's equation

Now that we understand the units in rings of integers of real quadratic fields, we can use this to understand the solutions of a famous Diophantine equation.

Definition 5.13. *Let $d > 0$ be non-square. The corresponding **Pell equation** is $x^2 - dy^2 = \pm 1$.*

We're interested in the integer solutions to this equation. By realising this as a norm equation, we can immediately link with the unit group of a certain ring.

Lemma 5.14. *Let $d > 1$ be non-square. Then $(x, y) \in \mathbb{Z}^2$ is a solution to the Pell equation $x^2 - dy^2 = \pm 1$ if and only if $(x + y\sqrt{d}) \in \mathbb{Z}[\sqrt{d}]^\times$.*

Proof. This is clear since:

$$x^2 - dy^2 = \pm 1 \iff (x + y\sqrt{d})(x - y\sqrt{d}) = \pm 1 \iff N(x + y\sqrt{d}) = \pm 1 \iff x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]^\times.$$

□

So solving Pell's equation is the same as finding the units of the ring $\mathbb{Z}[\sqrt{d}]$. This is something that we just did in Theorem 5.7 and Proposition 5.10...well almost!

There's a small issue, since those results only applied to the ring of integers of $\mathbb{Q}(\sqrt{d})$...which is great if d is square-free and $d \not\equiv 1 \bmod 4$. However, for other values of d the ring of integers of $\mathbb{Q}(\sqrt{d})$ is **not** $\mathbb{Z}[\sqrt{d}]$. It turns out that this isn't much of an issue though.

Theorem 5.15. *Let $d > 1$ be non-square. Then the ring $\mathbb{Z}[\sqrt{d}]$ has unit group $\mathbb{Z}[\sqrt{d}]^\times = \{\pm v^m | m \in \mathbb{Z}\}$ for some fundamental unit v.*

Proof. The exact same proof as Theorem 5.7 works for the ring $\mathbb{Z}[\sqrt{d}]$ (nowhere in the proof did we use any of the properties of the ring of integers). The existence of a fundamental unit can still be proved using continued fractions and is again omitted. □

We're now able to fully understand the solutions to Pell's equation.

Corollary 5.16. *Let $d > 1$ be non-square. Then there are infinitely many solutions to the Pell equation $x^2 - dy^2 = \pm 1$.*

Proof. The above lemma tells us that the integer solutions to the Pell equation $x^2 - dy^2 = \pm 1$ are in bijection with units in the ring $\mathbb{Z}[\sqrt{d}]$. The theorem that we just states shows that there are infinitely many units in this ring, and so there must be infinitely many solutions to the Pell equation. □

Example 5.17. *The above lemma tells us that the Pell equation $x^2 - 2y^2 = \pm 1$ is intimately related to the units of $\mathbb{Z}[\sqrt{2}]$. We saw earlier that $\mathbb{Z}[\sqrt{2}]^\times = \{\pm(1+\sqrt{2})^m | m \in \mathbb{Z}\}$. Let's use this to write down some solutions to Pell's equation:*

m	*Unit*	(x, y)
0	± 1	$\pm(1, 0)$
1	$\pm(1+\sqrt{2})$	$\pm(1, 1)$
2	$\pm(1+\sqrt{2})^2 = \pm(3+2\sqrt{2})$	$\pm(3, 2)$
3	$\pm(1+\sqrt{2})^3 = \pm(7+5\sqrt{2})$	$\pm(7, 5)$
-1	$\pm(1+\sqrt{2})^{-1} = \pm(1-\sqrt{2})$	$\pm(1, -1)$
-2	$\pm(1+\sqrt{2})^{-2} = \pm(3-2\sqrt{2})$	$\pm(3, -2)$
-3	$\pm(1+\sqrt{2})^{-3} = \pm(7-5\sqrt{2})$	$\pm(7, -5)$

As we saw in the introduction, we can describe all solutions recursively (since we can get to any unit by repeatedly multiplying by $(1+\sqrt{2})$ or it's inverse). The integer solutions to $x^2 - 2y^2 = \pm 1$ are given by $(x, y) = \pm(a, \pm b)$, where $(a, b) = (1, 0)$ or is a term in the recursion $(a_{n+1}, b_{n+1}) \mapsto (a_n + 2b_n, a_n + b_n)$ with starting value $(a_0, b_0) = (1, 1)$.

For the general Pell equation $x^2 - dy^2 = \pm 1$ with $d > 1$ non-square it is always possible to give such a recursion. Simply find the fundamental unit $\alpha + \beta\sqrt{d}$ and consider the recursion $(a_{n+1}, b_{n+1}) \mapsto (\alpha a_n + d\beta b_n, \beta a_n + \alpha b_n)$ with starting value $(a_0, b_0) = (\alpha, \beta)$. (See Exercises).

6 Fermat's Last Theorem (1 lecture)

To finish the course we're going to tackle one final challenge... Fermat's Last Theorem!

Theorem 6.1. *The only integer solutions to the equation $x^n + y^n = z^n$ satisfy $xyz = 0$.*

As mentioned at the start of the course, this is famously a very tough theorem to prove. It took around 350 years to get the full solution, and it involves many objects beyond this course (e.g. elliptic curves, modular forms, Galois representations). We're just going to prove the special case $n = 3$, which thankfully will be possible by using previously covered material!

The idea is to define the cube root of unity $\zeta = \frac{-1+\sqrt{-3}}{2}$ and factorise the equation as follows:

$$(x + y)(x + \zeta y)(x + \zeta^2 y) = z^3.$$

The ring $\mathbb{Z}[\zeta] = \mathbb{Z}\left[\frac{-1+\sqrt{-3}}{2}\right]$ is then the ring of integers of the number field $\mathbb{Q}(\sqrt{-3})$ and we have already seen that it is a UFD (see Exercises).

First we'll need to prove a few simple facts.

Lemma 6.2. *Let $\pi = \sqrt{-3} = 2\zeta + 1 \in \mathbb{Z}[\zeta]$.*

1. *π is prime.*
2. *The unit group of the above ring is $\mathbb{Z}[\zeta]^\times = \{\pm 1, \pm\zeta, \pm\zeta^2\}$.*
3. *$\pi = u(1 - \zeta)$ and $\pi = v(1 - \zeta^2)$ for units $u, v \in \mathbb{Z}[\zeta]^\times$.*
4. *Every $\alpha \in \mathbb{Z}[\zeta]$ satisfies $\alpha \equiv 0, \pm 1 \bmod \pi$.*
5. *If $\alpha \in \mathbb{Z}[\zeta]$ is such that $\pi \nmid \alpha$ then $\alpha^3 \equiv 1 \bmod 9$.*

Proof.

1. This follows since $N(\pi) = \sqrt{-3} \cdot (-\sqrt{-3}) = 3$ is a rational prime (and $\mathbb{Z}[\zeta]$ is a UFD so that all irreducibles are prime).
2. We have already seen this fact in Proposition 5.5.
3. This follows from the previous part and the fact that $\zeta(1-\zeta) = -\zeta^2(1-\zeta^2) = \zeta - \zeta^2 = \sqrt{-3}$.
4. If $\alpha = \gamma + \delta\zeta$ for some $\gamma, \delta \in \mathbb{Z}$ then $\alpha \equiv \gamma + \delta \bmod \pi$ (since $\pi|(1 - \zeta)$ implies $\zeta \equiv 1 \bmod \pi$). We then note that $3 = \pi\overline{\pi} \equiv 0 \bmod \pi$, so that $\gamma + \delta \equiv 0, \pm 1 \bmod \pi$.
5. By the previous part and the fact that $\pi \nmid \alpha$ we must have that $\alpha \equiv \pm 1 \bmod \pi$. Letting $\alpha = \pm 1 + \pi\beta$ for some $\beta \in \mathbb{Z}[\zeta]$ we now note that:

   $$\alpha^3 = (\pm 1 + \pi\beta)^3 = \pm 1 + 3\pi\beta \pm 3\pi^2\beta^2 + \pi^3\beta^3 = \pm 1 \mp 9\beta^2 + 3\pi(\beta - \beta^3).$$

   However, we have that $\beta \equiv 0, \pm 1 \bmod \pi$ (again by the previous part) and so $\beta^3 \equiv 0^3, (\pm 1)^3 \equiv 0, \pm 1 \equiv \beta \bmod \pi$. It follows that the RHS is congruent to $\pm 1 \bmod 9$.

□

Proof. (of Theorem 6.1 for $n = 3$) Suppose that $x, y, z \in \mathbb{Z}[\zeta]$ are non-zero and satisfy $x^3 + y^3 = z^3$. We may assume that x, y, z are pairwise coprime, since if a prime $\lambda \in \mathbb{Z}[\zeta]$ divides any pair of x, y, z then it must clearly divide all three. We would then be able to cancel λ^3 from both sides of the equation (and continue in this fashion until no common prime factors remain).

Suppose that $\pi \nmid xyz$. Then part five of the Lemma shows that $x^3, y^3, z^3 \equiv \pm 1 \bmod 9$. But $z^3 = x^3 + y^3 \equiv 0, \pm 2 \bmod 9$ gives us a contradiction. It follows that $\pi | xyz$. Since x, y, z are pairwise coprime, we must have that π divides exactly one of x, y, z (and WLOG we can assume that $\pi | z$, since if $x^3 + y^3 = z^3$ then $x^3 + (-z)^3 = (-y)^3$ and $y^3 + (-z)^3 = (-x)^3$).

Claim: There are no solutions to the equation $x^3 + y^3 = uz^3$ with $x, y, z \in \mathbb{Z}[\zeta]$ non-zero, pairwise coprime and $\pi | z$ to order n (i.e. n is such that $\pi^n | z$ but $\pi^{n+1} \nmid z$).

We can prove this claim by induction on n.

Base case: If $n = 1$ then we have already seen that $x^3 + y^3 \equiv 0, \pm 2 \bmod 9$. This implies that either $\pi^4 = 9 | z^3$ or $\pi \nmid z^3$. By unique factorization this would contradict that fact that π^3 exactly divides z^3.

Inductive Hypothesis: Suppose that the claim is true for $n = k \geq 1$ and suppose that we have a solution to $x^3 + y^3 = uz^3$ with $x, y, z \in \mathbb{Z}[\zeta]$ non-zero, pairwise coprime and $\pi | z$ to order $k + 1$. Consider factoring the equation over $\mathbb{Z}[\zeta]$:

$$(x + y)(x + \zeta y)(x + \zeta^2 y) = uz^3.$$

Now since $\pi | z$ and π is prime we must have that $\pi | (x + \zeta^i y)$ for some $i \in \{0, 1, 2\}$. This then implies that $\pi | (x + \zeta^j y)$ for each $j \in \{0, 1, 2\}$ since:

$$x + \zeta^j y = (x + \zeta^i y) + (\zeta^j y - \zeta^i y) = (x + \zeta^i y) + \zeta^j (1 - \zeta^{j-i}) y$$

and by part three of the lemma we have that $\pi | (1 - \zeta^{j-i})$.

So π actually divides all brackets on the LHS and we can write:

$$\left(\frac{x + y}{\pi}\right)\left(\frac{x + \zeta y}{\pi}\right)\left(\frac{x + \zeta^2 y}{\pi}\right) = u\left(\frac{z}{\pi}\right)^3.$$

Now, note that since $k + 1 \geq 2$, the RHS is still divisible by π^{3k} and so at least one bracket on the LHS is divisible by π still.

We claim that the three brackets on the LHS are coprime, so that exactly one of them is divisible by π^{3k}. To prove the claim, note that if $\delta \in \mathbb{Z}[\zeta]$ is an irreducible dividing $\frac{x+\zeta^i y}{\pi}$ and $\frac{x+\zeta^j y}{\pi}$ (for $i \neq j$) then δ divides

$$\left(\frac{x + \zeta^i y}{\pi}\right) - \left(\frac{x + \zeta^j y}{\pi}\right) = \frac{\zeta^i (1 - \zeta^{j-i}) y}{\pi} = \frac{\zeta^i u \pi y}{\pi} = \zeta^i u y$$

for some unit u. Similarly δ divides

$$\left(\frac{x + \zeta^i y}{\pi}\right) - \zeta^{i-j}\left(\frac{x + \zeta^j y}{\pi}\right) = \frac{(1 - \zeta^{i-j}) x}{\pi} = \frac{v \pi x}{\pi} = vx$$

for some unit v. We would then have that $\delta|x$ and $\delta|y$, contradicting the pairwise coprime assumption on x, y, z.

We can replace y with ζy or $\zeta^2 y$ without changing the original equation $x^3+y^3=z^3$ (since $\zeta^3=1$). This allows us assume that it is the first bracket that is divisible by π^{3k}.

It follows that:

$$\left(\frac{x+y}{\pi^{3k+1}}\right)\left(\frac{x+\zeta y}{\pi}\right)\left(\frac{x+\zeta^2 y}{\pi}\right) = u\left(\frac{z}{\pi^{k+1}}\right)^3,$$

with each of the brackets on the LHS being coprime in $\mathbb{Z}[\zeta]$ (and the RHS lying in $\mathbb{Z}[\zeta]$ too). By unique factorization, there must exist units $u_1, u_2, u_3 \in \mathbb{Z}[\zeta]^\times$ and coprime $z_1, z_2, z_3 \in \mathbb{Z}[\zeta]$ such that:

$$\frac{x+y}{\pi^{3k+1}} = u_1 z_1^3 \qquad \frac{x+\zeta y}{\pi} = u_2 z_2^3 \qquad \frac{x+\zeta^2 y}{\pi} = u_3 z_3^3.$$

We must have that $\pi \nmid z_1 z_2 z_3$, otherwise π would divide z to higher order than $k+1$.

A quick computation gives that $(x+y)+\zeta(x+\zeta y)+\zeta^2(x+\zeta^2 y)=0$, so that:

$$u_1\pi^{3k+1}z_1^3 + \zeta u_2 \pi z_2^3 + \zeta^2 u_3 \pi z_3^3 = 0.$$

Rearranging and cancelling π gives:

$$z_2^3 + \left(\frac{\zeta u_3}{u_2}\right) z_3^3 = \left(-\frac{u_1}{\zeta u_2}\right)\pi^{3k} z_1^3.$$

Since the quantities in brackets are both units, we can rewrite this as:

$$z_2^3 + u_1' z_3^3 = u_2' \pi^{3k} z_1^3,$$

for units $u_1', u_2' \in \mathbb{Z}[\zeta]^\times$. Finally, we note that π^3 divides the RHS, so that (with another use of the Lemma):

$$z_2^3 + u_1' z_3^2 \equiv \pm 1 \pm u_1' \equiv 0 \bmod \pi^3.$$

One can check that out of the six possibilities for the unit u_1', only $u_1' = \pm 1$ can satisfy this congruence. Thus $u_1' = \pm 1$, and it follows that:

$$z_2^3 + (\pm z_3)^3 = u_2'(\pi^k z_1)^3.$$

Since $z_1, \pm z_3$ and $\pi^k z_3$ are non-zero and coprime, and π^k divides $\pi^k z_3$ to order k, the inductive hypothesis tells us that this equation has no solutions, proving the claim...hence proving Fermat's Last Theorem in the case $n=3$! $\square$