MTH 561-562: Abstract Algebra 2021-2022
Course Notes Drew Armstrong

Contents

1 Complex Numbers

1.1 Cardano's Formula

One could say that algebra began with the study of quadratic equations. Given any numbers a, b, c we want to find all numbers x such that

$$ax^2 + bx + c = 0.$$

If $a = 0$ then there is nothing interesting to do, so let us assume that $a \neq 0$. First we divide both sides by a to obtain

$$\begin{aligned} x^2 + \frac{b}{a}x + \frac{c}{a} &= 0 \\ x^2 + \frac{b}{a}x &= -\frac{c}{a}. \end{aligned}$$

Now there is a famous trick called "completing the square." We add the the quantity $(b/2a)^2$ to both sides and observe that the left side factors:

$$\begin{aligned} x^2 + \frac{b}{a}x &= -\frac{c}{a} \\ x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 &= -\frac{c}{a} + \left(\frac{b}{2a}\right)^2 \end{aligned}$$

$$\begin{aligned}\left(x+\frac{b}{2a}\right)\left(x+\frac{b}{2a}\right)&=-\frac{c}{a}+\frac{b^2}{4a^2}\\ \left(x+\frac{b}{2a}\right)^2&=\frac{b^2-4ac}{4a^2}.\end{aligned}$$

Finally, we can take the square root of the left side and solve for x:

$$\begin{aligned}\left(x+\frac{b}{2a}\right)^2&=\frac{b^2-4ac}{4a^2}\\ x+\frac{b}{2a}&=\frac{\pm\sqrt{b^2-4ac}}{2a}\\ x&=-\frac{b}{2a}+\frac{\pm\sqrt{b^2-4ac}}{2a}\\ &=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.\end{aligned}$$

I'm sure that you already knew already this. But let me point out a subtlety that you may not have thought about. If $b^2-4ac\neq 0$ then the square root symbol $\sqrt{b^2-4ac}$ can refer to two different numbers. When $b^2-4ac>0$ then we usually assume that $\sqrt{b^2-4ac}$ refers to the positive real square root. However, if b^2-4ac is negative or non-real then it is not so clear what the symbol $\sqrt{b^2-4ac}$ should refer to. For example, we often write $i=\sqrt{-1}$ to refer to "the" square root of -1, but the number -1 actually has two square roots and there is no good way to distinguish between them. So we should really say:

> Let i denote an arbitrary symbol satisfying $i^2=-1$. Then the equation $x^2=1$ has exactly two solutions: i and $-i$, which are the two square roots of -1.

Later we will prove that any nonzero number of the form $a+b\sqrt{-1}$ has exactly two square roots, which are negatives of each other. With this in mind, here is a modern statement of the quadratic formula.

Modern Version of the Quadratic Formula

Let a,b,c be any numbers and let $\Delta=b^2-4ac$ denote the "discriminant" of the equation $ax^2+bx+c=0$. By completing the square we showed above that any solution has the form $x=(-b+\delta)/2a$, where δ is some number satisfying $\delta^2=\Delta$. Conversely, one can check that any x of this form is a solution. Thus we have one solution x for each square root of Δ. If $\Delta=0$ then $\delta=0$ is the only square root. Otherwise, if δ is an arbitrary square root of Δ then Δ has exactly two square roots: δ and $-\delta$. And the quadratic equation has exactly two solutions:

$$x=\frac{-b+\delta}{2a}\qquad\text{or}\qquad x=\frac{-b+(-\delta)}{2a}.$$

The quadratic formula was known to ancient civilizations. The next progress only came in the 1500s, when several Italian mathematicians discovered algorithms for the solution of cubic and quartic equations. These formulas were first published by Gerolamo Cardano in the *Ars Magna* (1545). For now I will just state the formula without proof.

Cardano's Formula (1545)

Let a, b, c, d be any numbers with $a \neq 0$ and consider the cubic equation

$$ax^3 + bx^2 + cx + d = 0.$$

To solve this we first divide both sides by a and then we substitute $x = y - b/(3a)$ to obtain the so-called "depressed form" of the equation:

$$y^3 + 3py + 2q = 0,$$

where[1]

$$p = \frac{3ac - b^2}{9a^2} \quad \text{and} \quad q = \frac{27a^2d - 9abc + 2b^3}{54a^3}.$$

Then Cardano's formula says that

$$y = \sqrt[3]{-q + \sqrt{q^2 + p^3}} + \sqrt[3]{-q - \sqrt{q^2 + p^3}}.$$

We could expand all of this to write a formula for x in terms of a, b, c, d, but that would look horrible.

This formula is quite difficult to interpret. In Cardano's time only real numbers were accepted, which led to two immediate problems:

(1) Sometimes there is an obvious solution but the formula does not see it.

(2) Sometimes there are 3 solutions but the formula only sees one of them.

These problems were eventually solved by the introduction of "complex numbers" of the form $a + b\sqrt{-1}$. The first hint of this idea was observed by Bombelli.

Bombelli's Example (1572)

[1] These complicated expressions are one of the reasons why the cubic equation is not studied in high school.

Consider the following cubic equation:

$$x^3 - 15x - 4 = 0.$$

One can easily check that $x = 4$. On the other hand, by applying Cardano's formula with $p = -5$ and $q = -2$ we obtain

$$\begin{aligned} x &= \sqrt[3]{-(-2) + \sqrt{(-2)^2 + (-5)^3}} + \sqrt[3]{-(-2) - \sqrt{(-2)^2 + (-5)^3}} \\ &= \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}. \end{aligned}$$

Cardano would say here that the formula gives no solution because square roots of negative numbers do not exist. Bombelli's idea was to just **pretend** that the expression $\sqrt{-1}$ is a number with the property $(\sqrt{-1})^2 = -1$ and to perform computations as usual. After some trial and error he observed that[2]

$$\begin{aligned} (2 + \sqrt{-1})^2 &= (2 + \sqrt{-1})(2 + \sqrt{-1}) \\ &= (2 + \sqrt{-1})(4 + 4\sqrt{-1} + (\sqrt{-1})^2) \\ &= (2 + \sqrt{-1})(4 + 4\sqrt{-1} - 1) \\ &= (2 + \sqrt{-1})(3 + 4\sqrt{-1}) \\ &= 6 + 11\sqrt{-1} + 4(\sqrt{-1})^2) \\ &= 6 + 11\sqrt{-1} - 4 \\ &= 2 + 11\sqrt{-1} \\ &= 2 + \sqrt{121}\sqrt{-1} \\ &= 2 + \sqrt{-121}. \end{aligned}$$

And a similar computation shows that $(2 - \sqrt{-1})^3 = 2 - \sqrt{-121}$. Therefore Bombelli concluded that Cardano's formula really does give the correct answer:

$$\begin{aligned} x &= \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}} \\ &= (2 + \cancel{\sqrt{-1}}) + (2 - \cancel{\sqrt{-1}}) \\ &= 4. \end{aligned}$$

In other words: The "real" solution 4 is obtained from Cardano's formula as a sum of two "imaginary" numbers.

In the next section I will give the modern interpretation of these computations.

[2] In the last step we have used the "formula" $\sqrt{ab} = \sqrt{a}\sqrt{b}$, which of course is not really a formula because it depends on the specific choices of the square roots.

1.2 Complex Numbers as a Ring

Bombelli observed that some issues with Cardano's formula can be resolved by pretending that the "imaginary" square roots of negative numbers actually exist. These ideas were slow to catch on, and were regarded by some as useless speculation well into the 1700s. The modern formulation is essentially the same as Bombelli's, just stated with more confidence. Let i be an abstract symbol. Then a *complex number* is an abstract symbol of the form $a + bi$, where a and b are real numbers. The set of real numbers is denoted by $\mathbb{R}$ and the set of complex numbers is denoted by

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}.$$

Let me emphasize that "$a + bi$" is only an abstract expression; the plus sign does not at first have anything to do with addition of real numbers because the symbol bi is not a real number. In order to make sense of this we will **define** addition and multiplication of the symbols "$a + bi$" by the following formulas:[3]

$$\begin{aligned}(a + bi) + (c + di) &:= (a + c) + (b + d)i,\\ (a + bi)(c + di) &:= (ac - bd) + (ad + bc)i.\end{aligned}$$

Perhaps it is not surprising that these operations turn out to behave just like the addition and multiplication of real numbers. In abstract algebra we capture this behavior with the following definition.

Definition of Rings

A *ring* is a set R together with two special elements $0, 1 \in R$ (called zero and one) and two binary operations $+, \cdot : R \times R \to R$ (called addition and multiplication), which satisfy the following eight axioms:

(A1) $\forall a, b \in R,\ a + b = b + a$ (commutative addition)

(A2) $\forall a, b, c \in R,\ a + (b + c) = (a + b) + c$ (associative addition)

(A3) $\forall a \in R,\ a + 0 = a$ (additive identity)

(A4) $\forall a \in R,\ \exists b \in R,\ a + b = 0$ (additive inversion)

(M1) $\forall a, b \in R,\ ab = ba$ (commutative multiplication)

(M2) $\forall a, b, c \in R,\ a(bc) = (ab)c$ (associative multiplication)

(M3) $\forall a \in R,\ a1 = a$ (multiplicative identity)

(D) $\forall a, b, c \in R,\ a(b + c) = ab + ac$ (distribution)

If we delete axiom (M1) then we obtain a structure called a *non-commutative ring*. In this course all rings will be commutative unless otherwise stated.

We can also define subtraction in a ring. Given any element $a \in R$, axiom (A4) tells us that there exists at least one element $b \in R$ with the property $a + b = 0$. In fact, there is

exactly one such element. Indeed, if $a+b=0$ and $a+b'=0$ then by combining axioms (A1), (A2), (A3) we obtain

$$b = b + 0 = b + (a + b') = (b + a) + b' = 0 + b' = c.$$

Since this element is unique we will denote it by the symbol "$-a$", and for any two elements $a, b \in R$ we will define the symbol

$$\text{“}a - b\text{”} := a + (-b).$$

In other words, a ring is a "number system" in which any two numbers can be added, subtracted and multiplied, and in which all of the usuals laws of arithmetic hold. One can check that the set of complex numbers $\mathbb{C}$ forms a ring with the operations defined above, and with the special elements $0 := 0 + 0i$ and $1 := 1 + 0i$.[4] This is the ultimate justification for referring to the symbols "$a + bi$" as "numbers". Here are the four most commonly discussed rings:

name	symbol	casual description
integers	$\mathbb{Z}$	$\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
rational numbers	$\mathbb{Q}$	$\{a/b : a, b \in \mathbb{Z}, b \neq 0\}$
real numbers	$\mathbb{R}$	{limits of sequences of rational numbers}
complex numbers	$\mathbb{C}$	$\{a + b\sqrt{-1} : a, b \in \mathbb{R}\}$

We can think of these as a nested sequence of "subrings"

$$\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

by identifying each fraction of the form $a/1$ with the integer a and by identifying the complex number of the form $a + 0i$ with the real number a. But let me observe that the rings $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ have an important extra property that $\mathbb{Z}$ does not have.

Definition of Fields

Let $(\mathbb{F}, +, \cdot, 0, 1)$ be a ring. We say that $\mathbb{F}$ is a *field* if it satisfies one further axiom:

(M4) $\forall a \in \mathbb{F}\backslash\{0\}, \exists b \in \mathbb{F}, ab = 1.$

In words: For any nonzero element $a \in \mathbb{F}$ there exists at least one element $b \in \mathbb{F}$ with the property $ab = 1$. In fact, there is **exactly one** such element. Indeed, if $ab = 1$ and $ab' = 1$ then by combining axioms (M1), (M2), (M3) we obtain

$$b = b1 = b(ab') = (ba)b' = 1b' = b'.$$

Since this element is unique we can give it the special name "a^{-1}", or "$1/a$". Then for

[3] The symbol := means "is defined as". It was adopted by mathematicians from the Pascal programming language.

any two elements $a, b \in \mathbb{F}$ with $b \neq 0$ we will define the notation

$$\text{“}a/b\text{”} = ab^{-1}.$$

You are familiar with the fact that rational numbers $\mathbb{Q}$ and the real numbers $\mathbb{R}$ are fields. Let me quickly observe that the ring of integers $\mathbb{Z}$ is **not** a field. For example, suppose for contradiction that there exists an integer $b \in \mathbb{Z}$ satisfying $2b = 1$. The integer b must be positive, which implies that $b \geqslant 1$ because there are no integers between 0 and 1. But then multiplying both sides by 2 gives a contradiction:

$$\begin{aligned} b &\geqslant 1 \\ 2b &\geqslant 2 \\ 1 &\geqslant 2. \end{aligned}$$

1.3 Complex Numbers as a Vector Space

So $\mathbb{Z}$ is a ring that is not a field and $\mathbb{Q}, \mathbb{R}$ are fields. In this section we will show that $\mathbb{C}$ is also a field, which is surprisingly difficult. Before proving this in the next section we need to say more about the relationship between $\mathbb{R}$ and $\mathbb{C}$. Recall that we view each real number a as a complex number by setting $a = a + 0i$. With this convention, the abstract symbol "$a + bi$" acquires a direct algebraic meaning:

$$\text{“}a + bi\text{”} = (a + 0i) + (b + 0i)(0 + 0i).$$

Of course this was the point all along. In order to formalize the relationship between $\mathbb{R}$ and $\mathbb{C}$ I will present another of the key concepts from twentieth century abstract algebra.

Definition of Vector Spaces and Dimension

A *vector space* consists of a set V (of vectors), a field $\mathbb{F}$ (of scalars), an operation $+ : V \times V \to V$ (called vector addition), and an operation $\cdot : \mathbb{F} \times V \to V$ (called scalar multiplication), which satisfy the following eight axioms:

(V1) $\forall \mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative addition)

(V2) $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (associative addition)

(V3) $\exists \mathbf{0} \in V, \forall \mathbf{u} \in V$, $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (zero vector)

(V4) $\forall \mathbf{u} \in V$, $\exists \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{0}$ (additive inversion)

(V5) $\forall \mathbf{u} \in V$, $1\mathbf{u} = \mathbf{u}$ (unit scalar)

(V6) $\forall a, b \in \mathbb{F}$, $\mathbf{u} \in V$, $a(b\mathbf{u}) = (ab)\mathbf{u}$ (associative multiplication)

(V7) $\forall a, b \in \mathbb{F}$, $\mathbf{u} \in V$, $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ (distribution)

[4] The proof is extremely boring.

(V8) $\forall a \in \mathbb{F}$, $\mathbf{u}, \mathbf{v} \in V$, $a(\mathbf{u}+\mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ (distribution)

We can also define subtraction of vectors. Given any $\mathbf{v} \in V$, axiom (V4) tells us that there exists at least one element $\mathbf{u} \in V$ satisfying $\mathbf{u}+\mathbf{v} = \mathbf{0}$. In fact, there is exactly one such element. Indeed, if $\mathbf{v}+\mathbf{u} = \mathbf{0}$ and $\mathbf{v}+\mathbf{u}' = \mathbf{0}$ then axioms (V1), (V2), (V3) imply that

$$\mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u} + (\mathbf{v}+\mathbf{u}') = (\mathbf{u}+\mathbf{v}) + \mathbf{u}' = \mathbf{0} + \mathbf{u}' = \mathbf{u}'.$$

We will call this unique element "$-\mathbf{v}$" and use it to define subtraction:

$$\text{“}\mathbf{u} - \mathbf{v}\text{”} := \mathbf{u} + (-\mathbf{v}).$$

We say that a vector space V over $\mathbb{F}$ is *n-dimensional* if there exists a set of n vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n \in V$ with the property that every vector $\mathbf{v} \in V$ has a **unique** expression of the form

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n \quad \text{with} \quad a_1, \ldots, a_n \in \mathbb{F}.$$

In this case we say that $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a *basis* for V over $\mathbb{F}$.

Remark: The definition of vector space does not include a way to multiply two vectors. Later we will discuss the definition of "inner product space", which includes a way to multiply two vectors to obtain a scalar. (Example: The dot product.) It is almost never possible to multiply two vectors to obtain another vector but we will see that the complex numbers are a special case.

The abstract definition of vector space is inspired by the following familiar example.

Prototype of a Vector Space: Cartesian Coordinates

Let $\mathbb{R}^n$ denote the set of ordered n-tuples of real numbers:

$$\mathbb{R}^n := \{\mathbf{x} = (x_1, \ldots, x_n) : x_i \in \mathbb{R} \text{ for all } i\}.$$

It is easy (and boring) to check that the following operations make the set $\mathbb{R}^n$ in to a vector space over the field of scalars $\mathbb{R}$:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_2 + y_2)$$
$$a \cdot (x_1, \ldots, x_n) := (ax_1, \ldots, ax_n).$$

As you know, we can view the vector $\mathbf{x}$ as a point in n-dimensional space. We can also view it as a directed line segment whose head is at the point $\mathbf{x}$ and whose tail is at the "origin" $\mathbf{0} = (0, \ldots, 0)$. Then the addition of vectors can be viewed as the familiar "head-to-tail" addition of directed line segments. This idea goes back at least to Isaac Newton, who used it to describe forces acting on rigid bodies.

It is not surprising that the vector space $\mathbb{R}^n$ is n-dimensional. To prove this, we can observe that the set of n vectors

$$\begin{aligned}\mathbf{e}_1 &= (1,0,0,\ldots,0,0)\\ \mathbf{e}_2 &= (0,1,0,\ldots,0,0)\\ &\vdots\\ \mathbf{e}_n &= (0,0,0,\ldots,0,1)\end{aligned}$$

is a basis of $\mathbb{R}^n$, called the *standard basis*. Indeed, for vector $\mathbf{x} = (x_1,\ldots,x_n)$ we have

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n,$$

and by definition these "coordinates" $x_1,\ldots,x_n$ are unique.

So what? The point of this section is that the complex numbers $\mathbb{C}$ naturally form a two-dimensional vector space over the field of real numbers $\mathbb{R}$.

$\mathbb{C}$ is a Two-Dimensional Vector Space over $\mathbb{R}$

We can view $\mathbb{C}$ as a vector space over $\mathbb{R}$ where $0 = 0 + 0i$ is the "zero vector" and where "vector addition" and "scalar multiplication" are given by the usual addition and multiplication of numbers:

$$\begin{aligned}(a+bi)+(c+di) &= (a+c)+(b+d)i\\ a(b+ci) &= (ab)+(ac)i.\end{aligned}$$

It is easy and boring to check that the eight vector space axioms hold in this situation. To see that this vector space is two-dimensional I claim that the set of two elements $1, i \in \mathbb{C}$ is a basis. Indeed, any complex number can be expressed in the form $a1 + bi$ for some $a, b \in \mathbb{R}$, and we only need to check that this representation is **unique**. For this purpose, suppose that we have $a + bi = c + di$ with $a, b, c, d \in \mathbb{R}$. Our goal is to show that $a = c$ and $b = d$. So let us suppose for contradiction that $b \neq d$. Then we have

$$\begin{aligned}a+bi &= c+di\\ 0+(b-d)i &= (c-a)+0i\\ 0+1i &= \left(\frac{c-a}{b-d}\right)+0i,\end{aligned}$$

which implies that i is a real number. But i is **not real** because any real number $a \in \mathbb{R}$

satisfies $a^2 \geqslant 0$, but $i^2 = -1 < 0$. This contradiction implies that $b = d$, hence also

$$\begin{aligned} a + bi &= c + di \\ a + \cancel{bi} &= c + \cancel{bi} \\ a &= c. \end{aligned}$$

In summary, we have

$$a + bi = c + di \quad \Longleftrightarrow \quad a = c \text{ and } b = d.$$

You might have noticed here that the vector space $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ is basically just the vector space $\mathbb{R}^2 = \{(a, b) : a, b \in \mathbb{R}\}$ in disguise. In technical jargon we will say that $\mathbb{C}$ and $\mathbb{R}^2$ are *isomorphic as vector spaces*. This just means that we have a one-to-one correspondence that preserves all of the vector space operations. In this case the one-to-one correspondence is particularly obvious:

$$\begin{array}{ccc} \mathbb{C} & \leftrightarrow & \mathbb{R}^2 \\ a + bi & \leftrightarrow & (a, b). \end{array}$$

The word "isomorphism" literally means "same structure". We use it in mathematics when two different mathematical structures are "essentially the same"; that is, when there is a one-to-one correspondence between their elements that preserves all of the relevant structure/operations.

1.4 Complex Numbers as a Field

By using scalar multiplication we can "divide" any complex number $a + bi \in \mathbb{C}$ by any nonzero real number $c \in \mathbb{R}$:

$$\frac{a + bi}{c} = \left(\frac{1}{c}\right)(a + bi) = \left(\frac{a}{c}\right) + \left(\frac{b}{c}\right) i.$$

The question is whether we can also divide by complex numbers:

$$\frac{a + bi}{c + di} = (\text{some real number?}) + (\text{some real number?})\, i.$$

This can be quite difficult unless you know a clever trick called "rationalizing the denominator". The idea is to multiply both the numerator and denominator of the hypothetical fraction "$(a + bi)/(c + di)$" by the "complex conjugate" of the denominator:

$$\begin{aligned} \frac{a + bi}{c + di} &= \frac{a + bi}{c + di} \cdot \frac{c - di}{c - di} \\ &= \frac{(a + bi)(c - di)}{(c + di)(c - di)} \\ &= \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} \\ &= \left(\frac{ac + bd}{c^2 + d^2}\right) + \left(\frac{bc - ad}{c^2 + d^2}\right) i. \end{aligned}$$

For this to work we require that $c^2 + d^2 \neq 0$, which will be true if $c + di \neq 0 + 0i$. Indeed, if $c + di \neq 0 + 0i$ then we must have $c \neq 0$ or $d \neq 0$, in which case $c^2 + d^2 > 0$. Thus we can divide by any nonzero complex number.

This trick of rationalizing the denominator is so useful that we turn it into a general concept.

Complex Conjugation and Absolute Value

For any complex number $\alpha = a + bi \in \mathbb{C}$ we define its *complex conjugate* $\alpha^* \in \mathbb{C}$ as follows:

$$(a + bi)^* := a - bi.$$

Then we define the *absolute value* $|\alpha| \in \mathbb{R}$ as the non-negative real square root of $a^2 + b^2 \in \mathbb{R}$ and we observe that

$$\alpha\alpha^* = (a + bi)(a - bi) = (a^2 + b^2) + 0i = a^2 + b^2 = |\alpha|^2.$$

For all complex numbers $\alpha, \beta \in \mathbb{C}$, I claim that the following properties hold:

- $\alpha = 0$ if and only if $|\alpha| = 0$.
- $\alpha = \alpha^*$ if and only if $\alpha \in \mathbb{R}$,
- $(\alpha + \beta)^* = \alpha^* + \beta^*$,
- $(\alpha\beta)^* = \alpha^*\beta^*$,
- $|\alpha\beta| = |\alpha||\beta|$.

You will prove all of these assertions on the homework. The final property (the multiplicativity of the absolute value) is probably the deepest fact about the complex numbers. It was first glimpsed by Diophantus of Alexandria (3rd century), who used the "two-square identity"

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$$

to study "Pythagorean triples of whole numbers", such as $3^2 + 4^2 = 5^2$ and $5^2 + 12^2 = 13^2$.

We can use the ideas of conjugation and absolute value to give a slicker proof that $\mathbb{C}$ is a field.

Multiplicative Inverses in $\mathbb{C}$

For any nonzero complex number $\alpha \in \mathbb{C}$ we have $|\alpha| \neq 0$. It follows that

$$\begin{aligned} \alpha\alpha^* &= |\alpha|^2 \\ \alpha(\alpha^*/|\alpha|^2) &= 1, \end{aligned}$$

so the multiplicative inverse of α has the explicit formula

$$\alpha^{-1} = \frac{\alpha^*}{|\alpha|^2}.$$

On the homework you will use the same ideas to show that the following set is a field:

$$\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}.$$

Later we will incorporate all of this into a general theory of "quadratic field extensions".

1.5 Complex Numbers as Linear Functions

The complex numbers are a central object in mathematics, which means that they can be viewed from many different angles. So far we have viewed $\mathbb{C}$ as a ring (specifically, a field) and as a two-dimensional vector space over $\mathbb{R}$. Recall that we have a bijection

$$\begin{array}{ccc} \mathbb{C} & \leftrightarrow & \mathbb{R}^2 \\ a + bi & \leftrightarrow & (a, b) \end{array}$$

that preserves the operations of vector addition and scalar multiplication. To be specific, the addition of vectors corresponds to addition of complex numbers and the scalar multiplication of vectors by real numbers corresponds to the usual multiplication of complex numbers by real numbers.

However, there is also a natural way to multiply any two complex numbers. What does this correspond to in $\mathbb{R}^2$? In general there is no sensible way to multiply two vectors in a vector space to obtain another vector, so this case must be very special. The key to understanding it is to express complex numbers in "polar form".

Polar Form of Complex Numbers

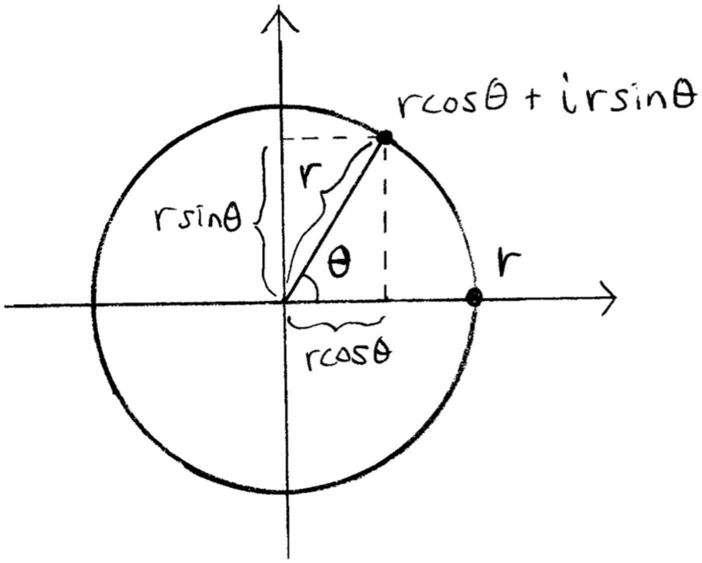
Based on the isomorphism $\mathbb{C} \cong \mathbb{R}^2$ we can view the complex number $a + bi$ as the point (a, b) in the Cartesian plane. But we can also express points of $\mathbb{R}^2$ in polar coordinates. That is, for any pair of real numbers (a, b), not both zero, there exist a unique pair of real numbers r and θ satisfying

$$a = r\cos\theta, \quad b = r\sin\theta, \quad r > 0 \quad \text{and} \quad \theta \in [0, 2\pi).$$

In other words, for any nonzero complex number $a + bi$, there exist unique real numbers $r > 0$ and $\theta \in [0, 2\pi)$ such that

$$a + bi = (r\cos\theta) + (r\sin\theta)i = r(\cos\theta + i\sin\theta).$$

In geometric terms, $r = |\alpha| = +\sqrt{a^2 + b^2}$ is the length of the vector (a, b) and and we view θ as the angle of the vector (a, b), measured counterclockwise from the "real axis":



Using these ideas, we have the following geometric interpretation of complex multiplication.

Geometric Interpretation of Complex Multiplication

Let $\alpha, \beta \in \mathbb{C}$ be nonzero complex numbers, thought of as vectors in the Cartesian plane $\mathbb{R}^2$. Suppose that α, β have lengths $r, s > 0$ and angles $\theta, \lambda \in [0, 2\pi)$, so that

$$\begin{aligned}\alpha &= r(\cos\theta + i\sin\theta),\\ \beta &= s(\cos\lambda + i\sin\lambda).\end{aligned}$$

Then I claim that the complex number $\alpha\beta$ has length rs and angle $\theta + \lambda$ (up to a suitable multiple of 2π). In other words:

the lengths multiply and the angles add.

Here is a quick and dirty proof, using the "angle sum identities" from trigonometry:

$$\begin{aligned}\alpha\beta &= r(\cos\theta + i\sin\theta)\cdot s(\cos\lambda + i\sin\lambda)\\ &= (rs)(\cos\theta + i\sin\theta)(\cos\lambda + i\sin\lambda)\\ &= (rs)[(\cos\theta\cos\lambda - \sin\theta\sin\lambda) + i(\cos\theta\sin\lambda + \sin\theta\cos\lambda)]\\ &= (rs)[\cos(\theta+\lambda) + i\sin(\theta+\lambda)].\end{aligned}$$

But this proof is not good because it seems like a coincidence. The true meaning of the theorem is revealed when we view complex numbers as “linear functions”.

Linear Functions and Matrices

Consider the vector space $\mathbb{R}^n$ over the field $\mathbb{R}$. We say that a function $L : \mathbb{R}^n \to \mathbb{R}^n$ is $\mathbb{R}$-*linear* if it preserves vector addition and scalar multiplication by $\mathbb{R}$. That is, for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and $a \in \mathbb{R}$ we must have

- $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$ (preserves addition)
- $L(a\mathbf{u}) = aL(\mathbf{u})$ (preserves scalar multiplication)

Equivalently, we can combine these by saying that L preserves “linear combinations”:[5]

$$L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v}).$$

I claim that there is a one-to-one correspondence between linear functions from $\mathbb{R}^n \to \mathbb{R}^n$ and $n \times n$ matrices with entries from $\mathbb{R}$:

$$\left\{ \begin{array}{c} \text{linear functions} \\ \text{from } \mathbb{R}^n \text{ to } \mathbb{R}^n \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} n \times n \text{ matrices with} \\ \text{entries from } \mathbb{R} \end{array} \right\}.$$

In order to find such a correspondence, we will identify each vector $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ with the corresponding $n \times 1$ column vector:

$$[\mathbf{u}] = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Then the standard basis vectors are written as follows:

$$[\mathbf{e}_1] = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, [\mathbf{e}_2] = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, [\mathbf{e}_n] = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

And the column $[\mathbf{u}]$ has a **unique** expression as a linear combination of basis vectors:

$$[\mathbf{u}] = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} u_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ u_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ u_n \end{pmatrix} = u_1[\mathbf{e}_1] + u_2[\mathbf{e}_2] + \cdots + u_n[\mathbf{e}_n].$$

Now for any linear function $L : \mathbb{R}^n \to \mathbb{R}^n$ we define the $n \times n$ matrix $[L] \in \mathbb{R}^{n\times n}$ whose ith column is the vector $L(\mathbf{e}_i) \in \mathbb{R}^n$:

$$[L] := \begin{pmatrix} | & | & & | \\ [L(\mathbf{e}_1)] & [L(\mathbf{e}_2)] & \cdots & [L(\mathbf{e}_n)] \\ | & | & & | \end{pmatrix}.$$

I claim that the assignment $L \mapsto [L]$ is a one-to-one correspondence. To prove this we will first show that the assignment is one-to-one. So let $L, M \in \mathbb{R}^n \to \mathbb{R}^n$ be two linear functions with the same matrix: $[L] = [M]$. By definition this means that $L(\mathbf{e}_i) = M(\mathbf{e}_i)$ for all i, because the two matrices have the same column vectors. For all vectors $\mathbf{u} \in \mathbb{R}^n$ it follows from the linearity of L and M that

$$\begin{aligned} L(\mathbf{u}) &= L(u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_n\mathbf{e}_n) \\ &= u_1L(\mathbf{e}_1) + u_2L(\mathbf{e}_2) + \cdots + u_nL(\mathbf{e}_n) \\ &= u_1M(\mathbf{e}_1) + u_2M(\mathbf{e}_2) + \cdots + u_nM(\mathbf{e}_n) \\ &= M(u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_n\mathbf{e}_n) \\ &= M(\mathbf{u}), \end{aligned}$$

hence $L = M$ as functions. Finally, we will show that the assignment is onto. So let Φ be any $n \times n$ matrix. We need to show that there exists some (necessarily unique) linear function $L_\Phi : \mathbb{R}^n \to \mathbb{R}^n$ with the property $\Phi = [L_\Phi]$. If Φ_i is the ith column vector of the matrix Φ then I claim that the following definition works:

$$L_\Phi(\mathbf{u}) := u_1\Phi_1 + u_2\Phi_2 + \cdots + u_n\Phi_n.$$

Indeed, it is easy to check that this function is linear. And the matrices $[L_\Phi]$ and Φ have the same column vectors because

$$L(\mathbf{e}_i) = 0\Phi_1 + \cdots + 0\Phi_{i-1} + 1\Phi_i + 0\Phi_{i+1} + \cdots + 0\Phi_n = \Phi_i.$$

In summary, the following pair of assignments are inverses:

$$\begin{array}{rcl} \{\text{linear functions } \mathbb{R}^n \to \mathbb{R}^n\} & \longleftrightarrow & \{n \times n \text{ matrices}\} \\ L & \mapsto & [L] \\ L_\Phi & \leftarrow\!\shortmid & \Phi. \end{array}$$

More generally, this entire line of reasoning gives a bijection between linear functions from $\mathbb{R}^n \to \mathbb{R}^m$ and $m \times n$ matrices, i.e., matrices with m rows and n columns.

That was quite abstract, so let's examine a few examples.

[5] Geometrically, a linear function must send the origin to itself and send parallelograms to parallelograms.

- **The Identity Matrix.** The identity function $I : \mathbb{R}^n \to \mathbb{R}^n$ defined by $I(\mathbf{u}) = \mathbf{u}$ is obviously linear. The corresponding matrix is called the *identity matrix*:

$$\begin{aligned}[I] &= \begin{pmatrix} | & | & & | \\ [I(\mathbf{e}_1)] & [I(\mathbf{e}_2)] & \cdots & [I(\mathbf{e}_n)] \\ | & | & & | \end{pmatrix} \\ &= \begin{pmatrix} | & | & & | \\ [\mathbf{e}_1] & [\mathbf{e}_2] & \cdots & [\mathbf{e}_n] \\ | & | & & | \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.\end{aligned}$$

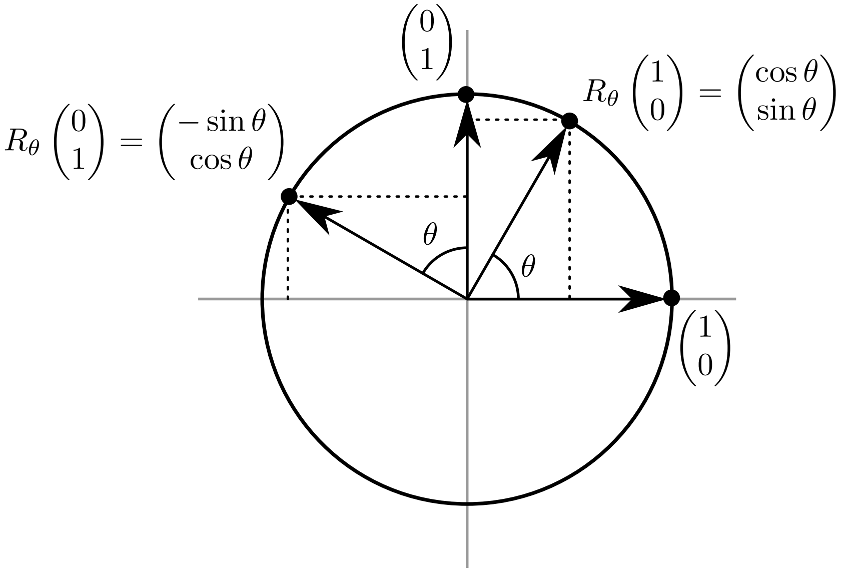
- **Scalar Matrices.** For any scalar $r \in \mathbb{R}$ the function $S_r : \mathbb{R}^n \to \mathbb{R}^n$ defined by $S_r(\mathbf{u}) = r\mathbf{u}$ is linear. The corresponding matrix is

$$\begin{aligned}[S_r] &= \begin{pmatrix} | & | & & | \\ [S_r(\mathbf{e}_1)] & [S_r(\mathbf{e}_2)] & \cdots & [S_r(\mathbf{e}_n)] \\ | & | & & | \end{pmatrix} \\ &= \begin{pmatrix} | & | & & | \\ [r\mathbf{e}_1] & [r\mathbf{e}_2] & \cdots & [r\mathbf{e}_n] \\ | & | & & | \end{pmatrix} \\ &= \begin{pmatrix} r & 0 & \cdots & 0 \\ 0 & r & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r \end{pmatrix}.\end{aligned}$$

  Note that this includes the identity matrix as a specific example when $r = 1$.

- **Rotation Matrices.** Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ denote the function that rotates every vector by angle θ, counterclockwise around the origin. It is easy to see that this function preserves vector addition and scalar multiplication, hence it is linear.

  What is the corresponding 2×2 matrix? The following diagram illustrates how the function R_θ acts on the standard basis vectors $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$:


$\begin{pmatrix}0\\1\end{pmatrix}$
$R_\theta\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}$
$R_\theta\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-\sin\theta\\\cos\theta\end{pmatrix}$
θ
θ
$\begin{pmatrix}1\\0\end{pmatrix}$

It follows that the matrix of the rotation function R_θ is

$$[R_\theta] = \begin{pmatrix} | & | \\ [R_\theta(\mathbf{e}_1)] & [R_\theta(\mathbf{e}_2)] \\ | & | \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Note that "rotation by zero" is the identity function, hence $[R_0]$ is the identity matrix.

Whenever there is a one-to-one correspondence between two different kinds of structures, for example between linear functions and matrices, it is important to ask how natural operations behave under this correspondence. I assume that you are familiar with the definition of matrix multiplication, but you may not be aware of the reason behind it.

Matrix Multiplication = Composition of Linear Functions

Recall from previous theorem that any two $n \times n$ matrices can be represented as $[L]$ and $[M]$, where $L, M : \mathbb{R}^n \to \mathbb{R}^n$ are linear functions. But linear functions can be composed, and it is easy to check that the composite function $L \circ M : \mathbb{R}^n \to \mathbb{R}^n$ is also linear, hence it corresponds to another $n \times n$ matrix $[L \circ M]$. By definition we say that this is the *matrix product* of $[L]$ and $[M]$ and we write

$$[L][M] := [L \circ M].$$

More generally, if $L : \mathbb{R}^m \to \mathbb{R}^\ell$ and $M : \mathbb{R}^n \to \mathbb{R}^m$ are linear functions then the matrices $[L]$ and $[M]$ are defined, with shapes $\ell \times m$ and $m \times n$, respectively. Since M maps **into** $\mathbb{R}^m$ and L maps **from** $\mathbb{R}^m$ the composite function $L \circ M : \mathbb{R}^n \to \mathbb{R}^\ell$ exists, and we can define the matrix $[L][M] : [L \circ M]$, which has shape $\ell \times n$. Let us investigate how to

compute the matrix entries of $[L][M]$ from the matrix entries of $[L]$ and $[M]$.

This is an extremely fruitful concept and there are many ways to describe it. I will use a standard notation from linear algebra. Let $A = (a_{ij})$ and $B = (b_{ij})$ be matrices where a_{ij}, b_{ij} are the entries of A, B in the ith row and jth column. Suppose that A has shape $\ell \times m$ and B has shape $m \times n$. Then the matrix AB is defined with shape $\ell \times n$ and its i, j entry is given as follows:

$$(i, j \text{ entry of } AB) = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

In various circumstances it is also useful to express this definition in terms of multiplications with row and column vectors:

$$\begin{aligned}
(i, j \text{ entry of } AB) &= (i\text{th row of } A)(j\text{th column of } B) \\
(i\text{th row of } AB) &= (i\text{th row of } A)B \\
(j\text{th column of } AB) &= A(j\text{th column of } B) \\
AB &= \sum_{k=1}^{m} (k\text{th column of } A)(k\text{th row of } B).
\end{aligned}$$

This notation takes some getting used to but you should make the effort because it is very important in all areas of mathematics.

The proof is not very interesting, but here is it.

Proof. Write $[L] = A = (a_{ij})$ and $[M] = B = (b_{ij})$, where $L : \mathbb{R}^m \to \mathbb{R}^\ell$ and $M : \mathbb{R}^n \to \mathbb{R}^m$ are linear, so that A has shape $\ell \times m$ and B has shape $m \times n$. By definition we have $AB = [L \circ M]$, so that[6]

$$\begin{aligned}
(j\text{th column of } AB) &= (j\text{th column of } [L \circ M]) \\
&= [(L \circ M)(\mathbf{e}_j)] \\
&= [L(M(\mathbf{e}_j))] \\
&= [L(j\text{th column of } M)] \\
&= \left[L \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix} \right] \\
&= [L(b_{1j}\mathbf{e}_1 + b_{2j}\mathbf{e}_2 + \cdots + b_{mj}\mathbf{e}_m)] \\
&= b_{1j}[L(\mathbf{e}_1)] + b_{2j}[L(\mathbf{e}_2)] + \cdots + b_{mj}[L(\mathbf{e}_m)]
\end{aligned}$$

[6] Forgive me for using the notation $\mathbf{e}_i$ to denote the basis vectors in both $\mathbb{R}^m$ and $\mathbb{R}^n$ even though these vectors have different numbers of entries.

$$
\begin{aligned}
&= \sum_{k=1}^{m} b_{kj}(k\text{th column of } [L]).\\
&= \sum_{k=1}^{m} b_{kj} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{\ell k} \end{pmatrix}\\
&= \begin{pmatrix} \sum_{k=1}^{m} a_{1k}b_{kj} \\ \vdots \\ \sum_{k=1}^{m} a_{\ell k}b_{kj} \end{pmatrix}.
\end{aligned}
$$

Since the i, j entry of AB is just the ith entry of the jth column we obtain the desired formula.

□

As an interesting example, let me present the "correct" proof of the angle sum trigonometric identities.

Correct Proof of the Angle Sum Trigonometric Identities

Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ denote the (linear) function that rotates each vector counterclockwise around the origin by angle θ. It is geometrically obvious that for all angles α, β we have

$$
\begin{array}{ccc}
R_\alpha \circ R_\beta & = & R_{\alpha+\beta} \\
(\text{rotate by } \beta \text{ then rotate by } \alpha) & = & (\text{rotate once by } \alpha + \beta).
\end{array}
$$

On the other hand, we showed that the rotation function R_θ corresponds to the matrix

$$
[R_\theta] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.
$$

By combining these observations with the definition of matrix multiplication we obtain

$$
\begin{aligned}
&\begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}\\
&= [R_{\alpha+\beta}]\\
&= [R_\alpha \circ R_\beta]\\
&= [R_\alpha][R_\beta]\\
&= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}\\
&= \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{pmatrix}.
\end{aligned}
$$

And comparing matrix entries gives

$$\begin{cases} \cos(\alpha+\beta) & = & \cos\alpha\cos\beta - \sin\alpha\sin\beta, \\ \sin(\alpha+\beta) & = & \sin\alpha\cos\beta + \cos\alpha\sin\beta. \end{cases}$$

There is no need to ever memorize these formulas. You only need to memorize the form of rotation matrix $[R_\theta]$ and use the obvious fact that $R_{\alpha+\beta} = R_\alpha \circ R_\beta$.

Finally, we obtain the main theorem of this section.

Complex Numbers as Linear Functions

For each complex number $\alpha \in \mathbb{C}$ we consider the function $L_\alpha : \mathbb{C} \to \mathbb{C}$ defined by:

$$L_\alpha(\beta) := \alpha\beta.$$

This function is called "multiply by α". If we view $\mathbb{C} = \mathbb{R}^2$ as a vector space then the function L_α is $\mathbb{R}$-linear since for all $b, c \in \mathbb{R}$ and $\beta, \gamma \in \mathbb{C}$ we have

$$L_\alpha(b\beta + c\gamma) = \alpha(b\beta + c\gamma) = b(\alpha\beta) + c(\alpha\gamma) = bL_\alpha(\beta) + cL_\alpha(\gamma).$$

Therefore it corresponds to a 2 × 2 matrix with real entries. To find this matrix, let $\alpha = a+bi$ and consider the standard basis vectors $1+0i$ and $0+1i$. Since $L_\alpha(1+0i) = a+bi$ and $L_\alpha(0+1i) = -b+ai$ it follows that

$$[L_\alpha] = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

But more is true. We observe that multiplication of complex numbers corresponds to composition of linear functions. In other words, for any $\alpha, \beta \in \mathbb{C}$ we have $L_{\alpha\beta} = L_\alpha \circ L_\beta$:

$$L_{\alpha\beta}(\gamma) = (\alpha\beta)(\gamma) = \alpha(\beta\gamma) = \alpha L_\beta(\gamma) = L_\alpha(L_\beta(\gamma)) = (L_\alpha \circ L_\beta)(\gamma).$$

Then by definition of matrix multiplication we have $[L_{\alpha\beta}] = [L_\alpha \circ L_\beta] = [L_\alpha][L_\beta]$ and it follows that multiplication of complex numbers can be viewed as matrix multiplication:

$$(a+bi)(c+di) = (ac-bd) + (ad+bc)i$$
$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{pmatrix}.$$

Finally, we observe that real numbers correspond to scalar matrices and complex numbers of length 1 correspond to rotation matrices:

$$[L_{r+0i}] = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \quad \text{and} \quad [L_{\cos\theta + i\sin\theta}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

It follows that complex numbers can be viewed as the set of (linear) functions $\mathbb{R}^2 \to \mathbb{R}^2$ that can be obtained by scaling and rotation.

This modern point of view was put forward by Hamilton in order to give a "real meaning" to the "imaginary numbers". Under this scheme we see that

$$\sqrt{-1} = (\text{rotate by } 90^\circ).$$

That's not imaginary at all.[7]

1.6 Euler's Formula and Roots of Unity

At the beginning of this chapter I mentioned the fact that the "square root function" $x \mapsto \sqrt{x}$ is not really a function. If x is real and positive then we could take $\sqrt{x}$ to be the unique real positive square root of x. But if x is a negative real number or a complex number then the symbol $\sqrt{x}$ represents two different complex numbers, and there is no good reason to prefer one over the other. Because of this non-uniqueness we must be careful when interpreting formulas such as

$$\sqrt{ab} = \sqrt{a}\sqrt{b}.$$

For example, if $a = b = -1$ then this formula seems to imply that

$$i^2 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1,$$

which is false. This caused significant confusion in the early days of complex numbers.

More generally, if $\alpha \in \mathbb{C}$ is a nonzero complex number then the expression $\sqrt[n]{\alpha}$ or $\alpha^{1/n}$ represents n distinct complex numbers. This was slowly clarified during the 1700s and it finally became transparent in the 1800s with the geometric interpretation of complex numbers. The first step was made by de Moivre in 1707.

[7] We have shown that $\mathbb{C}$ is a ring, a field, a real vector space, and a collection of 2×2 matrices with real entries. In very modern terms we could summarize this by saying that $\mathbb{C}$ is a two-dimensional commutative real division algebra with a two-dimensional faithful representation (and I could probably add more adjectives). Never mind. The point is that the complex numbers have a lot of interesting structure, which motivates all of the structures that we will discuss in this course.

De Moivre's Formula (1707)

For any angle θ and for any integer $n \geqslant 0$ we have

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta).$$

This is not difficult to prove once it is observed.[8] The hard part is to observe it in the first place. In fact, de Moivre stated the theorem in a much more complicated way because he did not use complex numbers. We'll return to this below.

The modern proof is essentially just that "n successive rotations by angle θ" is the same as "one single rotation by angle $n\theta$". This point of view was preceded by an interpretation using the language of Calculus.

Euler's Formula (1748)

For any complex number $\alpha \in \mathbb{C}$ Euler considered the following power series:

$$\exp(\alpha) := 1 + \alpha + \frac{\alpha^2}{2} + \frac{\alpha^3}{6} + \cdots = \sum_{k=0}^{\infty} \frac{\alpha^n}{n!}.$$

It turns out that this power series always converges. Furthermore, for any complex numbers $\alpha, \beta \in \mathbb{C}$ one can show that

$$\exp(\alpha)\exp(\beta) = \exp(\alpha + \beta).$$

The number $e := \exp(1) \approx 2.71828$ is today called *Euler's constant.* For any integer $n \geqslant 1$ we observe that

$$\exp(n) = \exp(1 + 1 + \cdots + 1) = \exp(1)^n = e^n.$$

For this reason it is standard to use the notation

$$\text{“}e^{\alpha}\text{”} := \exp(\alpha),$$

even though it is far from clear how to take "e to the power of π", for example. Using this language, Euler made the discovery that for any real number θ we have

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

[8] For example, it can be proved by induction using the angle sum trigonometric formulas.

which immediately gives a proof of de Moivre's formula:

$$(\cos\theta + i\theta)^n = (e^{i\theta})^n = e^{in\theta} = \cos(n\theta) + i\sin(n\theta).$$

Proof: I will assume, as Euler did, that the power series always converges. Rigorous treatment of convergence only emerged in the 1800s. To prove the identity $\exp(\alpha+\beta) = \exp(\alpha)\exp(\beta)$ we first recall the *binomial theorem*:

$$(\alpha+\beta)^m = \sum_{k+\ell=m} \frac{m!}{k!\ell!}\,\alpha^k\beta^\ell.$$

If we multiply the power series for $\exp(\alpha)$ and $\exp(\beta)$ then the binomial theorem gives the desired simplification:

$$\begin{aligned}
\exp(\alpha)\exp(\beta) &= \left(\sum_{k\geqslant 0}\frac{\alpha^k}{k!}\right)\left(\sum_{\ell\geqslant 0}\frac{\beta^\ell}{\ell!}\right)\\
&= \sum_{m\geqslant 0}\left(\sum_{k+\ell=m}\frac{\alpha^k}{k!}\frac{\beta^\ell}{\ell!}\right)\\
&= \sum_{m\geqslant 0}\frac{1}{m!}\left(\sum_{k+\ell=m}\frac{m!}{k!\ell!}\,\alpha^k\beta^\ell\right)\\
&= \sum_{m\geqslant 0}\frac{1}{m!}(\alpha+\beta)^m\\
&= \exp(\alpha+\beta).
\end{aligned}$$

Finally, to prove Euler's formula we use a direct computation:

$$\begin{aligned}
\exp(i\theta) &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \cdots\\
&= 1 + i\theta + \frac{-\theta^2}{2!} + \frac{-i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} + \cdots\\
&= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right).
\end{aligned}$$

Euler immediately recognized these as the power series expansions of $\cos\theta$ and $\sin\theta$, which had been discovered by Newton. □

Apart from being interesting and useful, Euler's formula allows us to simplify notation by writing $e^{i\theta}$ instead of $\cos\theta + i\sin\theta$. We will do this from now on.

Roots of Unity

Fix an integer $n \geqslant 1$ and consider the complex number $\omega = e^{2\pi i/n}$. I claim that the equation $x^n = 1$ has the complete solution

$$x = 1, \omega, \omega^2, \ldots, \omega^{n-1}.$$

To see this we first observe that

$$(\omega)^n = (e^{2\pi i/n})^n = e^{2\pi i} = \cos(2\pi) + i\sin(2\pi) = 1.$$

Thus for any integer k we have

$$(\omega^k)^n = (\omega^n)^k = 1^k = 1.$$

To see that this is the **complete** solution we must show that the n numbers ω^k with $k = 0, 1, \ldots, n-1$ are distinct. This follows from the fact that they represent distinct points of the complex plane.[9] Indeed, since the number $e^{i\theta}$ corresponds to the point $(\cos\theta, \sin\theta)$ in the Cartesian plane, we observe that $e^{i\alpha} = e^{i\beta}$ if and only if $\alpha - \beta$ is an integer multiple of 2π. It follows from this that for all integers $k, \ell \in \mathbb{Z}$ we have $\omega^k = \omega^\ell$ if and only if $k - \ell$ is a multiple of n.[10]

More generally, we can describe the nth roots of an arbitrary nonzero complex number $\alpha \in \mathbb{C}$ as follows. We first write $\alpha = re^{i\theta}$ in polar form, so that $r > 0$. Let $r' > 0$ denote the unique positive nth root of r and let $\alpha' := r'e^{i\theta/n}$. We observe that

$$(\alpha')^n = (r'e^{i\theta/n})^n = (r')^n(e^{i\theta/n})^n = re^{i\theta} = \alpha,$$

and we say that α' is the *principal nth root* of α. Then I claim that the equation $x^n = \alpha$ has the complete solution

$$x = \alpha', \alpha'\omega, \alpha'\omega^2, \ldots, \alpha'\omega^{n-1}.$$

Indeed, each of these is a solution because

$$(\alpha'\omega^k)^n = (\alpha')^n(\omega^k)^n = \alpha \cdot 1 = \alpha,$$

and they are distinct because $\alpha'\omega^k = \alpha'\omega^\ell$ if and only if $\omega^k = \omega^\ell$.

Geometrically, the $n$ths roots of α form a regular n-gon in the complex plane, centered at the origin.

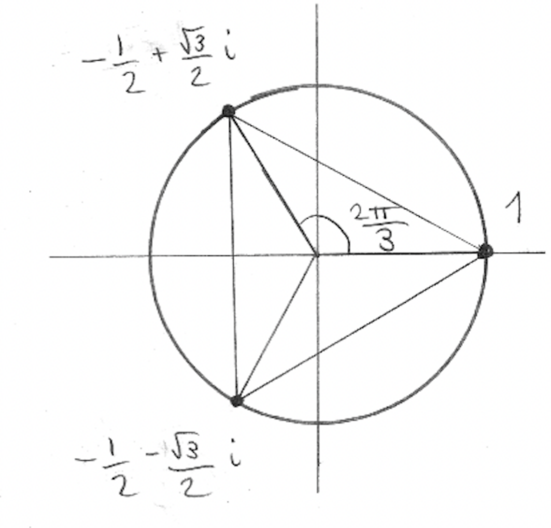
[9] We also need to know that an equation of degree n can have **no more than n roots**. You will prove this on the homework and we will discuss it more in the next section.

[10] This idea will reappear below when we discuss "modular arithmetic".

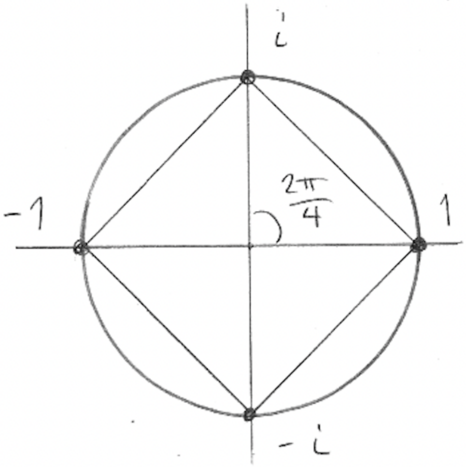
Examples:

- $n = 2$: Let $\omega = e^{2\pi i/2} = e^{\pi i} = -1$. Then the 2nd roots of 1 are
$$\omega^0 = 1 \quad \text{and} \quad \omega^1 = -1.$$
If α' is any square root of the nonzero complex number α, then the complete set of square roots is
$$\alpha'\omega^0 = \alpha' \quad \text{and} \quad \alpha'\omega^1 = -\alpha'.$$
That was pretty boring.

- $n = 3$: Let $\omega = e^{2\pi i/3} = \cos(2\pi/3) + i\sin(2\pi/3) = -1/2 + i\sqrt{3}/2 = (-1 + i\sqrt{3})/2$. Then the 3rd roots of 1 are
$$\begin{aligned}
\omega^0 &= 1, \\
\omega^1 &= (-1 + i\sqrt{3})/2, \\
\omega^2 &= e^{4\pi i/3} = \cos(4\pi/3) + i\sin(4\pi/3) = -1/2 - i\sqrt{3}/2 = (-1 - i\sqrt{3})/2.
\end{aligned}$$
Here is a picture:


$-\frac{1}{2} + \frac{\sqrt{3}}{2} i$
$\frac{2\pi}{3}$
1
$-\frac{1}{2} - \frac{\sqrt{3}}{2} i$

- $n = 4$: Let $\omega = e^{2\pi i/4} = e^{\pi i/2} = \cos(\pi/2) + i\sin(\pi/2) = i$. The 4th roots of unity are
$$\begin{aligned}
\omega^0 &= e^0 = 1, \\
\omega^1 &= e^{\pi i/2} = i, \\
\omega^2 &= e^{\pi i} = -1, \\
\omega^3 &= e^{3\pi i/2} = -i.
\end{aligned}$$
Here is a picture:



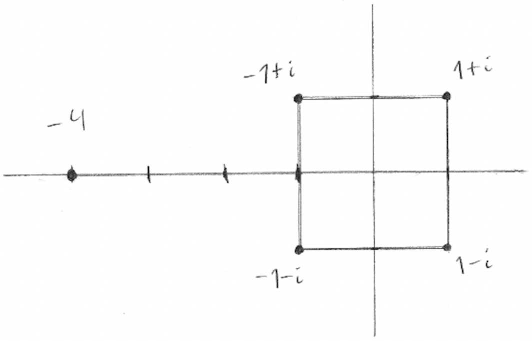
More generally, let's compute the 4th roots of $\alpha = -4$. First we express $\alpha = 4e^{\pi i}$ in polar form, so the principal 4th root is

$$\alpha' = \sqrt[4]{4} \cdot e^{\pi i/4} = \sqrt{2}[\cos(\pi/4) + i\sin(\pi/4)] = \sqrt{2}(1/\sqrt{2} + i\sqrt{2}) = 1 + i.$$

Then the complete set of 4th roots of -4 is

$$\begin{aligned}
\alpha'\omega^0 &= 1\alpha' = 1 + i, \\
\alpha'\omega^1 &= i\alpha' = -1 + i, \\
\alpha'\omega^2 &= -1\alpha' = -1 - i, \\
\alpha'\omega^3 &= -i\alpha' = 1 - i.
\end{aligned}$$

These form a square in the complex plane:



As an application, we can use these roots to factor the polynomial $x^4 + 4$:

$$x^4 + 4 = (x - (1 + i))(x - (-1 + i))(x - (-1 - i))(x - (1 - i)).$$

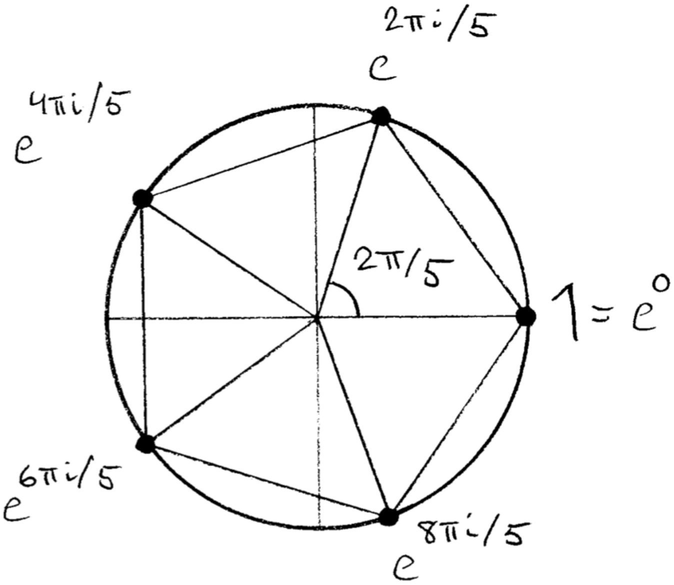
In 1702, Gottfried Leibniz claimed that the polynomial x^4+4 cannot be factored over the real numbers. However, we can show that he was wrong by grouping the four complex roots into "conjugate pairs":

$$\begin{aligned} x^4+4 &= [(x-(1+i))(x-(1-i))][(x-(-1+i))(x-(-1-i))] \\ &= (x^2-2x+2)(x^2+2x+2). \end{aligned}$$

- $n=5$: Let $\omega = e^{2\pi i/5} = \cos(2\pi/5) + i\sin(2\pi i/5)$. The 5th roots of unity are

$$\begin{aligned} \omega^0 &= 1, \\ \omega^1 &= e^{2\pi i/5} = \cos(2\pi/5) + i\sin(2\pi/5), \\ \omega^2 &= e^{4\pi i/5} = \cos(4\pi/5) + i\sin(4\pi/5), \\ \omega^3 &= e^{6\pi i/5} = \cos(6\pi/5) + i\sin(6\pi/5), \\ \omega^4 &= e^{8\pi i/5} = \cos(8\pi/5) + i\sin(8\pi/5), \end{aligned}$$

which correspond to the vertices of a regular pentagon in the Cartesian plane:



On the homework you will show that these numbers can also be expressed in terms of integers and square roots. For example, you will show that

$$\cos\left(\frac{2\pi}{5}\right) = \frac{-1+\sqrt{5}}{4}.$$

Is it always true that the roots of unity can be expressed in terms of integers and square roots? As a preview of things to come, let me mention the main theorem in this subject.

Preview of the Gauss-Wantzel Theorem

Consider an integer $n \geqslant 1$ and define the *phi-function*:[11]

$$\phi(n) := \#\{k \in \mathbb{Z} : 1 \leqslant k \leqslant n-1 \text{ and } \gcd(k,n) = 1\}.$$

This number is always even. Suppose that $\phi(n)/2 = m_1 m_2 \cdots m_k$ for some integers $m_1, \ldots, m_k \geqslant 2$. Then I claim that the number $\cos(2\pi/n)$ can be expressed in terms of integers and m_ith roots for the various i. If $\phi(n)$ is a power of 2 then there exists a formula for $\cos(2\pi/n)$ involving only integers and square roots.

For example, since 5 is prime, all of the numbers $1, 2, 3, 4$ are coprime to 5 and hence $\phi(5) = 4 = 2^2$. Since $\phi(5)$ is a power of 2, the theorem guarantees that $\cos(2\pi/5)$ can be expressed in terms of integers and square roots, as you will show on the homework.

The origin of the theorem is Gauss' discovery (at the age of 19) that the number $\cos(2\pi/17)$ can be expressed in terms of integers and square roots:

$$\cos\left(\frac{2\pi}{17}\right) = \frac{-1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + \sqrt{17 + 3\sqrt{17} - \sqrt{34 - 2\sqrt{17}} - 2\sqrt{34 + 2\sqrt{17}}}}{16}$$

According to the theorem, we know that such a formula is possible because $\phi(17) = 16 = 2^4$ is a power of 2. Gauss' discovery was surprising because it implies that the regular 17-gon can be constructed with straightedge and compass, a construction that was not known to the ancient Greeks.

In general, we will see that $\phi(n)$ is a power of 2 if and only if n can be expressed as a power of 2 times a product of distinct *Fermat prime numbers* of the form $p = 2^m + 1$. For example, $p = 17 = 2^4 + 1$ is a Fermat prime. Fermat had conjectured that **every** number of the form $2^m + 1$ is prime, but this turned out to be quite wrong. Today the only known Fermat primes are

$$3, 5, 17, 257, \text{ and } 65537,$$

and it is an open question whether there exist any others.

2 Introduction to Polynomials

2.1 Rings of Polynomials

We have talked about polynomials in an intuitive way, but we have not been careful with our definitions. Here is the modern, abstract, definition of polynomials.

[11] The notation $\gcd(k, n)$ represents the *greatest common divisor* of k and n. We will study this in detail in the next section.

Definition of Polynomials

Let $\mathbb{F}$ be a field and let "x" be an abstract symbol. By a *polynomial in x over* $\mathbb{F}$ we mean a formal expression

$$f(x) = \sum_{k\geqslant 0} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots ,$$

where the *coefficients* $a_0, a_1, a_2, \ldots$ are elements of $\mathbb{F}$ and only finitely many of these coefficients are nonzero. If a_n is the highest nonzero coefficient then we will say that $f(x)$ has *degree n* and we will write

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

For example:

$$\begin{aligned} \deg(x^2) &= 2, \\ \deg(7x^3 + 1) &= 3, \\ \deg(5) &= 0. \end{aligned}$$

The polynomials of degree 0 are just the nonzero constants. (For the degree of the zero constant, see below.) Let us denote the set of polynomials by

$$\mathbb{F}[x] = \{\text{polynomials in } x \text{ over } \mathbb{F}\}.$$

We can view this set as a ring by pretending that x is a number and performing arithmetic as usual. To be precise, we define addition and multiplication of polynomials as follows:

$$\begin{aligned} \left(\sum_{k\geqslant 0} a_k x^k\right) + \left(\sum_{k\geqslant 0} b_k x^k\right) &:= \sum_{k\geqslant 0} (a_k + b_k) x^k \\ \left(\sum_{k\geqslant 0} a_k x^k\right)\left(\sum_{\ell\geqslant 0} b_k x^k\right) &:= \sum_{m\geqslant 0}\left(\sum_{k+\ell=m} a_k b_\ell\right) x^m. \end{aligned}$$

The additive and multiplicative identity elements are the zero and one polynomials:

$$\begin{aligned} 0(x) &:= 0 + 0x + 0x^2 + 0x^3 + \cdots , \\ 1(x) &:= 1 + 0x + 0x^2 + 0x^3 + \cdots . \end{aligned}$$

However, we usually don't usually make distinction between the numbers $0, 1$ and the polynomials $0(x), 1(x)$. In fact, we can think of $\mathbb{F}$ as a subring of $\mathbb{F}[x]$ by identifying each element $a \in \mathbb{F}$ with the corresponding constant polynomial:

$$a = a + 0x + 0x^2 + 0x^3 + \cdots .$$

An important and basic fact about polynomials is the *additivity of degree*:

$$\deg(fg) = \deg(f) + \deg(g).$$

To prove this formula, suppose that $\deg(f) = m$ and $\deg(g) = n$. By definition this means that

$$\begin{aligned} f(x) &= a_m x^m + a_{m-1}x^{m-1} + \cdots a_1 x + a_0, \\ g(x) &= b_n x^n + b_{n-1}x^{n-1} + \cdots b_1 x + b_0, \end{aligned}$$

where $a_m \neq 0$ and $b_n \neq 0$. But then we have $a_m b_n \neq 0$ and

$$f(x)g(x) = a_m b_n x^{m+n} + \text{lower terms},$$

so that $\deg(fg) = m + n = \deg(f) + \deg(g)$. Strictly speaking, this formula only applies to nonzero polynomials. In order to make the formula true in general it is convenient to define the degree of the zero polynomial as follows:

$$\deg(0) := \text{“} -\infty\text{”}.$$

We don't think of this as a number, but just a symbol with the properties $-\infty < a$ and $-\infty + a = -\infty$ for all $a \in \mathbb{F}$.

Some Remarks:

- The ring $\mathbb{F}[x]$ is not a field. To see this it is enough to show that some nonzero element has no multiplicative inverse. We will show that $x \in \mathbb{F}[x]$ has no multiplicative inverse. Let us suppose for contradiction that there exists a polynomial $f(x) \in \mathbb{F}[x]$ satisfying $xf(x) = 1$. Then taking degrees gives

$$\begin{aligned} xf(x) &= 1 \\ \deg(x) + \deg(f) &= \deg(1) \\ 1 + \deg(f) &= 0 \\ \deg(f) &= -1, \end{aligned}$$

  which is a contradiction because there is no such thing as a polynomial of degree -1. In other words, we have shown that the expression $1/x$ is not a polynomial. We will call it a *rational expression.* Later we will consider the *field of rational expressions* $\mathbb{F}(x)$, which are basically fractions of polynomials.

- The set of polynomials $\mathbb{F}[x]$ can also be thought of as a *vector space over* $\mathbb{F}$ with scalar multiplication

$$a\left(\sum_{k\geqslant 0} b_k x^k\right) = \sum_{k\geqslant 0}(ab_k)x^k.$$

By convention we say that two polynomials are equal if and only if they have the same coefficients. This implies that the vector space $\mathbb{F}[x]$ is **infinite dimensional** with basis

$$1, x, x^2, x^3, \ldots.$$

Of course, we are accustomed to thinking of polynomials as functions, not just formal expressions. We will discuss the relationship between these points of view in the next section.

2.2 Descartes' Theorem

There is a deep analogy between the rings $\mathbb{Z}$ and $\mathbb{F}[x]$, which is based on the following theorem.[12]

Division With Remainder

(1) For all integers $a, b \in \mathbb{Z}$ with $b \neq 0$ there exist unique integers $q, r \in \mathbb{Z}$ (called the *quotient* and *remainder*) satisfying

$$\begin{cases} a = bq + r, \\ 0 \leqslant r < |b|. \end{cases}$$

(2) Let $\mathbb{F}$ be a field. Then for all polynomials $f(x), g(x) \in \mathbb{F}[x]$ with $g(x) \neq 0(x)$ there exist unique polynomials $q(x), r(x) \in \mathbb{F}[x]$ (called the *quotient* and *remainder*) satisfying

$$\begin{cases} f(x) = g(x)q(x) + r(x), \\ \deg(r) < \deg(g). \end{cases}$$

Note: The condition $\deg(r) < \deg(g)$ includes the possibility that the remainder is zero, i.e., that $\deg(r) = -\infty$.

The idea of the proof in both cases is to define an algorithm and to prove that this algorithm gives the desired result. We will prove existence here and you will prove uniqueness on the homework.

Proof for Integers: Let $a, b \in \mathbb{Z}$ with $b \neq 0$ and consider the set

$$S = \{a - qb : q \in \mathbb{Z}\} = \{\ldots, a - 2b, a - b, a, a + b, a + 2b, \ldots\} \subseteq \mathbb{Z}.$$

[12] Later we will make this analogy more precise when we discuss the concept of a *Euclidean domain*.

Let r be the smallest non-negative element of this set. By definition we know that $a = qb + r$ for some integer $q \in \mathbb{Z}$ and we also know that $0 \leqslant r$. It remains only to show that $r < |b|$. So let us assume for contradiction that $r \geqslant |b|$. Since $b \neq 0$ this implies that

$$0 \leqslant r - |b| < r.$$

On the other hand, we observe that $r - |b| = (a - qb) - |b| = a - (q \pm 1)b \in S$. Thus we have found a non-negative element of S that is strictly smaller than r. Contradiction. □

Proof for Polynomials Over a Field: Let $\mathbb{F}$ be a field and consider two polynomials $f(x), g(x) \in \mathbb{F}[x]$ with $g(x) \neq 0(x)$. Furthermore, consider the set

$$S = \{f(x) - q(x)g(x) : q(x) \in \mathbb{F}[x]\} \subseteq \mathbb{F}[x].$$

Let $r(x)$ be some element of S with **minimal degree** (allowing for the possibility that $r(x) = 0(x)$ and hence $\deg(r) = -\infty$). By definition we know that $f(x) = q(x)g(x) + r(x)$ for some $q(x) \in \mathbb{F}[x]$ and it remains only to show that $\deg(r) < \deg(g)$. So let us assume for contradiction that $\deg(r) \geqslant \deg(g)$. To be specific, since $g(x) \neq 0(x)$ we may write

$$g(x) = a_m x^n + \text{lower terms} \quad \text{and} \quad r(x) = b_n x^n + \text{lower terms},$$

where $a_m \neq 0$ and $m \leqslant n$. Then since $n - m \geqslant 0$ we may construct the following polynomial:[13]

$$\begin{aligned} h(x) &:= r(x) - \frac{b_n}{a_m} x^{n-m} g(x) \\ &= (b_n x^n + \text{lower terms}) - \frac{b_n}{a_m} x^{n-m} \left(a_m x^m + \text{lower terms}\right) \\ &= (b_n - b_n)\, x^n + \text{lower terms}. \end{aligned}$$

Note that the coefficient of x^n in $h(x)$ is zero, and hence $\deg(h) < n = \deg(r)$. On the other hand, we observe that $h(x)$ is an element of S:

$$\begin{aligned} h(x) &= r(x) - \frac{b_n}{a_m} x^{n-m} g(x) \\ &= (f(x) - q(x)g(x)) - \frac{b_n}{a_m} x^{n-m} g(x) \\ &= f(x) - \left(q(x) + \frac{b_n}{a_m} x^{n-m}\right) g(x) \in S. \end{aligned}$$

Thus $h(x)$ is an element of S with strictly smaller degree than $r(x)$. Contradiction. □

I assume you are familiar with long division of integers. Long division of polynomials is actually easier because it doesn't involve any "carrying". For example, suppose that $f(x) = 2x^4 - 6x^3 + x - 1$ and $g(x) = 2x^2 + 1$. The algorithm tells us first to multiply $g(x)$ by a

[13] Here we use that fact that $\mathbb{F}$ is a field to divide by a_m.

suitable “monomial” so that it has the same “leading term” as $f(x)$ and then subtract this from $f(x)$ to “eliminate” this leading term. To be specific, we multiply $g(x)$ by the monomial x^2 to obtain $2x^4 + x^2$ whose leading term matches $f(x)$. Then we repeat the process until it is impossible to continue:[14]

$$\begin{array}{rrrrrr}
 & & & x^2 & -3x & -\frac{1}{2} \\ \cline{2-6}
2x^2+1\big) & 2x^4 & -6x^3 & & +x & -1 \\
 & -2x^4 & & -x^2 & & \\ \cline{2-4}
 & & -6x^3 & -x^2 & +x & \\
 & & 6x^3 & & +3x & \\ \cline{3-5}
 & & & -x^2 & +4x & -1 \\
 & & & x^2 & & +\frac{1}{2} \\ \cline{4-6}
 & & & & 4x & -\frac{1}{2}
\end{array}$$

In the end we obtain a quotient $q(x) = x^2 - 3x - 1/2$ and a remainder $r(x) = 4x - 1/2$, which satisfy the desired properties:

$$\begin{cases} (2x^4 - 6x^3 + x - 1) = (2x^2 + 1)(x^2 - 3x - 1/2) + (4x - 1/2), \\ \deg(4x - 1/2) < \deg(2x^2 + 1). \end{cases}$$

Polynomial division with remainder was first used for theoretical purposes by René Descartes (1631) in his *Geometry.* The following theorem is the foundational property of polynomials, of similar importance to the Pythagorean theorem in geometry.

Descartes’ Factor Theorem (1631)

Consider a field $\mathbb{F}$, a polynomial $f(x) \in \mathbb{F}[x]$ and a constant $a \in \mathbb{F}$. Dividing $f(x)$ by $x - a$ gives

$$f(x) = (x - a)q(x) + r(x)$$

for some polynomials $q(x), r(x) \in \mathbb{F}[x]$ with $\deg(r) < \deg(x - a) = 1$. The condition on the degree implies that $r(x) = c$ for some constant $c \in \mathbb{F}$, either zero or nonzero. To determine this constant we substitute $x = a$:

$$\begin{aligned} f(a) &= (a - a)q(a) + c \\ f(a) &= 0q(a) + c \\ f(a) &= c. \end{aligned}$$

It follows from this that

$$f(a) = 0 \quad \Longleftrightarrow \quad f(x) = (x - a)q(x) \text{ for some polynomial } q(x).$$

[14] There are different ways to typeset this. I used a package to do it automatically, which I don’t like very much, but is much easier than doing it manually.

In other words, the constant $a \in \mathbb{F}$ is a root of $f(x)$ if and only if the polynomial $x - a$ is a divisor of $f(x)$. We will use this to prove by induction that

a polynomial $f(x) \in \mathbb{F}[x]$ of degree $n \geqslant 0$ can have at most n roots in $\mathbb{F}$.

Indeed, a polynomial of degree 0 is a nonzero constant, which has no roots. So let $\deg(f) = n \geqslant 1$. If $f(x)$ has no roots then we are happy because $0 \leqslant n$. Otherwise, $f(x)$ must have some root $f(a) = 0$ with $a \in \mathbb{F}$. From the above remarks this implies that $f(x) = (x - a)q(x)$ for some polynomial $q(x) \in \mathbb{F}[x]$, which must have degree $n - 1$:

$$\begin{aligned} \deg(f) &= \deg((x-a)q) \\ n &= \deg(x-a) + \deg(q) \\ n &= 1 + \deg(q). \end{aligned}$$

But if $b \neq a$ is any other root of $f(x)$ then we must have

$$\begin{aligned} f(x) &= (x-a)q(x) \\ f(b) &= (b-a)q(b) \\ 0 &= (b-a)q(b) \\ 0 &= q(b), \end{aligned}$$

which implies that b is also a root of $q(x)$. Finally, since $q(x)$ has degree $n - 1$ we may assume by induction that $q(x)$ has at most $n - 1$ roots in $\mathbb{F}$, which implies that $f(x)$ has at most $1 + (n - 1) = n$ roots in $\mathbb{F}$.

This theorem has the following useful consequence that we record for future reference.

Only the Zero Polynomial Can Have Infinitely Many Roots

If $f(x) = 0(x)$ is the zero polynomial then every element of the field $\mathbb{F}$ is a root of $f(x)$. If the field has infinitely many elements then the zero polynomial has infinitely many roots. On the other hand, any nonzero polynomial has a finite degree, so Descartes' Theorem implies that it has finitely many roots.

2.3 Polynomials: Functions or Formal Expressions?

In this class we have defined polynomials in terms of their coefficients and we have said that two polynomials are equal when they have the same coefficients:

$$\left(\sum_k a_k x^k\right) = \left(\sum_k b_k x^k\right) \quad \Longleftrightarrow \quad a_k = b_k \text{ for all } k.$$

On the other hand, given any formal polynomial expression $f(x) = \sum_k a_k x^k$ we can define a function by "substitution" or "evaluation":

$$\begin{array}{rccc} f: & \mathbb{F} & \to & \mathbb{F} \\ & \alpha & \mapsto & \sum_k a_k \alpha^k. \end{array}$$

The question I want to raise now is whether two polynomials with the same evaluations must have the same coefficients. In other words:

$$\left(\sum_k a_k \alpha^k\right) = \left(\sum_k b_k \alpha^k\right) \text{ for all } \alpha \in \mathbb{F} \quad \overset{?}{\Longleftrightarrow} \quad a_k = b_k \text{ for all } k.$$

To show you that this is not a silly question I will you show you an example of two polynomials with different coefficients that nevertheless define the same function. In order to do this I must also show you an example of a field with only finitely many elements.

The Field with Three Elements

Consider the set $\mathbb{F}_3 = \{0, 1, 2\}$ of three elements with the following algebraic operations:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

These operations are called "arithmetic mod 3" and we will discuss the details later. For now I only want to observe that the structure $(\mathbb{F}_3, +, \cdot, 0, 1)$ satisfies the axioms of a field, therefore we may consider the ring of polynomials $\mathbb{F}_3[x]$ with coefficients in $\mathbb{F}_3$.

Now let us consider the following two polynomials:

$$\begin{aligned} f(x) &= x + 0, \\ g(x) &= x^3 + 0x^2 + 0x + 0. \end{aligned}$$

Clearly these polynomials do not have the same coefficients, but the following table shows that they do have the same values:

α	$f(\alpha)$	$g(\alpha)$
0	0	$0^3 = 0$
1	1	$1^3 = 1$
2	2	$2^3 = 2$

That's not good. Luckily this problem does not occur when our field $\mathbb{F}$ has infinitely many elements.

Polynomials Over an Infinite Field

Let $\mathbb{F}$ be an infinite field and let $f(x), g(x) \in \mathbb{F}[x]$ be formal polynomial expressions:

$$f(x) = \sum_k a_k x^k \quad \text{and} \quad g(x) = \sum_k b_k x^k.$$

If f and g define the same function $\mathbb{F} \to \mathbb{F}$ then I claim that $f(x)$ and $g(x)$ have the same coefficients. That is, if $f(\alpha) = g(\alpha)$ for all $\alpha \in \mathbb{F}$ then I claim that $a_k = b_k$ for all k.

To prove this we define the polynomial expression

$$h(x) := f(x) - g(x) = \sum_k (a_k - b_k) x^k.$$

If we can show that $h(x)$ is the zero polynomial (i.e., the polynomial with all zero coefficients) then we will conclude $a_k - b_k = 0$ and hence $a_k = b_k$ for all k. But we have assumed that $f(\alpha) = g(\alpha)$ for all $\alpha \in \mathbb{F}$ and hence

$$h(\alpha) = f(\alpha) - g(\alpha) = 0 \quad \text{for all } \alpha \in \mathbb{F}.$$

In other words, every element of $\mathbb{F}$ is a root of $h(x)$. If the field $\mathbb{F}$ has infinitely many elements then the remark in the previous section shows that $h(x)$ is the zero polynomial, as desired.

So, at least in the case of polynomials over $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, there is no distinction between formal polynomial expressions and polynomial functions.

2.4 Concept of a Splitting Field

We now proceed to the subtleties of Descartes' Theorem. If $f(x) \in \mathbb{F}[x]$ and $\deg(f) = n \geqslant 0$ then we have proved that $f(x)$ has **at most** n distinct roots in the field $\mathbb{F}$. However, it is a possibility that there exist **less than** n distinct roots, and there are two ways this can happen:

- The roots might exist in a larger field. For example, the polynomial $x^2 + 1 \in \mathbb{R}[x]$ has no roots in $\mathbb{R}$ but it has two roots $\pm i$ in $\mathbb{C}$. And the polynomial $x^2 - 2 \in \mathbb{Q}[x]$ has no roots in $\mathbb{Q}$, but it has two roots $\pm\sqrt{2}$ in $\mathbb{R}$.
- There might exist repeated roots. For example, the polynomial $x^3 - x^2 - x + 1 = (x-1)^2(x+1)$ of degree three has only two distinct roots: $+1$ and -1. But the root $+1$ occurs with multiplicity 2. So it is still the case that $x^3 - x^2 - x + 1$ has three roots, "counted with multiplicity".

Concept of a Splitting Field

Consider a polynomial $f(x) \in \mathbb{F}[x]$ of degree $n \geqslant 0$ with coefficients in a field $\mathbb{F}$ and let $\mathbb{E} \supseteq \mathbb{F}$ be a larger field. We say that $f(x)$ *splits over* $\mathbb{E}$ if there exists elements $r_1, \ldots, r_n \in \mathbb{E}$, not necessarily distinct, such that

$$f(x) = (x - r_1)(x - r_2) \cdots (x - r_n).$$

In other words, $f(x)$ has n roots in $\mathbb{E}$, *counted with multiplicity.* Later we will show that such a field always exists, and in fact the minimal such field is unique up to isomorphism. The minimal field over which $f(x)$ splits is called the *splitting field of* $f(x) \in \mathbb{F}[x]$.

Let me also mention that the factorization of $f(x)$ into polynomials of degree 1, when it exists, is necessarily unique.[15] Indeed, suppose that we have

$$(x - r_1)(x - r_2) \cdots (x - r_n) = (x - s_1)(x - s_2) \cdots (x - s_n)$$

for some elements $r_1, \ldots, r_n, s_1, \ldots, s_n$ of a field $\mathbb{E}$. Evaluating each side at $x = s_1$ gives

$$\begin{aligned} (s_1 - r_1)(s_1 - r_2) \cdots (s_1 - r_n) &= (s_1 - s_1)(s_1 - s_2) \cdots (s_1 - s_n) \\ &= 0(s_1 - s_2) \cdots (s_1 - s_n) \\ &= 0, \end{aligned}$$

which implies that $s_1 - r_i = 0$ and hence $s_1 = r_i$ for some index i. After re-indexing the elements $s_1, \ldots, s_n$ if necessary we may assume that $r_1 = s_1$ and then we may cancel the common factor $x - r_1 = x - s_1$ from each side:[16]

$$\begin{aligned} \cancel{(x - r_1)}(x - r_2) \cdots (x - r_n) &= \cancel{(x - s_1)}(x - s_2) \cdots (x - s_n) \\ (x - r_2) \cdots (x - r_n) &= (x - s_2) \cdots (x - s_n). \end{aligned}$$

By repeating the argument (i.e., by using induction) we may re-index the remaining elements $s_2, \ldots, s_n$ so that $r_1 = s_1$, $r_2 = s_2$, ... and $r_n = s_n$, as desired.

Let me emphasize that the concept of the splitting field is relative to field of coefficients. Examples:

- The polynomial $x^2 + 1 \in \mathbb{R}[x]$ has splitting field $\mathbb{C} \supseteq \mathbb{R}$. Indeed, this polynomial splits over $\mathbb{C}$ because $x^2 + 1 = (x - i)(x + i)$ with $\pm i \in \mathbb{C}$. To see that $\mathbb{C}$ is the minimal such field, suppose that there exists another field $\mathbb{C} \supseteq \mathbb{E} \supseteq \mathbb{R}$ such that $x^2 + 1$ splits over $\mathbb{E}$.

[15] In the next section we will prove more generally that any polynomial over any field has a unique factorization into irreducible polynomials, not necessarily of degree 1.

[16] You will investigate "cancellation" on the homework.

By definition this means that

$$x^2 + 1 = (x - r_1)(x - r_2) \quad \text{for some } r_1, r_2 \in \mathbb{E}.$$

Then substituting $x = i$ gives

$$0 = (i - r_1)(i - r_2),$$

which implies that $i = r_1$ or $i = r_2$. Either way, we must have $i \in \mathbb{E}$. Finally, I claim that every complex number is in $\mathbb{E}$, so that $\mathbb{E} = \mathbb{C}$. Indeed, for any $a, b \in \mathbb{R}$ we have $a, b \in \mathbb{E}$ because $\mathbb{R} \subseteq \mathbb{E}$. Then since $a, b, i \in \mathbb{E}$ we have $a + bi \in \mathbb{E}$ because $\mathbb{E}$ is a ring. In summary:

The polynomial $x^2 + 1$ has splitting field $\mathbb{C}$ over $\mathbb{R}$.

- On the other hand, if we regard $x^2 + 1$ as an element of $\mathbb{Q}[x]$ then I claim that the splitting field is

$$\mathbb{Q}(i) := \{a + bi : a, b \in \mathbb{Q}\} \supseteq \mathbb{Q},$$

which is strictly smaller than $\mathbb{C}$ because, e.g., $\sqrt{2}$ is in $\mathbb{C}$ but not in $\mathbb{Q}(i)$. Indeed, it is easy to check that $\mathbb{Q}(i)$ is a subring of $\mathbb{C}$. It is also a field since for any rational numbers $a, b \in \mathbb{Q}$ we have

$$\frac{1}{a + bi} = \left(\frac{a}{a^2 + b^2}\right) + \left(\frac{-b}{a^2 + b^2}\right) i,$$

where the coefficients $a/(a^2 + b^2)$ and $-b/(a^2 + b^2)$ are also rational numbers. And the polynomial $x^2 + 1$ splits over $\mathbb{Q}$ because $\pm i \in \mathbb{Q}$. Finally, we need to show that $\mathbb{Q}(i)$ is the **smallest** extension of $\mathbb{Q}$ over which $x^2 + 1$ splits. The proof is the same as above. Suppose that $\mathbb{Q}(i) \supseteq \mathbb{E} \supseteq \mathbb{Q}$ for some some field $\mathbb{E}$ over which $x^2 + 1$ splits. Say $x^2 + 1 = (x - r_1)(x - r_2)$ for some $r_1, r_2 \in \mathbb{E}$. Then substituting $x = i$ shows that $i = r_1$ or $i = r_2$. In either case this implies that $i \in \mathbb{E}$. Then for any $a, b \in \mathbb{Q}$ we have $a + bi \in \mathbb{E}$ and hence $\mathbb{E} = \mathbb{Q}(i)$. In summary:

The polynomial $x^2 + 1$ has splitting field $\mathbb{Q}(i)$ over $\mathbb{Q}$.

On the homework you will find the splitting field of $x^2 - 2$ over $\mathbb{Q}$.

3 Unique Prime Factorization

3.1 Definition of Euclidean Domains

Before proceeding with topic of polynomial equations, we pause to develop some general theory. Much of the theory of (commutative) rings is based on a deep analogy between the ring of integers and rings of polynomials over fields:

$$\mathbb{Z} \approx \mathbb{F}[x]$$

In order to describe this analogy we must first develop the language of "divisibility".

Divisibility in a Ring

Let $(R, +, \cdot, 0, 1)$ be a ring. Then for all $a, b \in R$ we define the notation

$$a|b \quad\Longleftrightarrow\quad \text{there exists } k \in R \text{ such that } ak = b.$$

It is important to note that the symbol "$a|b$" represents a whole sentence. It means that "a divides b" or "b is divisible by a". We have the following basic properties:

- $1|a$ for all $a \in R$,
- $a|0$ for all $a \in R$,
- $a|b$ and $b|c$ imply $a|c$.

Indeed, we have $1|a$ because $1a = a$ and we have $a|0$ because $a0 = 0$. Now suppose that $a|b$ and $b|c$. By definition this means that $ak = b$ and $b\ell = c$ for some $k, \ell \in R$. But then we also have

$$a(k\ell) = (ak)\ell = b\ell = c,$$

which implies that $a|c$.

The properties of divisibility in a general ring can be quite wild. In order to model the properties of $\mathbb{Z}$ and $\mathbb{F}[x]$ we make a further restriction.

Definition of Integral Domains

We say that a ring $(R, +, \cdot, 0, 1)$ is an *integral domain* (or just a *domain*) if for all $a, b \in R$,

$$ab = 0 \quad\Longrightarrow\quad a = 0 \text{ or } b = 0.$$

For example, the rings $\mathbb{Z}$ and $\mathbb{F}[x]$ are integral domains. For a non-example, consider the ring $\mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}$ of "arithmetic mod 4" with the following addition and multiplication tables:[17]

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

This ring is not an integral domain because $2 \cdot 2 = 0$ but $2 \neq 0$.

Every field is an integral domain since if $ab = 0$ and $b \neq 0$ then b^{-1} exists and we can multiply both sides by b^{-1} to obtain

$$\begin{aligned} ab &= 0 \\ abb^{-1} &= 0b^{-1} \\ a &= 0. \end{aligned}$$

Similarly, if $ab = 0$ and $a \neq 0$ then we must have $b = 0$. Not every integral domain is a field; for example $\mathbb{Z}$ and $\mathbb{F}[x]$ are not fields. However, every integral domain satisfies *multiplicative cancellation*:

$$ac = bc \quad \text{and } c \neq 0 \quad \Longrightarrow \quad a = b.$$

To see this, we write

$$\begin{aligned} ac &= bc \\ ac - bc &= 0 \\ (a-b)c &= 0. \end{aligned}$$

If $c \neq 0$ then since R is an integral domain we have $a - b = 0$ and hence $a = b$.

The theory of divisibility in integral domains is closer to our intuition coming from $\mathbb{Z}$ and $\mathbb{F}[x]$. For example, suppose that some nonzero elements $a, b \in R$ satisfy $a|b$ and $b|a$. By definition this means that $ak = b$ and $b\ell = a$ for some $k, \ell \in R$ and hence

$$\begin{aligned} b\ell &= a \\ ak\ell &= a \\ ak\ell - a &= 0 \\ a(k\ell - 1) &= 0. \end{aligned}$$

Since $a \neq 0$ this implies that $k\ell - 1 = 0$ and hence $k\ell = 1$. This is more interesting than it looks because there may not be many elements in R that have a multiplicative inverse.

Definition of Units

Let R be a ring. We say that $u \in R$ is a *unit of R* if there exists a (necessarily unique) multiplicative inverse $u^{-1} \in R$. We denote the set of units by

$$R^\times = \{u \in R : \exists v \in R, uv = 1\}.$$

[17] It is not necessarily clear that these operations satisfy the ring axioms, but they do. We will discuss this in detail later.

For example, I claim that

$$\mathbb{Z}^\times = \{\pm 1\} \quad \text{and} \quad \mathbb{F}[x]^\times = \{\text{nonzero constants}\}.$$

To prove this for integers, we first observe that $\pm 1 \in \mathbb{Z}$ are units because $1 \cdot 1 = 1$ and $(-1)(-1) = 1$. To see that every unit is one of these, suppose that some nonzero integers $a, b \in \mathbb{Z}$ satisfy $ab = 1$. Since a, b are nonzero we have $|a|, |b| \geqslant 1$. But if $|a| \geqslant 2$ then we obtain a contradiction:

$$1 = |ab| = |a||b| \geqslant |a| \geqslant 2.$$

Hence $|a| = 1$, and a symmetric argument shows that $|b| = 1$.

To prove the result for polynomials, we first observe that each nonzero constant $a \in \mathbb{F}[x]$ is a unit whose inverse is the nonzero constant $1/a$. To see that every unit has this form, suppose that some nonzero $f(x), g(x) \in \mathbb{F}[x]$ satisfy $f(x)g(x) = 1$, so that

$$\deg(f) + \deg(g) = \deg(fg) = \deg(1) = 0.$$

Since $\deg(f), \deg(g) \geqslant 0$ this implies that $\deg(f) = \deg(g) = 0$ and hence $f(x), g(x)$ are nonzero constants, as desired.

Units are important for the theory of divisibility.

Definition of Association

For $a, b \in R$ in a ring we define the following notation:

$$a \sim b \quad \Longleftrightarrow \quad \text{there exists a unit } u \in R^\times \text{ such that } au = b.$$

Again, the symbol "$a \sim b$" represents a whole sentence. It says that "a is associate to b". You will check on the homework that this is an equivalence relation on the set R.

If R is an integral domain, then I claim that[18]

$$a \sim b \quad \Longleftrightarrow \quad a|b \text{ and } b|a.$$

Indeed, suppose that $a \sim b$ so that $au = b$ for some unit $u \in R^\times$. The equation $au = b$ implies that $a|b$ and the equation $bu^{-1} = a$ implies that $b|a$. Conversely, suppose that $a|b$ and $b|a$. By definition this means that $ak = b$ and $b\ell = a$ for some $k, \ell \in R$. Since $a \neq 0$ and since R is an integral domain, we have

$$\begin{aligned} b\ell &= a \\ ak\ell &= a \\ ak\ell - a &= 0 \end{aligned}$$

$$\begin{aligned} a(k\ell - 1) &= 0 \\ k\ell - 1 &= 0 \\ k\ell &= 1. \end{aligned}$$

This implies that $k, \ell \in R^\times$ and hence $a \sim b$.

For example, if $a, b \in \mathbb{Z}$ then since $\mathbb{Z}^\times = \{\pm 1\}$ we have $a \sim b$ if and only if $a = \pm b$. Hence

$$a|b \text{ and } b|a \text{ in } \mathbb{Z} \quad \Longleftrightarrow \quad a = \pm b.$$

And for nonzero polynomials $f(x), g(x) \in \mathbb{F}[x]$ we have

$$f(x)|g(x) \text{ and } g(x)|f(x) \text{ in } \mathbb{F}[x] \quad \Longleftrightarrow \quad f(x) = \lambda g(x) \text{ for some nonzero } \lambda \in \mathbb{F}.$$

There is one final property that the rings $\mathbb{Z}$ and $\mathbb{F}[x]$ have in common. Each of them has a notion of "division with remainder". The following definition is a little bit non-standard but it suffices for our purposes.[19]

Definition of Euclidean Domains

Let $(R, +, \cdot, 0, 1)$ be a ring. We say that R is a *Euclidean domain* if there exists a "size function" $N : R\backslash\{0\} \to \mathbb{N}$ satisfying the following two properties:

- For all nonzero $a, b \in R$ with $a|b$ we have $N(a) \leqslant N(b)$.
- For all $a, b \in R$ with $b \neq 0$, there exist some $q, r \in R$ (called quotient and remainder) satisfying the following two properties:

$$\begin{cases} a = bq + r, \\ r = 0 \text{ or } N(r) < N(b). \end{cases}$$

For example, we have already seen that the ring of integers $\mathbb{Z}$ with the size function $N(a) = |a|$ is a Euclidean domain. Indeed, to see that this N satisfies the desired property, consider some nonzero $a, b \in \mathbb{Z}$ with $a|b$. Since $b \neq 0$ this means that $ak = b$ for some nonzero k. Since k is nonzero we have $|k| \geqslant 1$ and then we multiply both sides of this inequality by the positive integer $|a|$ to obtain

$$1 \leqslant |k|$$

[18] Let us assume that a, b are both nonzero.

[19] Actually the concept of Euclidean domain is a bit awkward. The more elegant concept is a *principal ideal domain*, but we are not yet ready for that level of abstraction.

$$\begin{aligned}|a| &\leqslant |a||k| \\ |a| &\leqslant |ak| \\ |a| &\leqslant |b|.\end{aligned}$$

We have also seen that the ring of polynomials $\mathbb{F}[x]$ with size function $N(f) = \deg(f)$ is a Euclidean domain. Indeed, to see that this N satisfies the desired property, consider some nonzero $f(x), g(x) \in \mathbb{F}[x]$ with $f(x)|g(x)$. Since $g(x) \neq 0$ this means that $f(x)h(x) = g(x)$ for some nonzero $h(x)$. Then since f, g, h are all nonzero we have

$$\deg(f) \leqslant \deg(f) + \deg(h) = \deg(fh) = \deg(g).$$

Let me observe, however, that the abstract definition above is more compatible with $\mathbb{F}[x]$ than it is with $\mathbb{Z}$. Indeed, the usual statement of the division theorem for $\mathbb{Z}$ says that for all $a, b \in \mathbb{Z}$ with $b \neq 0$ there exist $q, r \in \mathbb{Z}$ with

$$\begin{cases} a = bq + r, \\ 0 \leqslant r \leqslant |b|. \end{cases}$$

This is not quite the same as saying that $r = 0$ or $|r| < |b|$ because it also includes the requirement that $r \geqslant 0$. But it makes no sense to say that $r \geqslant 0$ in a general Euclidean domain because the elements of a ring need not be ordered. For example, the elements of $\mathbb{F}[x]$ are **not** ordered; it makes no sense to say that $6x + 5 \geqslant 5x + 6$, or the other way around.

For this reason, quotients and remainders in a general Euclidean domain need not be unique. Luckily, we don't need them to be. Our purpose for defining Euclidean domains is to prove that every Euclidean domain has "unique prime factorization". For example, the integer 60 can be factored into prime integers in essentially only one way:

$$\begin{aligned}60 &= 2 \cdot 2 \cdot 3 \cdot 5 \\ &= 3 \cdot 2 \cdot 5 \cdot 2 \cdot 1 \cdot 1 \\ &= (-3) \cdot (-5) \cdot 2 \cdot 2 \\ &= \text{etc.}\end{aligned}$$

We can rearrange the factors and we can insert copies of 1 and -1 as we please, but this does not change the fact that there are "two copies of 2, one copy of 3 and one copy of 5". We will see that polynomials over a field also have unique prime factorization. For example, the polynomial $x^2 - 4 \in \mathbb{Q}[x]$ can be factored as

$$x^2 - 4 = (x-2)(x+2) = (-x+2)(-x-2) = (3x+6)\left(\frac{1}{3}x - \frac{2}{3}\right) = \text{etc.}$$

This time the prime factors are unique up to multiplication by nonzero constants, which are the units in the ring. Finally, let me note that the notion of "prime polynomial"[20] is relative

[20] The term "irreducible polynomial" is more common. This might come from the study of the ring $\mathbb{Z}[x]$, where we must distinguish between prime polynomials and prime coefficients.

to the field of coefficients. For example, the polynomial $x^2 - 2$ is prime as an element of $\mathbb{Q}[x]$ but it is not prime as an element of $\mathbb{R}[x]$ because $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$.

We will make all of this precise below.

3.2 The Euclidean Algorithm

In the pursuit of unique prime factorization we must first discuss greatest common divisors.

Definition of Greatest Common Divisors

Let R be a Euclidean domain with size function $N : R\backslash\{0\} \to \mathbb{N}$. For any two nonzero elements $a, b \in R$ we consider their set of *common divisors*

$$\text{Div}(a, b) = \{d \in R : d|a \text{ and } d|b\}.$$

We note that every common divisor d satisfies $N(d) \leqslant \min\{N(a), N(b)\}$ because $d|a$ implies that $dk = a$ for some k and hence $N(d) \leqslant N(dk) = N(a)$. Similarly, $N(d) \leqslant N(b)$.

Since the sizes of common divisors of a, b are bounded above by $\min\{N(a), N(b)\}$ it follows from the well-ordering property of the integers that there exist elements in $\text{Div}(a, b)$ of maximum size. Any such element will be called a *greatest common divisor* of a, b.

For example: Consider the set of common divisors of the integers 12 and 30:

$$\text{Div}(12, 30) = \{1, 2, 3, 6, -1, -2, -3, -6\}.$$

Thus, in this case, we have two greatest common divisors: 6 and -6.

More generally, we will prove below that any two greatest common divisors are associates. In the case of our two favorite Euclidean domains $\mathbb{Z}$ and $\mathbb{F}[x]$ this will allow us to make a further choice and to speak of **the** greatest common divisor.

Since the units of $\mathbb{Z}$ are ± 1, there will be exactly two greatest common divisors, and we will choose the positive one. Thus, for any nonzero integers $a, b \in \mathbb{Z}$ we define

$$\gcd(a, b) = \text{ the unique } \textbf{positive} \text{ common divisor of maximum absolute value.}$$

Since the units of $\mathbb{F}[x]$ are the nonzero constants, we can always scale our greatest common divisor so that the leading coefficient equals 1. [Jargon: A polynomial with leading coefficient 1 is called *monic*.] Thus, for any nonzero $f(x), g(x) \in \mathbb{F}[x]$ we define

$$\gcd(f, g) = \text{ the unique } \textbf{monic} \text{ common divisor of maximum degree.}$$

How can we prove that any two greatest common divisors are associate? We will do this by giving an algorithm to compute all of the elements of the set $\text{Div}(a, b)$. The proof that the algorithm works will involve the following lemmas.

Lemmas for the Euclidean Algorithm

(1) Let R be any ring and let $a, b, c, x \in R$ be elements satisfying $a = bx + c$. Then we have the following equality of sets:

$$\text{Div}(a, b) = \text{Div}(b, c).$$

(2) Let $a \in R$ be a nonzero element of a Euclidean domain. Since every element of R is a divisor, the common divisors of a and 0 are just the divisors of a:

$$\text{Div}(a, 0) = \text{Div}(a) = \{d \in R : d|a\}.$$

I claim that the maximum-sized divisors of a are exactly the associates of a.

Here is the algorithm.

The Euclidean Algorithm

Let R be a Euclidean domain with size function $N : R\backslash\{0\} \to \mathbb{N}$. For any nonzero $a, b \in R$, I claim that there exists a nonzero element $d \in R$ such that the common divisors of a and b are the same as the divisors of d:

$$\text{Div}(a, b) = \text{Div}(d).$$

Since these two sets are equal, their maximum-sized elements are the same. It then follows from Lemma (2) that any two greatest common divisors of a and b are associate to d, hence associate to each other.

To prove that such an element $d \in R$ exists we will actually give an efficient algorithm to compute it. To begin, we set $r_0 = b$ and then divide a by r_0 to obtain

$$a = r_0 q_1 + r_1, \quad \text{with } r_1 = 0 \text{ or } N(r_1) < N(r_0).$$

If $r_1 = 0$ then the algorithm stops. Otherwise, we divide r_0 by r_1 to obtain

$$r_0 = r_1 q_2 + r_2, \quad \text{with } r_2 = 0 \text{ or } N(r_2) < N(r_1).$$

If $r_2 = 0$ then the algorithm stops. Otherwise, we continue in the same fashion, to produce a sequence of nonzero remainders satisfying

$$N(r_0) > N(r_1) > N(r_2) > \cdots .$$

This process cannot continue forever because there cannot be an infinite decreasing sequence of non-negative integers. Hence there exists some index $n \geqslant 0$ such that $r_n \neq 0$ and $r_{n+1} = 0$. I claim that this r_n is the desired element d. Indeed, by repeated application of Lemma (1) we have

$$\mathrm{Div}(a, b) = \mathrm{Div}(a, r_0) = \mathrm{Div}(r_0, r_1) = \mathrm{Div}(r_1, r_2) = \cdots = \mathrm{Div}(r_n, 0) = \mathrm{Div}(r_n).$$

To summarize: If R is a Euclidean domain then we have shown that the greatest common divisor of two elements $a, b \in R$ is well-defined up to multiplication by units. Furthermore, we have given an algorithm to compute this greatest common divisor. If $N(a) \geqslant N(b)$ then Lamé's Theorem (which we will not prove) says that the algorithm takes no more than $5d + 2$ steps, where d is the number of decimal digits in $N(b)$. That's pretty fast.

3.3 The Extended Euclidean Algorithm

Let R be a Euclidean domain. In the last section we defined the greatest common divisor of two elements $a, b \in R$ (which we proved is unique up to multiplication by units) as the common divisor of maximum size. But you may see other definitions in the literature. Here we list three equivalent definitions.

I REGRET DOING IT THIS WAY. I SHOULD FIRST PROVE THAT A EUCLIDEAN DOMAIN IS A PID AND THEN DEFINE THE GCD FROM THERE.

Show that $aR + bR = dR$ for some d, which is unique up to multiplication by units. Show that d is a common divisor. Write $d = ax + by$ and $a = dq + r$ with $r = 0$ or $N(r) < N(d)$. If $r \neq 0$ then $r = a - dq \in aR + bR = dR$ so $N(d) \leqslant N(r)$. Contradiction. Hence d is a common divisor of a, b. Furthermore, if $e|a$ and $e|b$ then $e|d$ and hence $N(e) \leqslant N(d)$. It follows that d is a greatest common divisor.

Conversely, if e is a greatest common divisor of a, b then $e|d$ and hence $N(e) \leqslant N(d)$ and since d is a common divisor we have $N(d) \leqslant N(e)$, hence $N(d) = N(e)$. Finally, since $e|d$ and $N(e) = N(d)$ we have $d \sim e$.

Three Equivalent Definitions of GCD

Let R be a Euclidean domain with size function $N : R\backslash\{0\} \to \mathbb{N}$ and consider two nonzero elements $a, b \in R$. I claim that the following three definitions of *greatest common divisor* are equivalent:

(1) *A maximum-sized common divisor.* To be precise, consider the set $\mathrm{Div}(a,b)$ of common divisors. Then d is a greatest common divisor if $d \in \mathrm{Div}(a,b)$ and if for any $e \in \mathrm{Div}(a,b)$ we have $N(e) \leqslant N(d)$.

(2) *A maximally-divisible common divisor.* To be precise, we say that d is a greatest common divisor if $d \in \mathrm{Div}(a,b)$ and if for any $e \in \mathrm{Div}(a,b)$ we have $e|d$.

(3) *A minimum-sized nonzero R-linear combination.* To be precise, for any $a \in R$ we define the set of multiples $aR = \{ax : x \in R\}$ and for any two elements $a,b \in R$ we define the set of linear combinations:

$$aR + bR = \{ax + by : x, y \in R\}.$$

Note that $0 \in aR + bR$. We say that $d \neq 0$ is a greatest common divisor if $d \in aR + bR$ and if for all $e \in aR + bR$ we have $N(d) \leqslant N(e)$. This last definition is the least intuitive but it generalizes more naturally to rings that are not Euclidean.

The proof that these three definitions are equivalent will involve a modification of the Euclidean algorithm. In the original statement of the Euclidean algorithm we completely ignored the sequence of quotients $q_1, q_2, \ldots$. This time we will keep track of the information that is contained in the quotients.

Before presenting the general theorem I will give an example from the ring of integers. First we compute the greatest common divisor of 3094 and 2513 using the standard Euclidean algorithm, as described in the previous section:

$$\begin{array}{rcllrl}
3094 & = & 2513 \cdot 1 & + & 581 & \\
2513 & = & 581 \cdot 4 & + & 189 & \\
581 & = & 189 \cdot 3 & + & 14 & \\
189 & = & 14 \cdot 13 & + & 7 & \\
14 & = & 7 \cdot 2 & + & 0 & \text{STOP}
\end{array}$$

Hence from the lemma in the previous section we have:

$$\begin{aligned}
\mathrm{Div}(3094, 2513) &= \mathrm{Div}(2513, 581) \\
&= \mathrm{Div}(581, 189) \\
&= \mathrm{Div}(189, 14) \\
&= \mathrm{Div}(14, 7) \\
&= \mathrm{Div}(7, 0) \\
&= \mathrm{Div}(7).
\end{aligned}$$

Since the set of common divisors of 3094 and 2513 is equal to the set of divisors of 7, we conclude that the greatest common divisors are ± 7 and we choose the positive one:

$$\gcd(3094, 2513) = 7.$$

But note that we have ignored the sequence of quotients: $1, 4, 3, 13, 2$. What information do these numbers contain? I claim that we can use them to find a solution $x, y \in \mathbb{Z}$ to the following equation:[21]

$$3094x + 2513y = 7.$$

In order to do this we first consider the more general equation $ax + by = z$. This equation has two obvious solutions $(x, y, z) = (1, 0, 3094)$ and $(x, y, z) = (0, 1, 2513)$. It also has the useful property that any linear combination of solutions is still a solution. To be precise, consider the following set of triples of integers:

$$V = \{(x, y, z) \in \mathbb{Z}^3 : 3094x + 2513y = z\} \subseteq \mathbb{Z}^3.$$

If $\mathbf{x} = (x, y, z)$ and $\mathbf{x}' = (x', y', z')$ are any two elements of V then for any integers $r, s \in \mathbb{Z}$ I claim that the linear combination

$$r\mathbf{x} + s\mathbf{x}' = r(x, y, z) + s(x', y', z') = (rx + sx', ry + sy', rz + sz')$$

is also in the set V.[22] Indeed, by assumption we have $ax + by = z$ and $ax' + by' = z'$, hence

$$a(rx + sx') + b(ry + sy') = r(ax + by) + s(ax' + by') = rz + sz'.$$

The goal is to begin with the basic triples $\mathbf{x}_1 = (1, 0, 3094)$ and $\mathbf{x}_2 = (0, 1, 2513)$ and then to perform $\mathbb{Z}$-linear combinations until we obtain a triple of the form $(x, y, 7)$ for some integers $x, y \in \mathbb{Z}$. The Euclidean algorithm guarantees that this is always possible, and the sequence of quotients $1, 4, 3, 13, 2$ tells us exactly which linear combinations to perform. We record the computation in tabular form:

x	y	z	$\mathbf{x}$
1	0	3094	$\mathbf{x}_1$
0	1	2513	$\mathbf{x}_2$
1	-1	581	$\mathbf{x}_3 = \mathbf{x}_1 - 1\mathbf{x}_2$
-4	5	189	$\mathbf{x}_4 = \mathbf{x}_2 - 4\mathbf{x}_3$
13	-16	14	$\mathbf{x}_5 = \mathbf{x}_3 - 3\mathbf{x}_4$
-173	213	7	$\mathbf{x}_6 = \mathbf{x}_4 - 13\mathbf{x}_5$
359	-442	0	$\mathbf{x}_7 = \mathbf{x}_5 - 2\mathbf{x}_6$

Note that the values of z are precisely the sequence of remainders from the Euclidean algorithm, thus we stop when we reach a remainder of 0. The final nonzero remainder is the greatest common divisor and reading off the corresponding values of x and y tells us that

$$3094(-173) + 2513(213) = 7,$$

which solves the desired equation. Here is the general theorem. This result is also sometimes called *Bézout's Identity.*

[21] It will become clear later why we **want** to solve this equation.

[22] Jargon: The set $\mathbb{Z}^3$ is not quite a vector space because $\mathbb{Z}$ is not a field. Instead we call it a $\mathbb{Z}$-*module*. The fact that $V \subseteq \mathbb{Z}^3$ is closed under $\mathbb{Z}$-linear combinations makes it a $\mathbb{Z}$-*submodule*.

The Extended Euclidean Algorithm

Let R be a Euclidean domain with size function $N : R\backslash\{0\} \to \mathbb{N}$. For any nonzero $a, b \in R$ we showed in the previous section that there exists a greatest common divisor $\gcd(a, b) \in R$, which is unique up to multiplication by units. I claim now that there exist (non-unique) elements $x, y \in R$ satisfying[23]

$$ax + by = \gcd(a, b).$$

To prove the existence of such x, y we will actually give an algorithm to compute them. First, consider the set of triples $(x, y, z) \in R^3$ satisfying $ax + by = z$:

$$V = \{(x, y, z) \in R^3 : ax + by = z\} \subseteq R^3.$$

This set is closed under R-linear combinations,[24] since for any vectors $\mathbf{x} = (x, y, z)$ and $\mathbf{x}' = (x', y', z')$ in V and for any elements $r, r' \in R$, the vector $r\mathbf{x} + r'\mathbf{x}' = (rx + r'x', ry + r'y', rz + r'z')$ is also in V:

$$a(rx + r'x') + b(ry + r'y') = r(ax + by) + r'(ax' + by') = rz + rz'.$$

Our goal is to start with the basic vectors $\mathbf{x}_1 = (1, 0, a)$ and $\mathbf{x}_2 = (0, 1, b)$ in V and to form R-linear combinations until we obtain a vector of the form $(x, y, \gcd(a, b)) \in V$, from which it will follow that $ax + by = \gcd(a, b)$. To do this, we consider the steps in the usual (non-vector) Euclidean Algorithm:

$$\begin{aligned}
a &= bq_1 + r_1, \\
b &= r_1 q_2 + r_2, \\
r_1 &= r_2 q_3 + r_3, \\
&\vdots \\
r_{i-2} &= r_{i-1} q_i + r_i, \\
&\vdots \\
r_{n-2} &= r_{n-1} q_n + r_n, \\
r_{n-1} &= r_n q_{n+1} + 0,
\end{aligned}$$

where $r_n = \gcd(a, b)$. If we recursively define the vector $\mathbf{x}_{i+2} = \mathbf{x}_i - q_i \mathbf{x}_{i+1}$ then it will follow that $\mathbf{x}_{n+2} = (x, y, r_n)$ for some $x, y \in R$. Indeed, if we assume for induction that $\mathbf{x}_i = (x', y', r_{i-2})$ and $\mathbf{x}_{i+1} = (x'', y'', r_{i-1})$ for some $x', y', x'', y'' \in R$ then it follows that

$$\mathbf{x}_{i+2} = \mathbf{x}_i - q_i \mathbf{x}_{i+1} = (x' - q_i x'', y' - q_i y'', r_{i-2} - q_i r_{i-1}) = (x, y, r_i)$$

for some $x, y \in R$, as desired. Anyway, that's how a computer does it. A human would find it more convenient to organize all of the computations in a table:

x	y	z
1	0	a
0	1	b
1	$-q_1$	r_1
$-q_2$	$1+q_1q_2$	r_2
$1+q_2q_3$	$-q_1-q_3-q_1q_2q_3$	r_3
$\vdots$	$\vdots$	$\vdots$
something	something	$\gcd(a,b)$

In summary, for any nonzero elements a,b of a Euclidean domain and for any of their greatest common divisors d, there exist some elements x,y satisfying

$$ax+by=d.$$

This innocuous looking result unlocks the theory of prime factorization, as we will discuss in the next section. For now, we can use it to prove the equivalence of the three definitions of GCD discussed at the beginning of this section.

Proof that (1)⇔(2). Let d be a "maximally-divisible" common divisor of a and b. That is, suppose that $d|a$ and $d|b$, and suppose that for all e satisfying $e|a$ and $e|b$ we must have $e|d$. In this case we want to show that d is a "maximum-sized" common divisor. This follows immediately since for any other common divisor e we must have $e|d$, which implies that $N(e)\leqslant N(d)$. Conversely, let d be a "maximum-sized" common divisor of a and b. In order to show that d is "maximally-divisible" let e be any other common divisor. Our goal is to show that $e|d$. To do this we must use the result of the Extended Euclidean Algorithm just discussed. It tells us that there exist $x,y\in R$ satisfying

$$ax+by=d.$$

Then since $e|a$ and $e|b$ we have $ek=a$ and $e\ell=b$ for some $k,\ell\in R$, which implies that

$$d=ax+by=ekx+e\ell y=e(kx+\ell y),$$

and hence $e|d$. □

The third equivalent definition has significant theoretical importance so we will isolate it as a theorem.

[23] It doesn't matter which GCD we choose since if d is some GCD satisfying $d=ax+by$ then any other GCD has the form du for some unit $u\in R^\times$, hence $du=a(xu)+b(yu)$ for some $xu,yu\in R$.

[24] If R were a field then R^3 would be a vector space and we would call $V\subseteq R^3$ a vector subspace. If R is not a field then we use the more general terms R-module and R-submodule.

Bézout's Identity

Let $a, b \in R$ be any two nonzero elements of a Euclidean domain and let $d \in R$ be their greatest common divisor. Then I claim that

$$aR + bR = dR.$$

To explain this notation, $dR = \{dr : r \in R\}$ is the set of multiples of d and $aR + bR = \{ar + bs : r, s \in R\}$ is the set of "R-linear combinations" of a and b.

To prove this we must show both inclusions. To see that $aR + bR \subseteq dR$, consider any element $ar + bs \in aR + bR$. Since d is a common divisor of a and b we have $dk = a$ and $d\ell = b$ for some $k, \ell \in R$ and it follows that

$$ar + bs = dkr + d\ell s = d(kr + \ell s),$$

so that $ar + bs$ is an element of dR. Conversely, to see that $dR \subseteq aR + bR$, consider any element $dr \in dR$. From the Extended Euclidean Algorithm there exist $x, y \in R$ satisfying $ax + by = d$. It follows that

$$dr = (ax + by)r = a(xr) + b(yr),$$

so that dr is an element of $aR + bR$.

Proof that (1) and (2) are equivalent to (3). Let d be any GCD of a, b in the sense of definition (1) or (2). Then from the basic Euclidean Algorithm we know that the set of all GCDs of a and b are just the associates of d, and from Bézout's Identity just proved we have

$$aR + bR = dR.$$

It remains to show that the minimum-sized nonzero elements of dR are precisely the associates of d.[25] First of all, we note that d itself is a minimum-sized element of dR since $d = d1 \in dR$ and since any element dr satisfies $N(d) \leqslant N(dr)$. This also shows that $N(d)$ is the minimum size of an element of dR. Next we observe that any associate $e \sim d$ is a minimum-sized element of dR. Indeed, suppose that $e \sim d$ so that $d = eu$ and $e = du^{-1}$ for some unit $u \in R^\times$. This implies that $d|e$ (in particular, $e \in dR$) and $e|d$. Then from properties of the size function we have $N(d) \leqslant N(e)$ and $N(e) \leqslant N(d)$, hence $N(e) = N(d)$. It only remains to show that any minimum-sized element of dR is associate to d. For this, let $m = dk \in dR$ be any multiple of d satisfying $N(m) = N(d)$. If we can prove that $m|d$ then it will follow from the usual proof[26]

[25] In other words, we need to show that the minimum-sized multiples of d are the associates of d. Compare this to our lemma for the Euclidean Algorithm which says that the maximum-sized divisors of d are the associates of d, which you will prove on the homework. Pay attention because the proofs are almost identical.

[26] If $d|m$ and $m|d$ then we have $dk = m$ and $m\ell = d$ for some k, ℓ, which implies $m(1 - k\ell) = 0$. Since $m \neq 0$ this implies that $1 - k\ell = 0$ so that k, ℓ are units.

that $m \sim d$. So let us divide d by m to obtain $q, r \in R$ satisfying

$$\begin{cases} d = mq + r, \\ r = 0 \text{ or } N(r) < N(m). \end{cases}$$

If $r \neq 0$ then we must have $N(r) < N(m)$. On the other hand, we know that $r = d - mq = d - dkq = d(1 - kq)$ so that $d|r$ and hence $N(r) \geqslant N(d) = N(m)$. This contradiction shows that $r = 0$ and hence $m|d$. □

We end this section by considering the special case when $\gcd(a, b) = 1$.

Definition of Coprime

Let R be a Euclidean domain. We say that nonzero elements $a, b \in R$ are *coprime* (or *relatively prime*) when 1 is a greatest common divisor, hence the units $R^\times$ are the set of common divisors. In this case it is convenient to write

$$\gcd(a, b) = 1,$$

even though the GCD is not generally unique. If a, b are coprime then it follows from the Extended Euclidean Algorithm that we have

$$ax + by = 1$$

for some $x, y \in R$. Conversely, if such x, y exist then I claim that a, b are coprime. Indeed, suppose that $ax + by = 1$ and let d be any common divisor of a and b, so that $dk = a$ and $d\ell = b$ for some $k, \ell \in R$. It follows that

$$1 = ax + by = dkx + d\ell y = d(kx + \ell y),$$

and hence $d|1$. But the divisors of 1 are precisely the units.

3.4 Unique Prime Factorization

The previous section was fairly technical. The key result was the existence for any nonzero $a, b \in R$ in a Euclidean domain of elements $x, y \in R$ satisfying

$$ax + by = \gcd(a, b).$$

In this section we will exploit this result to prove the important *Fundamental Theorem of Arithmetic*, which says that elements of a Euclidean domain have “unique prime factorization”. Before stating the result we must define the word “prime”.

Definition of Prime

Recall that a positive integer $p \geqslant 2$ is called prime when its only positive divisors are 1 and itself. In a general Euclidean domain R we say that a nonzero, nonunit element $p \in R$ is *prime* when its only divisors are units and the associates of p. In other words:

$$d|p \quad \Longrightarrow \quad d \sim 1 \text{ or } d \sim p.$$

Let me also record a useful property of this definition. If a nonunit, nonzero element $a \in R$ is **not prime** then by definition it can be expressed as

$$a = bc \quad \text{where } b, c \text{ are not units and not associate to } a.$$

Applying the size function gives $N(b) \leqslant N(a)$ and $N(c) \leqslant N(a)$. But you will show on the homework that the maximum-sized divisors of a are the associates of a, hence in this situation we must have $N(b) < N(a)$ and $N(c) < N(a)$.

The reason for saying that units are not prime is purely conventional.[27] We do this so that factorization into primes will be unique. Indeed, the following factorizations of 60 should be considered the same:

$$60 = 2 \cdot 2 \cdot 3 \cdot 5 = 2 \cdot 2 \cdot 3 \cdot 5 \cdot 1 = 2 \cdot 2 \cdot 3 \cdot 5 \cdot 1 \cdot 1 = 2 \cdot 2 \cdot 3 \cdot 5 \cdot 1 \cdot 1 \cdot 1 \cdots .$$

We should also consider prime factorizations to be the same if they differ by rearranging the terms or inserting an even number of negative signs:

$$\begin{aligned} 60 &= 2 \cdot 2 \cdot 3 \cdot 5 \\ &= 3 \cdot 2 \cdot 5 \cdot 2 \\ &= (-3)(-2) \cdot 5 \cdot 2 \\ &= (-1) \cdot 5 \cdot 2 \cdot (-3) \cdot 2 \\ &= \text{etc.} \end{aligned}$$

The following theorem is sometimes called the *Fundamental Theorem of Arithmetic.*

Unique Prime Factorization

Let $a \in R$ be a nonzero, nonunit element of a Euclidean domain. Then:

(1) We can express a as a product of prime elements.

(2) The prime factors are unique up to permutations and multiplication by units.

[27] The reason for saying that 0 is not prime is more subtle and we won't discuss this.

In other words, in a Euclidean domain there is a concept of *prime multiplicity*. Given a prime element $p \in R$ there is a well-defined function $\nu_p : R\backslash\{0\} \to \mathbb{N}$ such that $\nu_p(a)$ is the multiplicity of the prime p in the factorization of a. For example, we have

$$\begin{aligned}\nu_2(60) &= 2,\\ \nu_3(60) &= 1,\\ \nu_5(60) &= 1,\\ \nu_7(60) &= 0.\end{aligned}$$

By convention we will also define $\nu_p(u) = 0$ for all primes p and units u.

Proof of (1). We will use induction on the size of a. If a is prime then we are done. Otherwise from the remarks above we can write $a = bc$ with $N(b) < N(a)$ and $N(c) < N(a)$. Since b and c are strictly smaller than a we can assume that each is a product of primes. Hence a is also a product of primes. □

For the proof of uniqueness we need the following famous lemma.

Euclid's Lemma

Let $p \in R$ be a prime element of a Euclidean domain. Then for all $a, b \in R$ we have

$$p|ab \quad\Longrightarrow\quad p|a \text{ or } p|b.$$

The proof is classic and it makes a good exam problem. If $p|(ab)$ and $p \nmid a$ then we will show that $p|b$. To do this we first observe that $\gcd(a,p) = 1$. Indeed, let d be any common divisor of a and p. Since $d|p$ and p is prime we must have $d \sim 1$ or $d \sim p$. But if $d \sim p$ then since $d|a$ we would have $p|a$. Contradiction. It follows that $d \sim 1$, hence the only common divisors of a and p are the units. In other words, we have $\gcd(a,p) = 1$, hence the Extended Euclidean Algorithm tells us that there exist $x, y \in R$ satisfying

$$ax + py = 1.$$

Now the trick is to multiply both sides by b and use the fact that $p|(ab)$ to write $ab = pk$ for some $k \in R$:

$$\begin{aligned}ax + py &= 1\\ abx + pby &= b\\ pkx + pby &= b\end{aligned}$$

$$p(kx + by) = b.$$

We conclude that $p|b$ as desired.

The hypothesis that p be prime is necessary. For example, we have $4|(6 \cdot 10)$ but $4 \nmid 6$ and $4 \nmid 10$. Now here is the proof of uniqueness.

Proof of Uniqueness. Suppose that we have

$$p_1 p_2 \cdots p_k = u q_1 q_2 \cdots q_\ell$$

for some prime elements $p_1, \ldots, p_k, q_1, \ldots, q_\ell \in R$ and unit $u \in R^\times$. In this case I claim that $k = \ell$ and that we can rearrange the factors so that $p_1 \sim q_1$, $p_2 \sim q_2$, $\ldots$, $p_k \sim q_k$. To see this we observe that p_1 divides the left hand side, so it also divides the right hand side:

$$p_1 | (q_1 q_2 \cdots q_\ell).$$

By applying induction to Euclid's Lemma we must have $p_1 | q_i$ for some i. After rearranging the factors if necessary we may assume that $p_1 | q_1$. Since q_1 is prime this implies that $p_1 \sim 1$ or $p_1 \sim q_1$. But $p_1 \sim 1$ is impossible because p_1, being prime, is not a unit. Hence we must have $p_1 \sim q_1$ so that $p_1 = u' q_1$ for some unit $u' \in R^\times$. Finally, we cancel p_1 from both sides:

$$\begin{aligned} p_1 p_2 \cdots p_k &= u q_1 q_2 \cdots q_\ell \\ p_1 p_2 \cdots p_k &= u u' p_1 q_2 \cdots q_\ell \\ p_2 \cdots p_k &= u u' q_2 \cdots q_\ell. \end{aligned}$$

And the result follows by induction. □

All of these ideas were implicit in Euclid's *Elements*, Book X. The explicit proof was first written down by Gauss in the case of integers. Simon Stevin was the first to observe that the same arguments apply to factorization of polynomials.

3.5 Irreducible Polynomials

Prime factorization in the ring $\mathbb{Z}$ is a familiar concept. However, since $\mathbb{F}[x]$ is also a Euclidean domain, the previous theorem also tells us that polynomials have unique prime factorization. You should be aware, however, that prime elements of the ring $\mathbb{F}[x]$ are more commonly called *irreducible polynomials.*

Definition of Irreducible Polynomials

Let $f(x)$ be a nonzero, nonconstant polynomial with coefficients in a field $\mathbb{F}$. We say that

$f(x)$ is *irreducible over* $\mathbb{F}$ if for all polynomials $g(x), h(x)$ with coefficients in $\mathbb{F}$ we have

$$f(x) = g(x)h(x) \quad \Longrightarrow \quad g(x) \text{ or } h(x) \text{ is constant.}$$

Note that we say "irreducible over $\mathbb{F}$" instead of just "irreducible". For example, the polynomial $x^2 + 1$ is **reducible** (i.e., not irreducible) over $\mathbb{C}$ because

$$x^2 + 1 = (x - i)(x + i).$$

However, I claim that x^2+1 is **irreducible** over $\mathbb{R}$. To see this, let us suppose for contradiction that $x^2 + 1 = g(x)h(x)$ for some nonconstant polynomials $g(x), h(x)$ with real coefficients. Taking degrees gives

$$2 = \deg(x^2 + 1) = \deg(g) + \deg(h),$$

which since $g(x), h(x)$ are nonconstant implies that $\deg(g) = \deg(h) = 1$. In particular, this tells us that $g(x) = ax + b$ for some real $a, b \in \mathbb{R}$ with $a \neq 0$, which implies that $-b/a \in \mathbb{R}$ is a real root of $x^2 + 1$ because

$$(-b/a)^2 + 1 = (a(-b/a) + b)\, h(-b/a) = 0 \cdot h(-b/a) = 0.$$

But we know that the polynomial $x^2 + 1$ has **no real roots** because any real number $\alpha \in \mathbb{R}$ satisfies $\alpha^2 \geqslant 0$ and hence $\alpha^2 + 1 \geqslant 1$.

These observations are quite useful so we record them as a theorem.

Irreducible Polynomials of Small Degree

Let $f(x)$ be a polynomial with coefficients in a field $\mathbb{F}$.

(1) If $\deg(f) = 1$ then $f(x)$ is irreducible over any field containing $\mathbb{F}$.

(2) If $\deg(f) = 2$ or 3 then I claim that

$$f(x) \text{ is reducible over } \mathbb{F} \quad \Longleftrightarrow \quad f(x) \text{ has a root in } \mathbb{F}.$$

To prove (1), suppose for contradiction that $\deg(f) = 1$ and that $f(x) = g(x)h(x)$ for some nonconstant $g(x), h(x)$ with roots in a field containing $\mathbb{F}$. Then taking degrees gives a contradiction:

$$1 = \deg(f) = \deg(g) + \deg(h) \geqslant 1 + 1 = 2.$$

To prove one direction of (2), let us suppose that $f(a)$ for some $a \in \mathbb{F}$. Then from Descartes' Theorem we have $f(x) = (x-a)g(x)$ for some $g(x) \in \mathbb{F}[x]$ of degree $\deg(f)-1$. Since $\deg(f) \geqslant 2$ this polynomial $g(x)$ is nonconstant and we conclude that $f(x)$ is reducible over $\mathbb{F}$, as desired. For the other direction of (2), let us suppose that $f(x)$ is reducible over $\mathbb{F}$, so that $f(x) = g(x)h(x)$ for some nonconstant $g(x), h(x)$ with coefficients in $\mathbb{F}$. Taking degrees gives

$$\deg(g) + \deg(h) = \deg(f) = 2 \text{ or } 3.$$

Since $\deg(g), \deg(h) \geqslant 1$ this implies that we must have $\deg(g) = 1$ or $\deg(h) = 1$. Without loss of generality, suppose that $\deg(g) = 1$, so that $g(x) = ax + b$ for some $a, b \in \mathbb{F}$ with $a \neq 0$. Then it follows that $-b/a \in \mathbb{F}$ is a root of $f(x)$:

$$f(-b/a) = (a(-b/a) + b)\, h(-b/a) = 0 \cdot h(-b/a) = 0.$$

For example, we have already discussed the prime factorization of $x^n - 1$ over $\mathbb{C}$:[28]

$$x^n - 1 = (x-1)(x-\omega)(x-\omega^2)\cdots(x-\omega^{n-1}),$$

And over $\mathbb{R}$:

$$x^n - 1 = \begin{cases} (x-1)\prod_{k=1}^{(n-1)/2}(x^2 - 2\cos(2\pi k/n)x + 1) & \text{if } n \text{ is odd,} \\ (x-1)(x+1)\prod_{k=1}^{(n-2)/2}(x^2 - 2\cos(2\pi k/n)x + 1) & \text{if } n \text{ is even.} \end{cases}$$

Indeed, for any integer $k \in \mathbb{Z}$ such that ω^k is not real, its complex conjugate ω^{-k} is also not real. It follows that the quadratic polynomial

$$(x - \omega^k)(x - \omega^{-k}) = x^2 - 2\cos(2\pi k/n) + 1$$

has no real roots, hence is irreducible over $\mathbb{R}$.

But this criterion does not work for polynomials of degree $\geqslant 4$. For example, we have seen that the polynomial $x^4 + 4$ has no real roots. Nevertheless, it is reducible over $\mathbb{R}$:

$$x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2).$$

In general it is quite difficult to prove that a given polynomial is irreducible. To give a taste of things to come, I will just show you the prime factorizations of $x^n - 1$ over $\mathbb{Q}$ for the first several values of n:

$$\begin{aligned} x^2 - 1 &= (x-1)(x+1) \\ x^3 - 1 &= (x-1)(x^2 + x + 1) \\ x^4 - 1 &= (x-1)(x+1)(x^2+1) \\ x^5 - 1 &= (x-1)(x^4 + x^3 + x^2 + x + 1) \\ x^6 - 1 &= (x-1)(x+1)(x^2 + x + 1)(x^2 - x + 1) \\ x^7 - 1 &= (x-1)(x^6 + x^5 + x^4 + x^3 + x^2 + 1) \\ x^8 - 1 &= (x-1)(x+1)(x^2+1)(x^4+1) \\ x^9 - 1 &= (x-1)(x^2 + x + 1)(x^6 + x^3 + 1). \end{aligned}$$

Do you notice any patterns here?

[28] Here we take $\omega = e^{2\pi i/n}$.

4 Some Number Theory

4.1 Modular Arithmetic

Before returning to the theory of polynomials in the next chapter, we pause to examine some consequences of unique prime factorization in the ring of integers. Some of this material was developed in the homework.

Definition of Equivalence Relations

Let S be a set. A *relation* on S is just a subset of the cartesian product set:

$$\mathscr{R} \subseteq S \times S = \{(x,y) : a, b \in S\}.$$

However, instead of writing $(x,y) \in \mathscr{R}$ we will write $x\mathscr{R}y$, "x is related to y" by $\mathscr{R}$. We will say that $\mathscr{R}$ is an *equivalence relation* when it satisfies the following three properties:

- $\forall x \in S, x\mathscr{R}x$ (reflexive)
- $\forall x, y \in S, x\mathscr{R}y$ implies $y\mathscr{R}x$ (symmetric)
- $\forall x, y, z \in S, x\mathscr{R}y$ and $y\mathscr{R}z$ imply $x\mathscr{R}y$ (transitive)

In this case will use a symbol such as $\sim$, $\simeq$, $\approx$, $\cong$ or $\equiv$ to emphasize that $\mathscr{R}$ behaves like an equals sign.

We have already seen one equivalence relation in this course. For elements $a, b \in R$ in a ring R we have defined the relation of *association*:

$$a \sim b \quad \Longleftrightarrow \quad \exists u \in R^\times, au = b.$$

Let us verify that this is, indeed, an equivalence:

- **Reflexive.** Since 1 is a unit we have $a1 = a$ and hence $a \sim a$.
- **Symmetric.** Suppose that $a \sim b$ so that $au = b$ for some unit $u \in R^\times$. By definition this means that u has a multiplicative inverse u^{-1}, so that $bu^{-1} = a$. Since the element u^{-1} is also a unit this implies that $b \sim a$.
- **Transitive.** Suppose that $a \sim b$ and $b \sim c$ so that $au = b$ and $bv = c$ for some units $u, v \in R^\times$. By definition this means that u and v have multiplicative inverses u^{-1} and v^{-1}. But then the product uv is also a unit with $(uv)^{-1} = u^{-1}v^{-1}$. Then since $a(uv) = (au)v = bv = c$ we conclude that $a \sim c$ as desired.

The next concept was introduced by Gauss in his *Disquisitiones Arithmeticae* (1801). We still use the same notation as he did.

Definition of Congruence Modulo and Integer

Fix an integer $n \geqslant 1$. Then for all integers $a, b \in \mathbb{Z}$ we define the following notation:

$$a \equiv b \mod n \quad \Longleftrightarrow \quad n|(a-b).$$

In this case we say that *a is congruent to b modulo n*. Let us verify that this is an equivalence relation on the set $\mathbb{Z}$:

- **Reflexive.** Since $n0 = a - a$ we have $n|(a-a)$ and hence $a \equiv a \bmod n$.
- **Symmetric.** Let $a \equiv b \bmod n$ so that $n|(a-b)$ and hence $a - b = nk$ for some $k \in \mathbb{Z}$. Then we have $b - a = n(-k)$ so that $n|(b-a)$ and hence $b \equiv a \bmod n$.
- **Transitive.** Let $a \equiv b \bmod n$ and $b \equiv c \bmod n$ so that $a - b = nk$ and $b - c = n\ell$ for some integers $k, \ell \in \mathbb{Z}$. Then we have

  $$a - c = (a-b) + (b-a) = nk + n\ell = n(k+\ell),$$

  so that $n|(a-c)$ and hence $a \equiv c \bmod n$.

The main reason for defining this relation is that it behaves well with respect to addition and multiplication of integers. To be precise, let us suppose that $a \equiv a' \bmod n$ and $b \equiv b' \bmod n$, so that $a - a' = nk$ and $b - b' = n\ell$ for some integers $k, \ell \in \mathbb{Z}$. Then we have

$$[(a+b) - (a'+b')] = (a-a') + (b-b') = nk + n\ell = n(k+\ell),$$

which implies that $a + b \equiv a' + b' \bmod n$, and we have

$$ab - a'b' = ab - ab' + ab' - a'b' = a(b-b') + (a-a')b' = an\ell + nkb' = n(a\ell + kb'),$$

which implies that $ab \equiv a'b' \bmod n$. This just means that we can perform arithmetic using the symbol $\equiv$ instead of $=$ and we won't get into trouble. For example, since $3 \equiv 13$ and $4 \equiv -6 \bmod 10$, we should also have $3 \cdot 4 \equiv 13 \cdot (-6) \bmod 10$. And, indeed,

$$13 \cdot (-6) \equiv -78 \equiv 2 \equiv 12 \equiv 3 \cdot 4 \mod 10.$$

We can use these operations to define a new family of finite rings.

The Ring $\mathbb{Z}/n\mathbb{Z}$ (i.e., Modular Arithmetic)

Fix an integer $n \geqslant 1$. I claim that every integer $a \in \mathbb{Z}$ is congruent mod n to a unique integer r in the set $\{0, 1, \dots, n-1\}$. Indeed, dividing a by n gives some $q, r \in \mathbb{Z}$ satisfying

$$\begin{cases} a = nq + r, \\ 0 \leqslant r < n, \end{cases}$$

and hence $a \equiv nq + r \equiv n0 + r \equiv r \bmod n$. To see that this integer r is unique, suppose that we have $a \equiv r \equiv r' \bmod n$ for some integers r, r' in the set $\{0, 1, \dots, n-1\}$. Our goal is to show that $r = r'$. First we observe that $r - r' \equiv a - a \equiv 0 \bmod n$, so that $n|(r - r')$. Now let us assume for contradiction that $r \neq r'$. Without loss of generality we can assume that $r' < r$ and hence $r - r' > 0$. But then the condition $n|(r - r')$ implies $n \leqslant r - r'$ and we obtain the desired contradiction:

$$r < n \leqslant r - r' \leqslant r.$$

In summary, we can define a ring structure on the finite set

$$\mathbb{Z}/n\mathbb{Z} = \{0, 1, 2, \dots, n-1\}.$$

The ring operations are addition and multiplication mod n and the special elements are 0 and 1. It is boring to check that the eight ring axioms are satisfied so we won't bother.

Remark: The theorem that every $a \in \mathbb{Z}$ is congruent mod n to a unique integer r in the set $\{0, 1, \dots, n-1\}$ is equivalent to the existence and uniqueness of remainders in the ring $\mathbb{Z}$. We previously proved the existence but we did not prove the uniqueness until now. Thus we could view $\mathbb{Z}/n\mathbb{Z} = \{0, 1, \dots, n-1\}$ as the set of possible remainders mod n. For this reason, the ring structure of $\mathbb{Z}/n\mathbb{Z}$ is sometimes called the *arithmetic of remainders*. More commonly it is called *modular arithmetic*.

4.2 Some Finite Fields

In the previous section we defined a family of finite rings $\mathbb{Z}/n\mathbb{Z}$, one for each positive integer $n \geqslant 1$. For example, here are the addition and multiplication tables for the ring $\mathbb{Z}/6\mathbb{Z}$:

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

The following identities are quite interesting:

$$2 \cdot 3 \equiv 3 \cdot 2 \equiv 4 \cdot 3 \equiv 3 \cdot 4 \equiv 0 \bmod 6.$$

They tell us that the ring $\mathbb{Z}/6\mathbb{Z}$ is not an integral domain, thus the theory developed in the previous chapter does not apply to it. The problem here is that the number 6 can be factored as $2 \cdot 3$. The situation is better for prime moduli.

The Ring $\mathbb{Z}/p\mathbb{Z}$ is a Field

Let $p \geqslant 2$ be a prime integer and consider the ring $\mathbb{Z}/p\mathbb{Z}$ of size p. Recall Euclid's Lemma, which says that

$$p|ab \quad \Longrightarrow \quad p|a \text{ or } p|b.$$

Since the statement $p|c$ is equivalent to $c \equiv 0 \bmod p$, this becomes

$$ab \equiv 0 \bmod p \quad \Longrightarrow \quad a \equiv 0 \bmod p \;\text{ or } b \equiv 0 \bmod p.$$

In other words, the ring $\mathbb{Z}/p\mathbb{Z}$ is an integral domain. You showed on a previous homework that every finite integral domain is a field. Let me reproduce the proof here. For any nonzero $a \in \mathbb{Z}/p\mathbb{Z}$ we consider the multiplication function $\mu_a : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ defined by $\mu_a(b) = ab$. Since $\mathbb{Z}/p\mathbb{Z}$ is an integral domain this function is injective:[29]

$$\begin{aligned} \mu_a(b) &\equiv \mu_a(c) \\ ab &\equiv ac \\ a(b-c) &\equiv 0 \\ (b-c) &\equiv 0 \\ b &\equiv c. \end{aligned}$$

But any injective function from a finite set to itself must also be surjective. Hence the element $1 \in \mathbb{Z}/p\mathbb{Z}$ is expressible as $\mu_a(b)$ for some $b \in \mathbb{Z}/p\mathbb{Z}$. In other words, each nonzero element $a \in \mathbb{Z}/p\mathbb{Z}$ has a multiplicative inverse

$$\begin{aligned} \mu_a(b) &\equiv 1 \\ ab &\equiv 1 \\ a^{-1} &\equiv b. \end{aligned}$$

The proof above tells us that inverses *exist* in the ring $\mathbb{Z}/p\mathbb{Z}$ but it does not tell us how to find them. Since there are only finitely many possibilities we could always just check them all. For example, to find the inverse of 3 mod 7 we could just multiply 3 by every element of $\mathbb{Z}/7\mathbb{Z}$:

$$\begin{aligned} 3 \cdot 1 &\equiv 3 \\ 3 \cdot 2 &\equiv 6 \end{aligned}$$

[29] All congruences are mod p.

$$\begin{aligned} 3 \cdot 3 &\equiv 9 \equiv 2 \\ 3 \cdot 4 &\equiv 12 \equiv 5 \\ 3 \cdot 5 &\equiv 15 \equiv 1 \\ 3 \cdot 6 &\equiv 18 \equiv 4. \end{aligned}$$

We see that $3 \cdot 5 \equiv 1 \bmod 7$ and hence $3^{-1} \equiv 5 \bmod 7$. In the worst case scenario this method will use $p-1$ computations to find the inverse of a nonzero element of $\mathbb{Z}/p\mathbb{Z}$.

Luckily we can do much better.

Computing Inverses in $\mathbb{Z}/p\mathbb{Z}$

Let $p \geqslant 2$ be prime and consider a nonzero element $a \in \mathbb{Z}/p\mathbb{Z}$. In other words, consider an integer $a \in \mathbb{Z}$ such that $p \nmid a$. Since p is prime this implies that $\gcd(p, a) = 1$, hence we can use the Extended Euclidean Algorithm to find some integers $x, y \in \mathbb{Z}$ such that

$$px + ay = 1.$$

Then reducing both sides of this equation mod p gives

$$1 \equiv px + ay \equiv 0x + ay \equiv ay$$

and it follows that $a^{-1} \equiv y \bmod p$. For example, we compute $346^{-1} \bmod 1009$.[30] We consider the set of triples (x, y, z) satisfying $1009x + 346y = z$. Then starting with the easy triples $(1, 0, 1009)$ and $(0, 1, 346)$ we perform linear combinations until we obtain a triple of the form $(x, y, 1)$:[31]

x	y	z
1	0	1009
0	1	346
1	−2	317
−1	3	29
11	−32	27
−12	35	2
167	−487	1

We conclude that $1009(167) + 346(-487) = 1$. Reducing this equation mod 1009 gives

$$1 \equiv 1009(167) + 346(-487) \equiv 0(167) + 346(-487) \equiv 346(-487),$$

and hence

$$346^{-1} \equiv -487 \equiv 522 \bmod 1009.$$

Just to be sure, let's check:

$$346 \cdot 522 \equiv 180612 \equiv 1009 \cdot 179 + 1 \equiv 0 \cdot 179 + 1 \equiv 1 \bmod 1009.$$

Note that this method only used 5 steps. In general, the Extended Euclidean Algorithm uses less than $\log_2(a)$ steps to compute the inverse of a mod p.

The results of computations in $\mathbb{Z}/p\mathbb{Z}$ have "pseudorandom" behavior. Even though the algorithm is perfectly deterministic, the results seem to bounce around randomly. For example, if we change a just a little bit then its inverse may change by a lot:

$$\begin{aligned} 346^{-1} &\equiv 522 \\ 347^{-1} &\equiv 410 \\ 348^{-1} &\equiv 519 \\ 349^{-1} &\equiv 717 \\ 350^{-1} &\equiv 320 \end{aligned}$$

There is no discernible pattern. This is one reason by modular arithmetic is used in cryptography. The next section will discuss a theorem that is at the heart of the most popular public-key cryptosystem.

4.3 The Euler-Fermat Theorem

Just as inverses behave pseudorandomly in the field $\mathbb{Z}/p\mathbb{Z}$, powers also behave pseudorandomly. For example, here are the first several powers of the element $346 \in \mathbb{Z}/1009\mathbb{Z}$:

$$\begin{aligned} 346^1 &\equiv 346 \\ 346^2 &\equiv 972 \\ 346^3 &\equiv 352 \\ 346^4 &\equiv 360 \\ 346^5 &\equiv 93 \\ 346^6 &\equiv 806 \\ 346^7 &\equiv 595 \end{aligned}$$

This sequence seems to have no pattern. But we know that this cannot go on forever because the set $\mathbb{Z}/1009\mathbb{Z}$ is finite. I claim that the sequence of powers will eventually hit 1 and then it cycle through the same sequence endlessly.

To prove this, we first establish an exponential notation for elements of $\mathbb{Z}/p\mathbb{Z}$. For any positive integer $n \geqslant 1$ and for any nonzero element $a \in \mathbb{Z}/p\mathbb{Z}$ we know that a^n is also nonzero mod p because $\mathbb{Z}/p\mathbb{Z}$ is a domain. Furthermore, the inverse of a^n is just $(a^{-1})^n$ because

$$a^n \cdot (a^{-1})^n \equiv \underbrace{aa\cdots a}_{n\text{ times}} \cdot \underbrace{a^{-1}a^{-1}\cdots a^{-1}}_{n\text{ times}} \equiv 1 \bmod p.$$

[30] My computer told me that 1009 is prime.

[31] Strictly speaking, we do not need to include the x column.

This suggests that we should define the notation a^n for **any integer value** of n, including zero and negative integers:

$$a^n = \begin{cases} a^n & n \geqslant 1, \\ 1 & n = 0, \\ (a^{-1})^{-n} & n \leqslant -1. \end{cases}$$

Finally, we observe that this notation satisfies the general rule

$$a^{m+n} \equiv a^m \cdot a^n \bmod p \qquad \text{for any integers } m, n \in \mathbb{Z}.$$

The following theorem illustrates the utility of this notation.

The Multiplicative Order of an Element

Let p be prime. For any nonzero $a \in \mathbb{Z}/p\mathbb{Z}$ we consider the sequence of powers mod p:

$$a, a^2, a^3, a^4, \ldots \in \mathbb{Z}/p\mathbb{Z}.$$

Since $\mathbb{Z}/p\mathbb{Z}$ is finite, some element of this sequence must be repeated. Let's say $a^k \equiv a^\ell$ mod p for some integers $1 \leqslant \ell < k$. Then multiplying both sides by $a^{-\ell}$ gives

$$\begin{aligned} a^k &\equiv a^\ell \\ a^k \cdot a^{-\ell} &\equiv a^\ell \cdot a^{-\ell} \\ a^{k-\ell} &\equiv 1. \end{aligned}$$

We have shown that $a^{k-\ell} \equiv 1$ mod p for some positive integer $k - \ell \geqslant 1$. The smallest such integer is called the *order of a mod p*:

$$\text{ord}_p(a) = \min\{r \geqslant 1 : a^r \equiv 1 \bmod p\}.$$

Thus the sequence of powers $a, a^2, a^3, \ldots$ mod p will reach 1 after $\text{ord}_p(a)$ steps, after which the sequence will repeat. For example, consider the powers of 3 mod 11:

k	3^k mod 11
1	3
2	9
3	5
4	4
5	1
6	3
7	9
8	5
9	4
10	1
$\vdots$	$\vdots$

We see from this table that $\text{ord}_{11}(3) = 5$, and the sequence repeats after every 5 steps.

We have proved the existence of the numbers $\text{ord}_p(a) \in \mathbb{N}$ for all nonzero elements $a \in \mathbb{Z}/p\mathbb{Z}$. It is difficult to predict the exact value of $\text{ord}_p(a)$ for a given value of a. However, in this section we will prove the important theorem that the order always divides $p - 1$:

$$\text{ord}_p(a)|(p-1) \quad \text{for all nonzero elements } a \in \mathbb{Z}/p\mathbb{Z}.$$

This theorem was stated by Pierre de Fermat in a letter to Frénicle de Bessy in 1640. After giving some examples, Fermat said: "I would send you the demonstration, if I did not fear it being too long."[32] This was a common way of communicating scientific discoveries at the time, since there were no scientific journals. The first published proofs of Fermat's theorem were given by Euler in the 1700s. We will present Euler's second proof from 1761 since it involves a concept that will be important in this course: the concept of a *group*. We will present the modern definition, even though this concept was not formalized until the late 1800s.

Informally, a group is a set with an invertible, associative, binary operation. The main examples are addition $+$, multiplication $\cdot$ and functional composition $\circ$. Each of these examples also has a special "identity element", which is 0 for addition, 1 for multiplication, and the identity function id for functional composition. Because functional composition is not commutative, we do not assume that a group operation is commutative.

The Concept of a Group

A *group* consists of a set G together with a binary operation $* : G \times G \to G$, which we write as $a * b$, and a special element $\varepsilon \in G$ satisfying the following three axioms:

[32] Oystein Ore, *Number theory and its history*, page 272.

(G1) $\forall a,b,c \in G,\ a*(b*c) = (a*b)*c$ (associative)

(G2) $\forall a \in G,\ a*\varepsilon = \varepsilon * a = a$ (identity)

(G3) $\forall a \in G,\ \exists b \in G,\ a*b=\varepsilon$ and $b*a=\varepsilon$ (inverses)

We say that the group $(G,*,\varepsilon)$ is *abelian* if it satisfies the additional axiom[33]

(G4) $\forall a,b \in G, a*b = b*a$ (commutative)

Axiom (G3) says that any element of a group has a two-sided inverse. In fact, this inverse must be unique. To see this, suppose that we have $a*b=b*a=\varepsilon$ and $a*c=c*a=\varepsilon$. It follows that

$$\begin{aligned} b &= b*\varepsilon && \text{(G2)}\\ &= b*(a*c) \\ &= (b*a)*c && \text{(G1)}\\ &= \varepsilon * c \\ &= c. && \text{(G2)} \end{aligned}$$

Since the inverse of a is unique, we give the name a^{-1}. This notation makes sense when $*$ is multiplication or functional composition, but is less appropriate when $*$ is addition. In that case we might sometimes write $-a$ for the inverse.

We have already seen some examples of groups. If $(R,+,\cdot,0,1)$ is a ring then the structure $(R,+,0)$ is an abelian group. The structure $(R,\cdot,1)$ is not a group[34] because it contains the element $0 \in R$ which has no multiplicative inverse, and it may contain other non-invertible elements. However, the set of units $(R^\times,\cdot,1)$ is an abelian group, called the *group of units* of the ring. The ring R is a field if and only if $R^\times = R\backslash\{0\}$.

So far we have not studied any examples of non-abelian groups. These kind of groups come from functional composition. Here are two of the prototypical examples:

- Given a field $\mathbb{F}$ and a positive integer $n \geqslant 1$ we define

  $$\mathrm{GL}_n(\mathbb{F}) = \text{the set of invertible } n\times n \text{ matrices with entries from } \mathbb{F}.$$

  This is a group, called a *general linear group*, with group operation given by matrix multiplication and identity element given by the $n \times n$ identity matrix.

[33] This is a peculiar notation. It would be more sensible to call this a *commutative group*. This "abelian" notation was introduced by Leopold Kronecker to commemorate from a theorem of Niels Henrik Abel, which says that a polynomial equation with a commutative "Galois group" is solvable by radicals. We will discuss this next semester.

[34] I don't want to overwhelm you with terminology, but a structure $(G,*,\varepsilon)$ satisfying axioms (G1) and (G2) is called a *monoid*.

- Invertible functions from a finite set to itself are called *permutations*. The permutations of a set form a group under composition, with the identity permutations as the identity element. The group of permutations of $\{1, 2, \ldots, n\}$ is called the *symmetric group* S_n.

Our discussion of multiplicative order generalizes to any group.

Order of a Group Element

Let $(G, *, \varepsilon)$ be a group. Then for any element $a \in G$ and for any integer $n \in \mathbb{Z}$ we define the exponential notation

$$a^n = \begin{cases} a * a * \cdots * a \text{ (n times)} & \text{if } n \geqslant 1 \\ \varepsilon & \text{if } n = 0 \\ a^{-1} * a^{-1} * \cdots * a^{-1} \text{ ($-n$ times)} & \text{if } n \leqslant -1 \end{cases}$$

One can check that this notation satisfies $a^{m+n} = a^m * a^n$ for all integers $m, n \in \mathbb{Z}$. We define the *order* of $a \in G$ as the minimum positive exponent r such that $a^r = \varepsilon$, or as ∞ if no such exponent exists:

$$\mathrm{ord}_G(a) = \min\{r \geqslant 1 : a^r = \varepsilon\} \in \mathbb{Z}_{\geqslant 1} \cup \{\infty\}.$$

If G is a finite then then I claim that $\mathrm{ord}_G(a)$ is finite. Indeed, in this case the sequence of powers $a, a^2, \ldots \in G$ must contain repetition, so that $a^k = a^\ell$ for some $k > \ell \geqslant 1$. Then we have

$$\begin{aligned} a^k &= a^\ell \\ a^k * a^{-\ell} &= a^\ell * a^{-\ell} \\ a^{k-\ell} &= a^0 \\ a^{k-\ell} &= \varepsilon \end{aligned}$$

for some positive integer $k - \ell \geqslant 1$.

The Euler-Fermat theorem shows us that the order of an element in a finite group is related to the size of the group. We will prove this in modern group-theoretic language but the ideas are due to Euler (1761). We will discuss afterwards how this abstract version implies the classical theorems of Euler and Fermat.

The Euler-Fermat Theorem

Let $(G, *, \varepsilon)$ be a finite abelian group. Then for all $a \in G$ we have[35]

$$a^{\#G} = \varepsilon.$$

To save space we will write $a * b = ab$ and $\varepsilon = 1$, but the proof is completely general. Consider the function $\mu_a : G \to G$ defined by $\mu_a(b) = ab$. This function is injective because every element of a group is invertible:

$$\begin{aligned} \mu_a(b) &= \mu_a(c) \\ ab &= ac \\ a^{-1}ab &= a^{-1}ac \\ b &= c. \end{aligned}$$

If G is finite then the function μ_a is also surjective. To be precise, suppose that $m = \#G$ and label the group elements as $G = \{b_1, b_2, \ldots, b_m\}$. Then we also have $G = \{ab_1, ab_2, \ldots, ab_m\}$ with the group elements possibly listed in a different order. Indeed, every element b_j has the form ab_i for some i because μ_a is surjective, and $ab_i = ab_j$ implies $b_i = b_j$ because μ_a is injective. Now we "multiply" all of the group elements together in two different ways:

$$\begin{aligned} b_1 b_2 \cdots b_m &= (ab_1)(ab_2)\cdots(ab_m) \\ \cancel{b_1 b_2 \cdots b_m} &= a^m \cancel{b_1 b_2 \cdots b_m} \\ 1 &= a^m. \end{aligned}$$

Euler's original application was to the group of units of the finite ring $\mathbb{Z}/n\mathbb{Z}$. I claim that

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{a \in \mathbb{Z}/n\mathbb{Z} : \gcd(a, n) = 1\}.$$

Indeed, if $\gcd(a, n) = 1$ then from Bézout's Identity we have $ax + ny = 1$ for some $x, y \in \mathbb{Z}$. It follows that

$$\begin{aligned} ax + ny &= 1 \\ ax - 1 &= n(-y) \\ n &\mid (ax - 1) \\ ax &\equiv 1 \bmod n, \end{aligned}$$

and hence $a \in \mathbb{Z}/n\mathbb{Z}$ is a unit. Conversely, suppose that $a \in \mathbb{Z}/n\mathbb{Z}$ is a unit, so that $ab \equiv 1 \bmod n$ for some $b \in \mathbb{Z}$. By definition this means that $ab - 1 = nk$ for some $k \in \mathbb{Z}$. If $d \in \mathbb{Z}$

[35] In fact, this theorem also holds for finite non-abelian groups, but the proof is harder.

is any common divisor of a and n then the equation $1 = ab - nk$ implies that $d|1$ and hence $d = \pm 1$. In other words, $\gcd(a, n) = 1$.

Euler's Totient Theorem

For any integer $n \geqslant 1$ we define *Euler's totient function*[36]

$$\phi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times = \#\{a \in \mathbb{Z} : 1 \leqslant a < n \text{ and } \gcd(a, n) = 1\}.$$

Since $\phi(n)$ is the size of the abelian group $(\mathbb{Z}/n\mathbb{Z})^\times$, the previous theorem tells us that

$$a^{\phi(n)} \equiv 1 \bmod n \text{ for all } a \in (\mathbb{Z}/n\mathbb{Z})^\times.$$

In other words,

$$a^{\phi(n)} \equiv 1 \bmod n \text{ for all } a \in \mathbb{Z} \text{ such that } \gcd(a, n) = 1.$$

If p is prime then $\mathbb{Z}/p\mathbb{Z}$ is a field. In other words, every nonzero element of $\mathbb{Z}/p\mathbb{Z}$ is a unit:

$$\begin{aligned} (\mathbb{Z}/p\mathbb{Z})^\times &= (\mathbb{Z}/p\mathbb{Z})\backslash\{0\} \\ \#(\mathbb{Z}/p\mathbb{Z})^\times &= \#(\mathbb{Z}/p\mathbb{Z}) - 1 \\ \phi(p) &= p - 1. \end{aligned}$$

Thus we recover the original theorem of Fermat, which was Euler's goal.

Fermat's Little Theorem

Let p be prime so that $\gcd(a, p) = 1$ if and only if $p \nmid a$. Then since $\phi(p) = p - 1$, Euler's totient theorem tells us that

$$a^{p-1} \equiv 1 \bmod p \;\text{ for all } a \in \mathbb{Z} \text{ such that } p \nmid a.$$

We can clean this up a bit by multiplying both sides by a to obtain

$$a^p \equiv a \bmod p,$$

which is true for any integer $a \in \mathbb{Z}$ whatsoever.

[36] This notation was introduced by James Joseph Sylvester in 1879. Sylvester is famous for introducing ridiculous mathematicial terminology, a small percentage of which has become standard. For example, Sylvester introduced the term *matrix* for a rectangular array of numbers, his reasoning being that such an array is a "womb" that gives birth to determinants. True story.

This result is called *Fermat's Little Theorem* in order to distinguish it from *Fermat's Last Theorem.*[37] Fermat, being an amateur mathematician working in a time before scientific journals, left behind few proofs. Euler later supplied proofs for most of Fermat's claimed results and disproved at least one.[38] But Euler was unable to prove or disprove the following.

Fermat's Last Theorem

For all positive integers a, b, c, n with $n \geqslant 3$ we have

$$a^n + b^n \neq c^n.$$

This problem became famous and inspired many fundamental concepts in number theory. It was finally proved in 1993 by Andrew Wiles and appeared on the front page of the New York Times. A gap in the proof led to some panic but Wiles was able to patch the gap with his student Richard Taylor, and a correct proof appeared in 1994. The ideas of this proof are far beyond the scope of our course.

4.4 The Chinese Remainder Theorem

Recall Euler's totient function:

$$\phi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times = \#\{a \in \mathbb{Z} : 1 \leqslant a < n \text{ and } \gcd(a,n) = 1\}.$$

We proved last time that

$$a^\phi(n) \equiv 1 \bmod n \quad \text{ for all integers } a \in \mathbb{Z} \text{ satisfying } \gcd(a,n) = 1.$$

If p is prime then since $\phi(p) = p - 1$ we obtain Fermat's little theorem:

$$a^{p-1} \equiv 1 \bmod p \quad \text{ for all integers } a \in \mathbb{Z} \text{ satisfying } p \nmid a.$$

But what if n is not prime? In this section we will prove the following formula:

$$\phi(n) = n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

where the product is taken over all prime divisors $p|n$. This result seems intuitively plausible. Indeed, we observe that $\gcd(a,n) \neq 1$ if and only if a and n share a prime factor. Thus we

[37] I do not know the origin of these names.

[38] Fermat had claimed that the number $2^{2^n} + 1$ is prime for all integers $n \geqslant 0$. Euler shows that $2^{2^5} + 1$ is not prime, and no other *Fermat prime* has ever been found. So this is a case where Fermat was completely wrong.

wish to remove all multiples of the prime factors of n. We can remove the multiples of p by multiplying n with $(1 - 1/p)$. Then, presumably, we can remove the multiples of another prime factor q by multiplying the result with $(1-1/q)$. But this is not so simple because some multiples of q are also multiples of p.

The underlying issue is today expressed in terms of a general property of rings called the "Chinese Remainder Theorem".[39] The first example of the theorem appeared in the fourth century text *Sun Zu Suan Jing* (Master Sun's Mathematical Manual):

> There are certain things whose number is unknown. If we count them by threes, we have two left over; by fives, we have three left over; and by sevens, two are left over. How many things are there?

In modern language, we are looking for integer solutions $c \in \mathbb{Z}$ to the following system of congruences:

$$\begin{cases} c \equiv 2 \bmod 3, \\ c \equiv 3 \bmod 5, \\ c \equiv 2 \bmod 7. \end{cases}$$

Instead of just solving this one problem we will develop the general theory. The idea is to compare the set $\mathbb{Z}/mn\mathbb{Z}$ with the cartesian product set $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. To be specific, we consider the function sending the congruence class a mod mn to the pair of congruence classes $(a \text{ mod } m, a \text{ mod } n)$. Here is an example with $m = 2$ and $n = 3$:

a mod 6	$(a \text{ mod } 2, a \text{ mod } 3)$
0	$(0,0)$
1	$(1,1)$
2	$(0,2)$
3	$(1,0)$
4	$(0,1)$
5	$(1,2)$

Note that each ordered pair on the right appears exactly once, which happens because 2 and 3 are coprime. Indeed, we see that the first coordinate cycles through $\{0,1\}$ while the second coordinate cycles through $\{0,1,2\}$. Since 2 and 3 are coprime there is no repetition. We will be more precise about this below.

In practical terms, this example tells us that each system of congruences $c \equiv a$ mod 2 and $c \equiv b$ mod 3 has a unique solution mod 6. For example, the final row of the table tells us that

$$\left\{ \begin{array}{lcl} c & \equiv & 1 \bmod 2 \\ c & \equiv & 2 \bmod 3 \end{array} \right\} \quad \Longleftrightarrow \quad c \equiv 5 \bmod 6.$$

In general, we would like a recipe to send a pair of congruence classes mod m and n to a unique congruence class mod mn. This is what the Chinese Remainder Theorem does. Actually, the term "Chinese Remainder Theorem" refers to a collection of ideas, which I will break into a

[39] The theorem was named by Leonard Dickson in 1929 and this notation has become standard.

few pieces. The proof will use two lemmas, which are only slight modification of things that we already know.

Lemmas for the Chinese Remainder Theorem

(1) If $\gcd(m,n)=1$ then $m|c$ and $n|c$ imply $(mn)|c$.

(2) If $ax+by=1$ then $\gcd(a,b)=1$.

To prove (1), let $\gcd(m,n)=1$ so that $mx+ny=1$ for some $x,y\in\mathbb{Z}$. If $mk=c$ and $n\ell=c$ for some $k,\ell\in\mathbb{Z}$ then

$$\begin{aligned}(mx+ny)c&=c\\ mxc+nyc&=c\\ mxn\ell+nymk&=c\\ mn(x\ell+yk)&=c.\end{aligned}$$

To prove (2), let $ax+by=1$. If $dk=a$ and $d\ell=b$ then

$$1=ax+by=dkx+d\ell y=d(kx+\ell y).$$

In other words, any common divisor of a and b must be a divisor of 1. Hence $\gcd(a,b)=1$.

Remark: It is always possible to use unique prime factorization to prove things like this. But there is a general rule when writing proofs that one should not use a deeper theorem to prove a shallower theorem. This helps minimize the risk of circular reasoning.

Chinese Remainder Theorem, Part I

Let integers $m,n\geqslant 1$ satisfy $\gcd(a,b)=1$ and consider the following function:

$$\begin{array}{rrcl}\varphi: & \mathbb{Z}/mn\mathbb{Z} & \to & \mathbb{Z}/m\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z}\\ & a \bmod mn & \mapsto & (a \bmod m, a \bmod n).\end{array}$$

To save space we could write $\varphi(a)=(a,a)$, as long as we are clear that the input is a congruence class mod mn and the output is ordered pair of congruence classes mod m and n. I claim that φ is a bijection.

What needs to be proved?

- **Well-Defined?**[40] First we should check that the definition is not affected by changing a to another integer a' satisfying $a \equiv a' \bmod mn$. Indeed, if $a \equiv a' \bmod mn$, so that $a - a' = mnk$ for some $k \in \mathbb{Z}$, then we have $a - a' = m(nk)$, which implies that $a \equiv a' \bmod m$ and $a - a' = n(mk)$, which implies that $a \equiv a' \bmod n$.
- **Injective?** Suppose that $a \equiv b \bmod m$ and $a \equiv b \bmod n$, so that $m|(a-b)$ and $n|(a-b)$. Then from Lemma (1) we have $mn|(a-b)$, so that $a \equiv b \bmod mn$.
- **Surjective?** We have an injective function from the set $\mathbb{Z}/mn\mathbb{Z}$ to the set $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Since these sets have the same size mn any injective function must also be surjective.

It follows that the function $\varphi : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ has an inverse function:

$$\begin{array}{rrcl} \varphi^{-1}: & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} & \to & \mathbb{Z}/mn\mathbb{Z} \\ & (a \bmod m, b \bmod n) & \mapsto & ? \bmod mn. \end{array}$$

But it is not at all clear how to express the output as a function of the input (a, b).

Chinese Remainder Theorem, Part 2

Let integers $m, n \geqslant 1$ satisfy $\gcd(m, n) = 1$, so we can use the Extended Euclidean Algorithm to find some (non-unique) integers $x, y \in \mathbb{Z}$ satisfying

$$mx + ny = 1.$$

I claim that the inverse of the function $\varphi(a \bmod mn) = (a \bmod m, a \bmod n)$ from $\mathbb{Z}/mn\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ can be computed as follows:[41]

$$\varphi^{-1}(a \bmod m, b \bmod n) = any + bmx \bmod mn.$$

In concrete terms, we have the following solution to a system of two congruences:

$$\left\{ \begin{array}{rcl} c & \equiv & a \bmod m \\ c & \equiv & b \bmod n \end{array} \right\} \quad \Longleftrightarrow \quad c \equiv any + bmx \bmod mn.$$

To prove this we only need to check that $\varphi(any + bmx) = (a, b)$. In other words, we need to check that

$$\begin{aligned} any + bmx &\equiv a \bmod m, \\ any + bmx &\equiv b \bmod n. \end{aligned}$$

[40] Students usually have difficulty with the concept of "well-definedness". The idea is that a function whose input is an equivalence class must not be affected by changing the representative from this class.

We only need to check one of these because they are symmetric. All congruences in the following computation are mod m:

$$\begin{aligned} any + bmx &\equiv any + b0x \\ &\equiv any \\ &\equiv a(1 - mx) \\ &\equiv a(1 - 0x) \\ &\equiv a. \end{aligned}$$

For example, when $m = 2$ and $n = 3$ we can take $x = -1$ and $y = 1$, so that $any + bmx = 3a - 2b$, and hence[42]

$$\left\{ \begin{array}{ccc} c & \equiv & a \bmod 2 \\ c & \equiv & b \bmod 3 \end{array} \right\} \quad \Longleftrightarrow \quad c \equiv 3a - 2b \bmod 6.$$

We can use the same method to solve multiple simultaneous congruences by induction. Recall Sun Zu's system of congruences:

$$\left\{ \begin{array}{ccl} c & \equiv & 2 \bmod 3, \\ c & \equiv & 3 \bmod 5, \\ c & \equiv & 2 \bmod 7. \end{array} \right.$$

First we take $m = 3$ and $n = 5$ and observe that $3(2) + 5(-1) = 1$, so that

$$\left\{ \begin{array}{ccc} c & \equiv & 2 \bmod 3 \\ c & \equiv & 3 \bmod 5 \end{array} \right\} \Longleftrightarrow c \equiv 2 \cdot 5(-1) + 3 \cdot 3(2) \equiv 8 \bmod 15.$$

Hence we have

$$\left\{ \begin{array}{ccc} c & \equiv & 2 \bmod 3 \\ c & \equiv & 3 \bmod 5 \\ c & \equiv & 2 \bmod 7 \end{array} \right\} \quad \Longleftrightarrow \quad \left\{ \begin{array}{ccc} c & \equiv & 8 \bmod 15 \\ c & \equiv & 2 \bmod 7 \end{array} \right\}.$$

Then we take $m = 15$ and $n = 7$ and observe that $15(1) + 7(-2) = 1$, so that

$$\left\{ \begin{array}{ccc} c & \equiv & 8 \bmod 15 \\ c & \equiv & 2 \bmod 7 \end{array} \right\} \quad \Longleftrightarrow \quad c \equiv 8 \cdot 7(-2) + 2 \cdot 15(1) \equiv 23 \bmod 105.$$

On the homework will you investigate a method to solve a system of multiple congruences in one step. It is not any faster but it is slightly more beautiful.

[41] Over the years I have settled on this mnemonic because *any* is a word and *bmx* is a type of bicycle that was popular in my childhood.

[42] We could equally well take $x = 2$ and $y - 1$. The solution would look different but it would be the same.

We end this section by using the Chinese Remainder Theorem to compute Euler's totient function. We have seen that the the following function is well-defined for any integers $m, n \geqslant 1$:

$$\begin{array}{rrcl} \varphi : & \mathbb{Z}/mn\mathbb{Z} & \to & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ & a \bmod mn & \mapsto & (a \bmod m, a \bmod n). \end{array}$$

But this is not just a function between sets. We know that $\mathbb{Z}/mn\mathbb{Z}$ is a ring and we can also view $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ as a ring by defining addition and multiplication componentwise:

$$\begin{aligned} (a \bmod m, b \bmod n) + (a \bmod m, b \bmod n) &= (a + a' \bmod m, b + b' \bmod n), \\ (a \bmod m, b \bmod n) \cdot (a \bmod m, b \bmod n) &= (aa' \bmod m, bb' \bmod n). \end{aligned}$$

The "zero" and "one" elements of this ring are $(0, 0)$ and $(1, 1)$. Since the function φ preserves this ring structure we say that φ is a *ring homomorphism*. When $\gcd(m, n) = 1$ we also know that φ is a bijection, in which case we say it is a *ring isomorphism*. The final piece of the Chinese Remainder Theorem says that this ring isomorphism restricts to a *group isomorphism* between the groups of units. I won't bother to use this language in the official statement. We will be much more systematic about homomorphisms next semester.

Chinese Remainder Theorem, Part 3

Let integers $m, n \geqslant 1$ satisfy $\gcd(m, n) = 1$, so the function $\varphi(a) = (a, a)$ defines a bijection:

$$\varphi : \mathbb{Z}/mn\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

I claim that this restricts to a bijection:

$$\varphi : (\mathbb{Z}/mn\mathbb{Z})^\times \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times.$$

Hence the domain and codomain have the same size, which gives us the following identity for Euler's totient function:

$$\phi(mn) = \#(\mathbb{Z}/mn\mathbb{Z})^\times = \#(\mathbb{Z}/m\mathbb{Z})^\times \cdot \#(\mathbb{Z}/n\mathbb{Z})^\times = \phi(m)\phi(n).$$

What needs to be checked? We only need to show that a is a unit mod mn if and only if a is a unit mod m and n separately:

$$\gcd(a, mn) = 1 \quad \Longleftrightarrow \quad \gcd(a, m) = 1 \text{ and } \gcd(a, n) = 1.$$

For one direction, suppose that $\gcd(a, mn) = 1$ so that $ax + mny = 1$ for some $x, y \in \mathbb{Z}$. Then since $ax + m(ny) = 1$, Lemma (2) implies that $\gcd(a, m) = 1$ and since $ax + n(my) = 1$, Lemma (2) implies that $\gcd(a, n) = 1$. Conversely, suppose that $\gcd(a, m) = 1$ and $\gcd(a, n) = 1$, hence there exist integers $x, y, x', y' \in \mathbb{Z}$ satisfying $ax + my = 1$ and $ax' + ny' = 1$. Multiplying these equations gives

$$(ax + my)(ax' + ny') = 1$$

$$a(xx' + xny' + myx') + mn(yy') = 1,$$

and it follows from Lemma (2) that $\gcd(a, mn) = 1$.

Finally, we will prove the formula from the beginning of the section. Consider the prime factorization of an integer $n \geqslant 1$:

$$n = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}.$$

Applying the previous result gives

$$\phi(n) = \phi(p_1^{n_1})\phi(p_2^{n_2}) \cdots \phi(p_k^{n_k}).$$

But now we are stuck. It is **not** true that $\phi(p^2) = \phi(p)\phi(p)$ because p is not coprime to p. We need to find a way to compute $\phi(p^m)$ when p is prime. I claim that

$$\phi(p^m) = p^m - p^{m-1} = p^m \left(1 - \frac{1}{p}\right).$$

To see this, we first observe that

$$\gcd(a, p^m) = 1 \iff p \nmid a.$$

Indeed, since p is prime the only divisors of p^m are the powers of p. If $p \nmid a$ then a is also not divisible by any power of p, hence a and p^m have no common divisor. Conversely, if $p|a$ then p is a nontrivial common divisor of a and p^m.

Recall that $\phi(p^m)$ is the number of integers between 1 and p^m that are coprime to p^m. By the previous remark these are just the integers that are not divisible by p. So our goal is to count the integers between 1 and p^m that are not divisible by p. But it is easier to count the integers that **are** divisible by p. Indeed, there are p^{m-1} multiples of p in this range:

$$1p, 2p, 3p, \ldots, (p^{m-1})p.$$

Then throwing away these multiples of p gives $\phi(p^m) = p^m - p^{m-1}$ as desired.

We conclude that

$$\begin{aligned}
\phi(n) &= \phi(p_1^{n_1})\phi(p_2^{n_2}) \cdots \phi(p_k^{n_k}) \\
&= p_1^{n_1}\left(1 - \frac{1}{p_1}\right) p_2^{n_2}\left(1 - \frac{1}{p_2}\right) \cdots p_k^{n_k}\left(1 - \frac{1}{p_k}\right) \\
&= p_1^{n_1} p_2^{n_2} p_k^{n_k}\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\left(1 - \frac{1}{p_k}\right) \\
&= n \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right) \\
&= n \prod_{p|n}\left(1 - \frac{1}{p}\right),
\end{aligned}$$

where the product is taken over the prime divisors of n.

5 Partial Fractions

5.1 Leibniz' Mistake

After our detour through number theory, we return to the theory of polynomials over a field. Because $\mathbb{Z}$ and $\mathbb{F}[x]$ are both examples of Euclidean domains we will find that some of the theorems have already been proved. In particular, in this section we will see that the method of partial fractions from calculus is basically equivalent to the Chinese Remainder Theorem from number theory.

The goal of this chapter is to prove the following theorem. There are many equivalent statements; for now we will state the original version.

The Fundamental Theorem of Algebra (Original Version)

Every non-constant polynomial $f(x) \in \mathbb{R}[x]$ can be expressed as

$$f(x) = p_1(x)p_2(x)\cdots p_k(x),$$

where $p_i(x) \in \mathbb{R}[x]$ and $\deg(p_i) = 1$ or 2 for all i.

We will see that this result is highly non-trivial. Several generations of mathematicians (including Euler) tried and failed to give a rigorous proof. Even the first generally accepted proofs had logical gaps that were not completely filled until the late 1800s.

The fundamental theorem is so difficult that Gottfried Leibniz, one of the two founders of Calculus, temporarily convinced himself that it is false. In 1702, Leibniz wrote a paper on the integration of rational expressions $f(x)/g(x)$ where $f(x), g(x) \in \mathbb{R}[x]$. If the denominator $g(x)$ could be factored into polynomials of degrees 1 and 2 then Leibniz knew that the integral could be solved by means of the following two basic integrals:

$$\int x^n dx = \begin{cases} x^{n+1}/(n+1) & \text{if } n \neq -1 \\ \log|x| & \text{if } n = -1 \end{cases} \qquad \text{and} \qquad \int \frac{1}{x^2+1} dx = \arctan(x).$$

For example, consider the integral

$$\int \frac{x^5}{x^4 - 2x^3 + 2x^2 - 2x + 1}\, dx.$$

By inspection we see that $x = 1$ is a root of the denominator, which then factors as

$$x^4 - 2x^3 + 2x^2 - 2x + 1 = (x-1)^2(x^2+1).$$

After knowing this, one can use the method of partial fractions to compute[43]

$$\frac{x^5}{x^4-2x^3+2x^2-2x+1} = x+2+\frac{2}{x-1}+\frac{1/2}{(x-1)^2}-\frac{1/2}{x^2+1},$$

and then the integral is straightforward:

$$\int \frac{x^5}{x^4-2x^3+2x^2-2x+1}\,dx = \frac{x^2}{2}+2x+2\log|x-1|-\frac{1/2}{x-1}-\frac{1}{2}\arctan(x).$$

However, Leibniz claimed that not all real polynomials can be so factored. As an example he gave the polynomial x^4+a^4, where a is a real number. In his words:[44]

> *Therefore $\int \frac{dx}{x^4+a^4}$ cannot be reduced to the squaring of the circle or the hyperbola by our analysis above, but founds a new kind of its own.*

To see that this is wrong, we will compute the 4th roots of $-a^4$ for any positive number $a>0$. First we write $-a^4$ in polar form as

$$-a^4 = a^4 e^{i\pi}.$$

Thus the principal 4th root is

$$ae^{i\pi/4} = a\left[\cos(\pi/4)+i\sin(\pi/4)\right] = \frac{a}{\sqrt{2}}(1+i),$$

and since $1, i, -1, -i$ are the 4th roots of unity, the remaining 4th roots of $-a^4$ are

$$\begin{aligned} ae^{i\pi/4}i &= a(i-1)/\sqrt{2}, \\ ae^{i\pi/4}(-1) &= a(-1-i)/\sqrt{2}, \\ ae^{i\pi/4}(-i) &= a(-i+1)/\sqrt{2}. \end{aligned}$$

Then grouping these roots into conjugate pairs gives the following factorization:

$$\begin{aligned} x^4+a^4 &= \Big[(x-a(1+i)/\sqrt{2})(x-a(1-i)/\sqrt{2})\Big]\Big[(x-a(-1+i)/\sqrt{2})(x-a(-1-i)/\sqrt{2})\Big] \\ &= (x^2-a\sqrt{2}x+a^2)(x^2+a\sqrt{2}+a^2). \end{aligned}$$

If Leibniz had found this factorization then he would have been able to compute the integral. To illustrate the method we will examine the simplest case $a=\sqrt{2}$. I claim that there exist real numbers A, B, C, D such that[45]

$$\begin{aligned} \frac{1}{x^4+4} &= \frac{1}{(x^2-2x+2)(x^2+2x+2)} \\ &= \frac{A+Bx}{x^2-2x+2}+\frac{C+Dx}{x^2+2x+2}. \end{aligned}$$

[43] We will discuss this method in detail below.

[44] "Squaring the circle" refers to arctan and "squaring the hyperbola" refers to log.

[45] This follows from a general theorem on partial fractions which we will prove below.

To find these numbers we could add the fractions on the right hand side and then equate the coefficients in the numerator to the numerator $1 = 1 + 0x + 0x^2 + 0x^3$ on the left side. This would lead to a system of four linear equations in four unknowns, which is not too difficult to solve. However, we will use a more general method that is common to all Euclidean Domains.

First we will apply the Extended Euclidean Algorithm in the ring $\mathbb{R}[x]$ to obtain some polynomials $\alpha(x), \beta(x) \in \mathbb{R}[x]$ satisfying

$$(x^2 + 2x + 2)\alpha(x) + (x^2 - 2x + 2)\beta(x) = 1.$$

The method here is exactly the same as for integers, though the calculations are a bit more involved. Consider the set of triples of polynomials

$$V = \{(\alpha(x), \beta(x), \gamma(x)) \in \mathbb{R}[x]^3 : f(x)\alpha(x) + g(x)\beta(x) = \gamma(x)\},$$

which is closed under $\mathbb{F}[x]$-linear combinations.[46] Then beginning with the basic triples $(1, 0, x^2 + 2x + 2)$ and $(0, 1, x^2 - 2x + 1)$ we perform the steps of the Euclidean Algorithm to obtain a triple of the form $(\alpha(x), \beta(x), \gamma(x))$, where $\gamma(x)$ is the greatest common divisor. In this case we find that $\gamma(x) = 1$:

$\alpha(x)$	$\beta(x)$	$\gamma(x)$
1	0	$x^2 + 2x + 2$
0	1	$x^2 - 2x + 2$
1	-1	$4x$
$-x/4 + 1/2$	$x/4 + 1/2$	2
$-x/8 + 1/4$	$x/8 + 1/4$	1

To get from the third to the fourth row we need to compute the quotient and remainder of $x^2 - 2x + 2$ mod $4x$:

$$\begin{array}{r l}
 & \frac{1}{4}x - \frac{1}{2} \\ \cline{2-2}
4x \big) & x^2 - 2x + 2 \\
 & -x^2 \\ \cline{2-2}
 & \quad -2x \\
 & \quad 2x \\ \cline{2-2}
 & \qquad\quad 2
\end{array}$$

Then the fourth row equals the second row minus $(x/4 - 1/2)$ times the third row. In the last step we just scaled everything by $1/2$ to obtain the monic GCD. In conclusion, we have have

$$1 = \frac{1}{8}(2 - x)(x^2 + 2x + 2) + \frac{1}{8}(2 + x)(x^2 - 2x + 2).$$

Then we divide both sides by $x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2)$ to obtain the desired partial fraction expansion:

$$\frac{1}{(x^2 + 2x + 2)(x^2 - 2x + 2)} = \frac{(2 - x)/8 \cdot (x^2 + 2x + 2)}{(x^2 + 2x + 2)(x^2 - 2x + 2)} + \frac{(2 + x)/8 \cdot (x^2 - 2x + 2)}{(x^2 + 2x + 2)(x^2 - 2x + 2)}$$

[46] We say that V is an $\mathbb{F}[x]$-module.

$$\frac{1}{x^4+4} = \frac{(2-x)/8}{x^2-2x+2} + \frac{(2+x)/8}{x^2+2x+2}$$

At this point, Leibniz would easily have computed the integral in terms of log and arctan. Since it is not easy for me, and since this is not a Calculus class, I will just tell you the answer that my computer gives:

$$\int \frac{dx}{x^4+4} = \frac{\arctan(x+1)+\arctan(x-1)}{8} + \frac{\log(x^2+2x+2)-\log(x^2-2x+2)}{16}.$$

5.2 Fractions

In the previous section we discussed "rational expressions" $f(x)/g(x)$ where $f(x)$ and $g(x)$ are polynomials. Since we have been careful to study polynomials from an abstract point of view, we should do the same for rational expressions. The construction of fractions of polynomials is completely analogous to the construction of fractions of integers. More generally, for any integral domain R there is a well-defined "field of fractions" $\text{Frac}(R)$. In this section we will study the formal details of this construction. Then we will have earned the right to treat fractions informally for the rest of the course.

For any ring R we may consider the set of "fractional expressions":

$$\text{Frac}(R) = \{a/b : a, b \in R, b \neq 0\}.$$

At first we do not attach any meaning to the abstract symbol "a/b". Of course, our goal is to treat these symbols in the same way that we do fractions of integers. The first difficulty is that many different-looking looking symbols correspond to the same "value":

$$\frac{1}{2} = \frac{-1}{-2} = \frac{7}{14} = \frac{-13}{-26} = \cdots$$

From past experience, we know that two fractions a/b and c/d are equal if and only if the integers ad and bc are equal. Thus we define the following relation over a general ring R:

$$\frac{a}{b} \sim \frac{c}{d} \text{ in Frac}(R) \quad\iff\quad ad = bc \text{ in } R.$$

Our first goal is to verify that $\sim$ is an equivalence relation on the set $\text{Frac}(R)$:

- **Reflexive.** For all $a, b \in R$ we have $ab = ba$, which implies that $a/b \sim b/a$.
- **Symmetric.** Suppose that $a/b \sim c/d$ for some $a, b, c, d \in R$, which by definition means that $ad = bc$. But then we have $cb = da$, which implies that $c/d = a/b$. Here we assumed that R is a **commutative ring**.
- **Transitive.** Suppose that we have $a/b = c/d$ and $c/d = e/f$ for some $a, b, c, d, e, f \in R$. By definition, this means that $ad = bc$ and $cf = de$. In this case we wish to show that $af = be$, so that $a/b = e/f$. For this we will use the associative and commutative properties of R:

$$d(af) = (ad)f = (bc)f = b(cf) = b(de) = d(be).$$

Now we might be stuck. However, if R is an **integral domain**, then since d is nonzero (because it is the denominator of the fraction c/d) we may cancel it from both sides to obtain $af = be$ as desired.

Here is a summary.

Equivalence of Fractions

If R is an integral domain then the relation

$$\frac{a}{b} \sim \frac{c}{d} \quad \Longleftrightarrow \quad ad = bc$$

is an equivalence on the set of fractional expressions $\text{Frac}(R) = \{a/b : a, b \in R, b \neq 0\}$.

Now recall that fractions of integers can be added and multiplied as follows:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd} \quad \text{and} \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

So we will do the same for fractions over an arbitrary domain.

Addition and Multiplication of Fractions

For any domain R and for any fractions $a/b, c/d \in \text{Frac}(R)$ we define

$$\frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd} \quad \text{and} \quad \frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}.$$

Note that the denominators are nonzero because $b \neq 0$ and $d \neq 0$ imply $bd \neq 0$ in a domain. More subtly, we must check that these operations are compatible with equivalence. In other words, if $a/b \sim a'/b'$ and $c/d \sim c'/d'$ then we must check that $(a/b) + (c/d) \sim (a'/b') + (c'/d')$ and $(a/b) \cdot (c/d) \sim (a'/b')(c'/d')$.

Proof. We have assumed that $ab' = a'b$ and $cd' = c'd$. It follows that

$$\begin{aligned}(ad+bc)(b'd') &= (ad)(b'd') + (bc)(b'd')\\ &= (ab')(dd') + (cd')(bb')\\ &= (a'b)(dd') + (c'd)(bb')\\ &= (a'd' + b'c')(bd),\end{aligned}$$

so that $(ad+bc)/(bd) \sim (a'd'+b'c')/(b'd')$, and

$$(ac)(b'd') = (ab')(cd') = (a'b)(c'd) = (a'c')(bd),$$

so that $(ac)/(bd) \sim (a'c')/(b'd')$.

Note that this proof is similar in spirit to the proof that addition and multiplication of integers is compatible with congruence mod n. In that case we obtained the ring $\mathbb{Z}/n\mathbb{Z}$ of congruence classes. So in this case we expect to get a "ring of fractions". In fact, we get more.

The Field of Fractions of a Domain

For any domain R we have a field of fractions[47]

$$(\text{Frac}(R), \sim, +, \cdot, 0/1, 1/1).$$

In other words, we have a ring of fractions with the operations $+, \cdot$ defined above, where $0/1$ is the additive identity and $1/1$ is the multiplicative identity. It is quite tedious to check the eight ring axioms, so we won't. In order to see that this is also a field, consider any "nonzero" fraction $a/b \not\sim 0/1$. By definition this means that $a1 \neq b0$, or $a \neq 0$. It follows that the fraction b/a exists, and we check that

$$\frac{a}{b} \cdot \frac{b}{a} \sim \frac{ab}{ba} \sim \frac{1}{1}.$$

In other words, $(a/b)^{-1} = b/a$.

It is common to "identify" the fraction of integers $a/1$ with the integer a, and thus to view the domain $\mathbb{Z}$ as a subring of the field $\mathbb{Q}$. In order to make this formal, it is more correct to say that the function $\mathbb{Z} \to \mathbb{Q}$ defined by $a \mapsto a/1$ is an *injective ring homomorphism*.[48] This observation leads to the so-called "universal property" of fractions.

Universal Property of Fractions

Let R be a domain with field of fractions $\text{Frac}(R)$. Then the following function is an

[47] There is a subtle point that the elements of this field are not formal fractions, but equivalence classes of formal fractions. Similarly, the elements of the ring $\mathbb{Z}/n\mathbb{Z}$ are not integers, but congruence classes of integers mod n. I don't want to be more precise about this right now.

[48] Is it disheartening to learn that you never really understood fractions in the first place?

injective ring homomorphism:

$$\begin{array}{rccl} \varphi : & R & \to & \mathrm{Frac}(R) \\ & a & \mapsto & a/1. \end{array}$$

This gives us a ring isomorphism between R and the subring $\{a/1 : a \in R\} \subseteq \mathrm{Frac}(R)$. It is common to abuse notation and to say that R *is a subring of* $\mathrm{Frac}(R)$.

More generally, let $\mathbb{F} \supseteq R$ be a field containing R as a subring. Then we obtain an injective ring homomorphism:

$$\begin{array}{rccl} \mu : & \mathrm{Frac}(R) & \to & \mathbb{F} \\ & a/b & \mapsto & ab^{-1}. \end{array}$$

This gives a ring isomorphism between $\mathrm{Frac}(R)$ and the subring $\{ab^{-1} : a, b \in R, b \neq 0\} \subseteq \mathbb{F}$. Furthermore, we observe that the map $\mu \circ \varphi : R \to \mathbb{F}$ is just the identity:

$$a \mapsto a/1 \mapsto a1^{-1} = a.$$

In colloquial terms, these results just say that

$\mathrm{Frac}(R)$ *is "the smallest field that contains R".*

Unfortunately, the messing about with arrows is necessary to make this colloquial idea precise, and therefore to prove anything about it.

This theorem is quite abstract so you can mostly forget about it for now; I just wanted to put it in front of your eyes.

Proof. The function φ is injective because $a/1 \sim b/1$ implies $a = b$ and it is a ring homomorphism because $a/1 + b/1 \sim (a+b)/1$ and $(a/1)(b/1) \sim (ab)/1$. The function μ is well-defined[49] because

$$a/b \sim a'/b' \quad \Rightarrow \quad ab' = a'b \quad \Rightarrow \quad ab^{-1} = a'(b')^{-1} \quad \Rightarrow \quad \mu(a/b) = \mu(a'/b').$$

Here we have used the facts that $\mathbb{F}$ is a field and b, b' are nonzero. And μ is injective because each of the implications above is reversible. Finally, μ is a ring homomorphism because

$$\begin{aligned} \mu(a/b) + \mu(c/d) &= ab^{-1} + cd^{-1} \\ &= (ab^{-1})(dd^{-1}) + (cd^{-1})(bb^{-1}) \\ &= (ad)(b^{-1}d^{-1}) + (bc)(b^{-1}d^{-1}) \\ &= (ad + bc)(b^{-1}d^{-1}) \end{aligned}$$

[49] For any function defined on a set of equivalence classes, one must check that the value of the function does not depend on the class representative used to compute it.

$$
\begin{aligned}
&= (ad+bc)(bd)^{-1} \\
&= \mu(a/b + c/d)
\end{aligned}
$$

and

$$\mu(a/b)\mu(c/d) = (ab^{-1}(cd^{-1}) = (ac)(b^{-1}d^{-1}) = (ac)(bd)^{-1} = \mu((a/b)(c/d)).$$

□

We have now earned the right to use fractional notation over an arbitrary integral domain; for example, over the ring of polynomials $\mathbb{F}[x]$.

The Field of Rational Functions

For any field $\mathbb{F}$ we use the following notation

$$\mathbb{F}(x) := \text{Frac}(\mathbb{F}[x])$$

and we call this the *field of rational functions over* $\mathbb{F}$.

Recall that for infinite fields $\mathbb{F}$ there is no difference between formal polynomial expressions $\mathbb{F}[x]$ and polynomial functions $\mathbb{F} \to \mathbb{F}$.[50] Unfortunately, the situation is more complicated for "rational functions". For example, the rational function $1/(x^2+1) \in \mathbb{R}(x)$ defines a perfectly good function $\mathbb{R} \to \mathbb{R}$, but if we think of $1/(x^2+1)$ as an element of $\mathbb{C}(x)$ then it does not define a function $\mathbb{C} \to \mathbb{C}$ because it is not defined at $x = i$ or $x = -i$. We won't worry too much about this.

5.3 Partial Fractions

In this section we will prove the general theorem on partial fractions in Euclidean domains, and relate this to the Chinese Remainder Theorem from the previous chapter. First we prove a vector version of the Euclidean Algorithm.

Bézout's Identity for Vectors

Let R be a Euclidean domain with size function $N : R\backslash\{0\} \to \mathbb{N}$. For any nonzero elements $a_1, \ldots, a_n \in R$ we consider the set of common divisors:

$$\text{Div}(a_1, \ldots, a_n) = \{d \in R : d|a_i \text{ for all } i\}.$$

Since $d|a_i$ and $a_i \neq 0$ imply $N(d) \leqslant N(a_i)$, the set $\text{Div}(a_1, \ldots, a_n)$ has some element of

[50] Proof: If $f(x)$ and $g(x)$ define the same function then $f(x) - g(x)$ is a polynomial with infinitely many roots, hence it is the zero polynomial.

maximum size. If d is such an element, I claim that there exist $x_1, \ldots, x_n \in R$ satisfying

$$d = a_1x_1 + \cdots + a_nx_n.$$

Proof. To prove this, we consider the set of R-linear combinations

$$a_1R + \cdots + a_nR = \{a_1x_1 + \cdots a_nx_n : x_i \in R \text{ for all } i\}.$$

Since $a \neq 0$ implies $N(a) > 0$, this set has some nonzero element e of minimum size, and we will write $e = a_1x_1 + \cdots a_nx_n$ for some $x_i \in R$. If we can show that $d \sim e$ are associates then we will have $d = ue$ for some unit $u \in R^\times$ and hence

$$d = a_1(ux_1) + \cdots + a_n(ux_n)$$

as desired.

To prove that $d \sim e$ it is enough to show that $d|e$ and $N(d) = N(e)$, since we know from a previous homework that the maximum sized divisors of e are just the associates of e. To show that $d|e$ let us write $dk_i = a_i$ for some $k_i \in R$, which is possible because d is a common divisor of $a_1, \ldots, a_n$. Then we have

$$\begin{aligned} e &= a_1x_1 + \cdots + a_nx_n \\ &= d_1k_1x_1 + \cdots + d_nk_nx_n \\ &= d(k_1x_1 + \cdots + k_nx_n), \end{aligned}$$

which implies that $d|e$ and hence also $N(d) \leqslant N(e)$.

Next we will show that e is a common divisor of $a_1, \ldots, a_n$, from which it will follow that $N(e) \leqslant N(d)$ because d is a maximum sized common divisor. To show that $e|a_i$ for all i, we use the Division Theorem two find $q_i, r_i \in R$ satisfying

$$\begin{cases} a_i = eq_i + r_i, \\ r = 0 \text{ or } N(r_i) < N(e). \end{cases}$$

If $r_i \neq 0$ then we must have $N(r_i) < N(e)$. But this leads to a contradiction because

$$\begin{aligned} r_i &= a_i - eq_i \\ &= a_i - (a_1x_1 + \cdots a_nx_n)q_i \\ &= a_1(-x_1q_i) + \cdots + a_i(1 - x_iq_i) + \cdots + a_n(-x_nq_i) \end{aligned}$$

is an element of $a_1R + \cdots a_nR$ and e is supposed to be an element of this set with minimum size. Therefore we must have $r_i = 0$ for all i. □

If d is a greatest common divisor of elements $a_1, \ldots, a_n$ in a Euclidean domain then we have just proved that there exist elements $x_1, \ldots, x_n \in R$ satisfying

$$d = a_1x_1 + \cdots + a_nx_n.$$

But we have not yet given an algorithm to find such elements.

The Euclidean Algorithm for Vectors

Let R be a Euclidean domain. For any nonzero elements $a_1,\ldots,a_n \in R$ we will write $\gcd(a_1,\ldots,a_{n-1})$ to denote some maximum sized common divisor of $a_1,\ldots,a_{n-1}$. I claim that we have the following equality of sets:

$$\text{Div}(a_1,\ldots,a_n) = \text{Div}(\gcd(a_1,\ldots,a_{n-1}),a_n).$$

If we can show this then it will follow by induction that

$$\text{Div}(a_1,\ldots,a_n) = \text{Div}(d)$$

for some element $d \in R$, which will imply that the maximum sized common divisors of $a_1,\ldots,a_n$ are just the associates of d. In other words:

The GCD of $a_1,\ldots,a_n$ is unique up to multiplication by units.

To prove the equality of sets, we first use Bézout's Identity to write $e = \gcd(a_1,\ldots,a_{n-1}) = a_1x_1 + \cdots a_{n-1}x_{n-1}$ for some elements $x_i \in R$. Now let d be an element of the right set so that $dk_i = a_i$ for some elements $k_1,\ldots,k_n \in R$. Then we have

$$e = dk_1x_1 + \cdots dk_{n-1}x_{n-1} = d(k_1x_1 + \cdots k_{n-1}x_{n-1}),$$

so that $d|e$. Since we also have $d|a_n$ it follows that d is an element of the right set. Conversely, let d be an element of the right set so that $d|e$ and $d|a_n$. Since e is a common divisor of $a_1,\ldots,a_{n-1}$ we can write $e\ell_i = a_i$ for some $\ell_1,\ldots,\ell_{n-1} \in R$ and since $d|e$ we can write $dk = e$ for some $k \in R$. It follows that $a_i = e\ell_i = dk\ell_i = d(k\ell_i)$ and hence $d|a_i$ for all i from 1 to $n-1$. Since we also have $d|a_n$ it follows that d is an element of the left set.

This theorem allows us to use the notation $\gcd(a_1,\ldots,a_n)$ without confusion since the GCD is essentially unique. Then the equality of sets

$$\text{Div}(a_1,\ldots,a_n) = \text{Div}(\gcd(a_1,\ldots,a_{n-1}),a_n).$$

implies the equality (up to units) of greatest common divisors:

$$\gcd(a_1,\ldots,a_n) = \gcd\left(\gcd(a_1,\ldots,a_{n-1}),a_n\right).$$

As the title of the theorem implies, we can turn this identity into a recursive algorithm to find elements $x_1,\ldots,x_n \in R$ satisfying

$$\gcd(a_1,\ldots,a_n) = a_1x_1 + \cdots + a_nx_n.$$

In the base case $n = 2$ we can just use the Extended Euclidean Algorithm from Chapter 3. For $n \geqslant 3$, let us assume that we have already found $x_1', \ldots, x_n' \in R$ satisfying

$$\gcd(a_1, \ldots, a_{n-1}) = a_1 x_1' + \cdots + a_{n-1} x_{n-1}'.$$

We can also use the Extended Euclidean Algorithm to find $x, y \in R$ such that

$$\gcd(a_1, \ldots, a_n) = \gcd(a_1, \ldots, a_{n-1})x + a_n y.$$

Then putting these together gives

$$\begin{aligned} \gcd(a_1, \ldots, a_n) &= \gcd(a_1, \ldots, a_{n-1})x + a_n y \\ &= (a_1 x_1' + \cdots + a_{n-1} x_{n-1}')x + a_n y \\ &= a_1(x_1' x) + \cdots a_{n-1}(x_{n-1}' x) + a_n y, \end{aligned}$$

as desired.

Let's compute an example. Consider the numbers $a_1 = 35$, $a_2 = 63$ and $a_3 = 45$, which satisfy

$$\gcd(35, 63, 45) = \gcd(\gcd(35, 63), 45) = \gcd(7, 45) = 1.$$

Since $\gcd(35, 63) = 7$ we begin by looking for $x, y \in \mathbb{Z}$ such that $735x + 63y$:

$$\begin{array}{r|c|l} 0 & 1 & 63 \\ 1 & 0 & 35 \\ -1 & 1 & 28 \\ 2 & -1 & 7 \end{array}$$

We find that $7 = 35(2) + 63(-1)$. Then since $\gcd(7, 45) = 1$ we look for $x, y \in \mathbb{Z}$ such that $1 = 7x + 45y$:

$$\begin{array}{r|c|l} 0 & 1 & 45 \\ 1 & 0 & 7 \\ -6 & 1 & 3 \\ 13 & -2 & 1 \end{array}$$

We find that $1 = 7(13) + 45(-2)$. Then combining the two equations gives

$$\begin{aligned} 1 &= 7(13) + 45(-2) \\ &= [35(2) + 63(-1)]\,(13) + 45(-2) \\ &= 35(26) + 63(-13) + 45(-2). \end{aligned} \tag{$*$}$$

I have secretly chosen this example to also provide an introduction to partial fractions. Note that the integer 315 has prime factorization

$$351 = 3^3 \cdot 5 \cdot 7.$$

The general idea of partial fractions is that a factorization of a denominator leads to a sum of fractions. In the case where the denominator is 315 we will be able to write

$$\frac{1}{315} = \frac{A}{3^2} + \frac{B}{3^1} + \frac{C}{5} + \frac{D}{7} + E,$$

for some integers $A, B, C, D, E \in \mathbb{Z}$ satisfying $0 \leqslant A, B < 3$, $0 \leqslant C < 5$ and $0 \leqslant D < 7$.[51] How can we find these integers? It turns out that the hard work has already been done. First we divide the previous equation $(*)$ by 315 to obtain

$$\begin{aligned}\frac{1}{315} &= \frac{35(26) + 63(-13) + 45(-2)}{315}\\ &= \frac{26}{9} + \frac{-13}{5} + \frac{-2}{7}\end{aligned}$$

Thus we have separated the fraction into its "coprime parts". Next we have to clean things up. For each fraction of the form a/p^k with p prime, we first divide a by p and then we successively divide each quotient by p to obtain

$$\begin{aligned} a &= pq_1 + r_1 & 0 &\leqslant r_1 < p,\\ q_1 &= pq_2 + r_2 & 0 &\leqslant r_2 < p,\\ q_2 &= pq_3 + r_3 & 0 &\leqslant r_3 < p,\\ &\vdots \\ q_{k-1} &= pq_k + r_k & 0 &\leqslant r_k < p. \end{aligned}$$

Then putting everything together gives

$$\begin{aligned} a &= r_1 + q_1 p\\ &= r_1 + r_2 p + q_2 p^2\\ &= r_1 + r_2 p + r_3 p^2 + q_3 p^3\\ &\vdots\\ &= r_1 + r_2 p + r_3 p^2 + \cdots + r_k p^{k-1} + q_k p^k \end{aligned}$$

and hence

$$\frac{a}{p^k} = \frac{r_1}{p^k} + \frac{r_2}{p^{k-1}} + \cdots + \frac{r_k}{p} + q_k.$$

Applying this to the partial fractions in our example gives

$$\begin{aligned} 26/9 &= 2/9 + 2/3 + 2,\\ -13/5 &= 2/5 - 3,\\ -2/7 &= 5/7 - 1, \end{aligned}$$

[51] In fact, one can show that the integers A, B, C, D are unique, and that $E = 0$. But we will not prove this because these properties do not generalize to other Euclidean domains. See *Partial fractions in Euclidean domains* by Packard and Wilson.

and then adding these gives

$$\frac{1}{351} = \left(\frac{2}{9} + \frac{2}{3} + 2\right) + \left(\frac{2}{5} - 3\right) + \left(\frac{5}{7} - 1\right) = \frac{2}{9} + \frac{2}{3} + \frac{2}{5} + \frac{5}{7} - 2,$$

which has the desired form.

The general story for Euclidean domains works exactly the same way. The most difficult part is to show that we can always find an equation similar to $(*)$ above. This is established by the following slightly tricky lemma. (We will use this same lemma in the next section to generalize the Chinese Remainder Theorem to multiple moduli.)

Lemma for Partial Fractions and CRT

Let R be a Euclidean domain and consider some elements $n_1, \ldots, n_k \in R$ such that $\gcd(n_i, n_j) = 1$ for all $i \neq j$. (We say that these elements are *pairwise coprime.*) Now for each element $1 \leqslant i \leqslant k$ we consider the element

$$\hat{n}_i = n_1 \cdots n_{i-1} n_{i+1} \cdots n_k \in R.$$

In this case I claim that the elements $\hat{n}_1, \ldots, \hat{n}_k$ are *jointly coprime* (which is a weaker condition[52] than being pairwise coprime):

$$\gcd(\hat{n}_1, \hat{n}_2, \ldots, \hat{n}_k) = 1.$$

The general proof is hard to write down, so we first consider the smallest cases. When $k = 2$ we have $\hat{n}_1 = n_2$ and $\hat{n}_2 = n_1$, so that $\gcd(\hat{n}_1, \hat{n}_2) = \gcd(n_2, n_1) = \gcd(n_1, n_2) = 1$, as desired. When $k = 3$ we have $\hat{n}_1 = n_2 n_3$, $\hat{n}_2 = n_1 n_3$ and $\hat{n}_3 = n_1 n_2$, and our goal is to show that

$$\left\{ \begin{array}{rcl} \gcd(n_1, n_2) & = & 1 \\ \gcd(n_1, n_3) & = & 1 \\ \gcd(n_2, n_3) & = & 1 \end{array} \right\} \implies \gcd(n_2 n_3, n_1 n_3, n_1 n_2) = 1.$$

To this end, suppose for contradiction that there exists a common prime divisor p of $n_2 n_3$, $n_1 n_3$ and $n_1 n_2$. Now there are two cases:

- Suppose that $p \nmid n_3$. Since p is prime with $p | n_1 n_3$ and $p | n_2 n_3$ we must have $p | n_1$ and $p | n_2$, which gives the contradiction $\gcd(n_1, n_2) \neq 1$.
- Suppose that $p | n_3$. Since p is prime and $p | n_1 n_2$ we must also have $p | n_1$ or $p | n_2$. If $p | n_1$ then we obtain the contradiction that $\gcd(n_1, n_3) \neq 1$ and if $p | n_2$ then we obtain the contradiction that $\gcd(n_2, n_3) \neq 1$.

[52] Consider the integers $2, 3, 4$. These are jointly coprime because they have no common prime divisor. But they are not pairwise coprime because $\gcd(2, 4) \neq 1$.

In any case, we obtain a contradiction, which proves that n_1n_2, n_1n_3, n_1n_2 have no common prime divisor, as desired. The general case is the same but the notation becomes a mess.

Proof of the Lemma. We have already shown that the statement holds for k less than 4. So let us assume that $k \geqslant 4$ and consider some elements $n_1, \ldots, n_k$ and $\hat{n}_1, \ldots, \hat{n}_k$ as in the statement of the lemma. In order to use induction, we also define elements $\tilde{n}_1, \ldots, \hat{n}_{k-1}$ by

$$\tilde{n}_i = n_1 \ldots n_{i-1} n_{i+1} \cdots n_{k-1}.$$

Since the elements $n_1, \ldots, n_k$ are pairwise coprime, so are the elements $n_1, \ldots, n_{k-1}$. Thus by induction we may assume that $\gcd(\tilde{n}_1, \ldots, \tilde{n}_{k-1}) = 1$. In order to show that $\gcd(\hat{n}_1, \ldots, \hat{n}_k) = 1$, we assume for contradiction that there exists a prime element p such that $p|\hat{n}_i$ for all i. There are two cases:

- Suppose that $p \nmid n_k$ and observe that $\hat{n}_i = \tilde{n}_i n_k$ for all $1 \leqslant i \leqslant k-1$. Since p is prime and $p|\hat{n}_i$ for all $1 \leqslant i \leqslant k-1$, it follows that $p|\tilde{n}_i$ for all $1 \leqslant i \leqslant k-1$, which contradicts the fact that $\gcd(\tilde{n}_1, \ldots, \tilde{n}_{k-1}) = 1$.
- Suppose that $p|n_k$ and observe that $\hat{n}_k = n_1 n_2 \cdots n_{k-1}$. Since p is prime and $p|\hat{n}_k$ this implies that $p|n_i$ for some $1 \leqslant i \leqslant k-1$, which gives the contradiction $\gcd(n_i, n_k) \neq 1$.

In any case, we obtain a contradiction, which proves that $\hat{n}_1, \ldots, \hat{n}_k$ have no common prime divisor, as desired. □

Theorem of Partial Fractions

Let R be a Euclidean domain with size function $N : R\backslash\{0\} \to \mathbb{N}$ and consider any nonzero, nonunit element $n \in R$. Suppose that n has unique prime factorization

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k},$$

for some distinct primes $p_1, \ldots, p_k$. Then we can write

$$\frac{1}{n} = m + \left(\frac{r_{1,1}}{p_1} + \frac{r_{1,2}}{p_1^2} + \cdots + \frac{r_{1,e_1}}{p_1^{e_1}} \right) + \cdots + \left(\frac{r_{k,1}}{p_k} + \frac{r_{k,2}}{p_k^2} + \cdots + \frac{r_{k,e_k}}{p_k^{e_k}} \right),$$

for some elements $m, r_{i,j} \in R$ where $r_{i,j} = 0$ or $N(r_{i,j}) < N(p_i)$.

Proof of the Theorem. Let $n_i = p_i^{e_i}$ for all $1 \leqslant i \leqslant k$, so that $n = n_1 n_2 \cdots n_k$, and define

$$\hat{n}_i = n_1 \cdots n_{i-1} n_{i+1} \cdots n_k,$$

which is the unique element of R satisfying $\hat{n}_i n_i = n$. Since the primes p_i, p_j are distinct we have $\gcd(n_i, n_j) = 1$ for all $i \neq j$ and it follows from the lemma that

$$\gcd(\hat{n}_1, \ldots, \hat{n}_k) = 1.$$

It follows from Bézout's Identity for Vectors that there exist some $x_1, \ldots, x_k \in R$ satisfying

$$1 = x_1\hat{n}_1 + x_2\hat{n}_2 + \cdots + x_k\hat{n}_k.$$

Hence we can divide both sides by n to obtain

$$\frac{1}{n} = x_1\frac{\hat{n}_1}{n} + \cdots + x_k\frac{\hat{n}_k}{n} = \frac{x_1}{n_1} + \cdots + \frac{x_k}{n_k}.$$

Finally, we consider the fraction $x_i/n_i = x_i/p_i^{e_i}$. By dividing x_i by p_i and then successively dividing each quotient by p_i (as in the example above), we can write

$$\frac{x_i}{p_i^{e_i}} = q_{i,e_i} + \frac{r_{i,1}}{p_i} + \frac{r_{i,2}}{p_i^2} + \cdots + \frac{r_{i,e_i}}{p_i^{e_i}}.$$

for some elements $q_{i,e_i}, r_{i,1}, r_{i,2}, \ldots, r_{i,e_i} \in R$ with $r_{i,j} = 0$ or $N(r_{i,j}) < N(p_i)$. Then adding all of these expressions together gives the desired result, with $m = \sum_{i=1}^{k} q_{i,e_i} \in R$. □

We gave a motivating example in the ring $\mathbb{Z}$, but the main applications of this theorem come from rings of polynomials $\mathbb{F}[x]$. Let's consider the case when $\mathbb{F} = \mathbb{R}$.

Let $f(x) \in \mathbb{R}[x]$ be a non-constant polynomial with real coefficients and suppose we can write

$$f(x) = p_1(x)^{d_1} \cdots p_k(x)^{d_k} q_1(x)^{e_1} \cdots q_\ell(x)^{e_\ell},$$

where $p_i(x), q_j(x) \in \mathbb{R}[x]$ are irreducible over $\mathbb{R}$ with $\deg(p_i) = 1$ and $\deg(q_j) = 2$ for all i, j. In this case, the theorem of partial fractions tells us that

$$\frac{1}{f(x)} = g(x) + \sum_{i=1}^{k}\sum_{j=1}^{d_i} \frac{a_{i,j}}{p_i(x)^j} + \sum_{i=1}^{\ell}\sum_{j=1}^{e_i} \frac{b_{i,j} + c_{i,j}x}{q_i(x)^j},$$

for some polynomial $g(x) \in \mathbb{R}[x]$ and some real numbers $a_{i,j}, b_{i,j}, c_{i,j} \in \mathbb{R}$. One can show that each term in this sum can be integrated in terms of log and arctan, as Leibniz knew. In the next chapter we will prove that every real polynomial can indeed be factored in this way.

5.4 Generalized Chinese Remainder Theorem

To end this chapter, we show that the theorem of partial fractions is intimately related to the Chinese Remainder Theorem. Recall the system of congruences from Master Sun's Mathematical Manual:

$$\begin{cases} c \equiv 2 \bmod 3, \\ c \equiv 3 \bmod 5, \\ c \equiv 2 \bmod 7. \end{cases}$$

We previously solved this by combining the congruences two-by-two. The technology developed in the previous section will now allow us to give a more elegant one-step solution.

Generalized Chinese Remainder Theorem

Consider some positive integers $n_1, \ldots, n_k \in \mathbb{Z}$ and let $n = n_1 \cdots n_k$. Then we have a ring homomorphism defined as follows:

$$\begin{array}{rccc} \varphi : & \mathbb{Z}/n\mathbb{Z} & \to & \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z} \\ & a \bmod n & \mapsto & (a \bmod n_1, \ldots, a \bmod n_k). \end{array}$$

If $\gcd(n_i, n_j) = 1$ for all $i \neq j$ then this homomorphism is invertible. Furthermore, if we define $\hat{n}_i = n_1 \cdots n_{i-1} n_{i+1} \cdots n_k$ as above then from the previous section we know that there exist some (non-unique) integers $x_1, \ldots, x_k \in \mathbb{Z}$ such that

$$\hat{n}_1 x_1 + \hat{n}_2 x_2 + \cdots + \hat{n}_k x_k = 1.$$

In this case, I claim that we can compute the inverse of φ as follows:

$$\varphi^{-1}(a_1, a_2, \ldots, a_k) = a_1 \hat{n}_1 x_1 + \cdots + a_k \hat{n}_k x_k \bmod n.$$

Please compare this to the formula $\varphi^{-1}(a, b) = any + bmx$ in the case of two moduli.

Proof. First we observe that φ is well-defined. Indeed, suppose that $a \equiv b \bmod n$, so that $n|(a-b)$. Then since $n_i|n$ we must have $n_i|(a-b)$ and hence $a \equiv b \bmod n_i$ for all i. In other words, we have $\varphi(a) = \varphi(b)$.

Now suppose that $\gcd(n_i, n_j) = 1$ for all $i \neq j$, so from the previous section we know that $\hat{n}_1 x_1 + \cdots + \hat{n}_k x_k = 1$ for some integers $x_1, \ldots, x_k$. We can use this to prove that φ is injective, as follows. First suppose that $\varphi(a) = \varphi(b)$ so that $a \equiv b \bmod n_i$ for all i. Let's say $n_i d_i = a - b$ for some integers d_i. But then we have

$$\begin{aligned} (a-b) &= (\hat{n}_1 x_1 + \hat{n}_2 x_2 + \cdots + \hat{n}_k x_k)(a-b) \\ &= \hat{n}_1 x_1 (a-b) + \hat{n}_2 x_2 (a-b) + \cdots + \hat{n}_k x_k (a-b) \\ &= \hat{n}_1 x_1 n_1 d_1 + \hat{n}_2 x_2 n_2 d_2 + \cdots + \hat{n}_k x_k n_k d_k \\ &= (\hat{n}_1 n_1) x_1 d_1 + (\hat{n}_2 n_2) x_2 d_2 + \cdots + (\hat{n}_k n_k) x_k d_k \\ &= n x_1 d_1 + n x_2 d_2 + \cdots + n x_k d_k \\ &= n(x_1 d_1 + x_2 d_2 + \cdots + x_k d_k), \end{aligned}$$

so that $a \equiv b \bmod n$. Then since φ is an injective function between sets of the same size it must be invertible.

More precisely, I claim that $\varphi^{-1}(a_1, \ldots, a_k) = a_1 \hat{n}_1 x_1 + \cdots + a_k \hat{n}_k x_k$. To see this, we first observe hat $n_j | \hat{n}_i$ and hence $\hat{n}_i \equiv 0 \bmod n_j$ for all $i \neq j$. Furthermore, we have $\hat{n}_i x_i = 1 - \sum_{i \neq j} \hat{n}_j x_j \equiv 1 - 0 \equiv 1 \bmod n_i$ for all i. Finally, we conclude that

$$a_1 \hat{n}_1 x_1 + \cdots + a_k \hat{n}_k x_k \equiv a_1 0 + \cdots a_{i-1} 0 + a_i 1 + a_{i+1} 0 + \cdots + a_k 0 \equiv a_i \bmod n_i$$

for all i, as desired. □

To see how this works, we apply it to Master Sun's system of congruences. Let $(n_1, n_2, n_3) = (3, 5, 7)$ so that $(\hat{n}_1, \hat{n}_2, \hat{n}_3) = (35, 21, 15)$. Since $(3, 5, 7)$ are pairwise coprime it follows that $(35, 21, 15)$ are jointly coprime, so there exist $x_1, x_2, x_3 \in \mathbb{Z}$ satisfying $35x_1 + 21x_2 + 15x_3 = 1$. In order to find such x_1, x_2, x_3 we must use a recursive method. To be precise, we will use the Extended Euclidean Algorithm and the fact that

$$\gcd(35, 21, 15) = \gcd(\gcd(35, 21), 15) = \gcd(7, 15) = 1.$$

First we find some $x', y' \in \mathbb{Z}$ such that $7x' + 15y' = 1$:

$$\begin{array}{r|c|l} 0 & 1 & 15 \\ 1 & 0 & 7 \\ -2 & 1 & 1 \end{array}$$

We see that $7(-2) + 15(1) = 1$. Then we find some $x'', y'' \in \mathbb{Z}$ such that $35x'' + 21y'' = 7$:

$$\begin{array}{r|c|l} 1 & 0 & 35 \\ 0 & 1 & 21 \\ 1 & -1 & 14 \\ -1 & 2 & 7 \end{array}$$

We see that $35(-1) + 21(2) = 7$. Then we put these together to obtain

$$1 = 7(-2) + 15(1) = [35(-1) + 21(2)]\,(-2) + 15(1) = 35(2) + 21(-4) + 15(1).$$

Thus we can take $(x_1, x_2, x_3) = (2, -4, 1)$. Finally, since $(\hat{n}_1 x_1, \hat{n}_2 x_2, \hat{n}_3 x_3) = (70, -84, 15)$, we obtain an explicit description for the inverse of φ:

$$\begin{array}{rccc} \varphi^{-1}: & \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} & \to & \mathbb{Z}/105\mathbb{Z} \\ & (a_1, a_2, a_3) & \mapsto & 70a_1 - 84a_2 + 15a_3. \end{array}$$

In Master Sun's case we have $(a_1, a_2, a_3) = (2, 3, 2)$, so that

$$\varphi^{-1}(2, 3, 2) = 70(2) - 84(3) + 15(2) \equiv 23 \bmod 105.$$

In other words, we have

$$\left\{ \begin{array}{ccc} c & \equiv & 2 \bmod 3 \\ c & \equiv & 3 \bmod 5 \\ c & \equiv & 2 \bmod 7 \end{array} \right\} \Leftrightarrow \varphi(c) = (2, 3, 2) \Leftrightarrow c = \varphi^{-1}(2, 3, 2) \equiv 23 \bmod 105.$$

Of course, we solved this system before. The advantage of the new method is that we can tweak the input $(a_1, a_2, a_3) = (2, 3, 2)$ without doing the work again. For example, since $70(2) - 84(4) + 15(2) \equiv 44 \bmod 105$ we have

$$\left\{ \begin{array}{ccc} c & \equiv & 2 \bmod 3 \\ c & \equiv & 4 \bmod 5 \\ c & \equiv & 2 \bmod 7 \end{array} \right\} \Leftrightarrow c \equiv 44 \bmod 105.$$

6 The Fundamental Theorem of Algebra

6.1 Equivalent Statements of the FTA

The goal of this chapter is to prove the Fundamental Theorem of Algebra (FTA). The original statement of the theorem (mentioned in Section 5.1) claims that every non-constant polynomial $f(x) \in \mathbb{R}[x]$ can be expressed as

$$f(x) = p_1(x)p_2(x)\cdots p_k(x),$$

where $p_i(x) \in \mathbb{R}[x]$ and $\deg(p_i) = 1$ or 2. As we have seen, if this version of the FTA is true then any rational expression can be integrated in terms of log and arctan. The proof is quite involved, and will require an entire chapter to understand.

In this section we seek to increase our understanding of the **statement** of the FTA. To this end we will prove the equivalence of several different statements.

Equivalent Statements of the FTA

The following six statements are logically equivalent:

(1$\mathbb{R}$) Every non-constant $f(x) \in \mathbb{R}[x]$ has a root in $\mathbb{C}$.

(2$\mathbb{R}$) Every non-constant $f(x) \in \mathbb{R}[x]$ can be expressed as

$$f(x) = p_1(x)p_2(x)\cdots p_k(x),$$

where $p_i(x) \in \mathbb{R}[x]$ and $\deg(p_i) = 1$ or 2.

(3$\mathbb{R}$) Every prime element of $\mathbb{R}[x]$ has degree 1 or 2.

(1$\mathbb{C}$) Every non-constant $f(x) \in \mathbb{C}[x]$ has a root in $\mathbb{C}$.

(2$\mathbb{C}$) Every non-constant $f(x) \in \mathbb{C}[x]$ splits over $\mathbb{C}$.

(3$\mathbb{C}$) Every prime element of $\mathbb{C}[x]$ has degree 1.

It is straightforward to prove that the three statements (1$\mathbb{C}$), (2$\mathbb{C}$) and (3$\mathbb{C}$) are equivalent. We will refer to any of these three as the $\mathbb{C}$FTA.

Proof (Equivalent Forms of $\mathbb{C}$FTA).

(1$\mathbb{C}$)$\Rightarrow$(2$\mathbb{C}$): Consider some non-constant $f(x) \in \mathbb{C}[x]$. By assumption there exists $\alpha_1 \in \mathbb{C}$ such that $f(\alpha_1) = 0$, hence by Descartes' Theorem we can write

$$f(x) = (x - \alpha_1)g(x)$$

for some $g(x) \in \mathbb{C}[x]$. If $g(x)$ is constant then we are done. Otherwise, there exists some $\alpha_2 \in \mathbb{C}$ such that $g(\alpha_2) = 0$. Then by Descartes' Theorem we have $g(x) = (x - \alpha_2)h(x)$ and hence

$$f(x) = (x - \alpha_1)g(x) = (x - \alpha_1)(x - \alpha_2)h(x).$$

By continuing in this way[53] we conclude that $f(x)$ splits over $\mathbb{C}$.

(2$\mathbb{C}$)⇒(3$\mathbb{C}$): Let $p(x)$ be a prime element of $\mathbb{C}[x]$. Since units are not prime we know that $p(x)$ is non-constant. Hence by assumption we can write

$$p(x) = c(x - \alpha_1) \cdots (x - \alpha_n)$$

for some $c, \alpha_1, \ldots, \alpha_n \in \mathbb{C}$. Since $p(x)$ divides the product $\prod_i (x - \alpha_i)$, and since $p(x)$ is prime, we know from Euclid's Lemma that $p(x)|(x - \alpha_i)$ for some i. It follows that $\deg(p) \leqslant \deg(x - \alpha_i) = 1$, which implies that $\deg(p) = 1$.

(3$\mathbb{C}$)⇒(1$\mathbb{C}$): Every non-constant $f(x) \in \mathbb{C}[x]$ has a unique prime factorization in $\mathbb{C}[x]$:

$$f(x) = p_1(x)p_2(x) \cdots p_k(x).$$

By assumption, each prime $p_i(x)$ has degree 1. In particular, we have $p_1(x) = ax + b$ for some $a, b \in \mathbb{C}$ with $a \neq 0$, and hence $-b/a \in \mathbb{C}$ is a root of $f(x)$. □

The equivalence of the statements (1$\mathbb{R}$), (2$\mathbb{R}$) and (3$\mathbb{R}$) is a bit less straightforward since it uses some properties of complex conjugation. We will refer to any of these three statements as the $\mathbb{R}$FTA. Our proof of equivalence will use the following lemma.

Lemma for the $\mathbb{R}$FTA

For any extension of fields $\mathbb{E} \supseteq \mathbb{F}$ we have an extension of rings $\mathbb{E}[x] \supseteq \mathbb{F}[x]$. If there exist $f(x), p(x) \in \mathbb{F}[x]$ and $q(x) \in \mathbb{E}[x]$ such that $f(x) = p(x)q(x)$ then I claim that in fact $q(x) \in \mathbb{F}[x]$.

Indeed, we know from the Division Theorem in $\mathbb{F}[x]$ that there exist $q'(x), r'(x) \in \mathbb{F}[x]$ satisfying $f(x) = p(x)q'(x) + r'(x)$ and $\deg(r') < \deg(p)$. But now we have $f(x) = p(x)q(x) + 0$ and $f(x) = p(x)q'(x) + r'(x)$ in the ring $\mathbb{E}[x]$ and it follows from the uniqueness of quotients in $\mathbb{E}[x]$ that $q(x) = q'(x) \in \mathbb{F}[x]$.

Proof (Equivalent Forms of $\mathbb{R}$FTA).

(1$\mathbb{R}$)⇒(2$\mathbb{R}$): Consider some non-constant $f(x) \in \mathbb{R}[x]$. By assumption there exists $\alpha \in \mathbb{C}$ such that $f(\alpha) = 0$. If $\alpha \in \mathbb{R}$ then by Descartes' Theorem we can write $f(x) = (x - \alpha)g(x)$ for

[53] We could also phrase this as a formal proof by induction.

some $g(x) \in \mathbb{R}[x]$. If $\alpha \notin \mathbb{C}$ then since the coefficients of $f(x)$ are real we also have $f(\alpha^*) = 0$ with $\alpha \neq \alpha^*$ and it follows from Descartes' Theorem that

$$f(x) = (x - \alpha)(x - \alpha^*)g(x)$$

for some $g(x) \in \mathbb{C}[x]$. But in fact I claim that $g(x) \in \mathbb{R}[x]$. To see this we let $p(x) = (x-\alpha)(x-\alpha^*) = x^2-(\alpha+\alpha^*)x+\alpha\alpha^*$, which has real coefficients. Then since $f(x) = p(x)g(x)$ with $f(x), p(x) \in \mathbb{R}[x]$ and $g(x) \in \mathbb{C}[x]$ we conclude from the Lemma that in fact $g(x) \in \mathbb{R}[x]$. In summary, we have shown that any non-constant $f(x) \in \mathbb{R}[x]$ satisfies $f(x) = p(x)g(x)$ for some $p(x), g(x) \in \mathbb{R}[x]$ with $\deg(p) = 1$ or 2. Now the result follows by induction.

(2ℝ)⇒(3ℝ): Let $p(x)$ be a prime element of $\mathbb{R}[x]$. Since units are not prime we know that $p(x)$ is non-constant. Hence we can write

$$p(x) = q_1(x) \cdots q_k(x),$$

where $q_i(x) \in \mathbb{R}[x]$ and $\deg(q_i) = 1$ or 2 for all i. Since $p(x)$ divides the product $\prod_i q_i(x)$, and since $p(x)$ is prime, we know from Euclid's Lemma that $p(x)|q_i(x)$ for some i. It follows that $\deg(p) \leqslant \deg(q_i)$, which implies that $\deg(p) = 1$ or 2.

(3ℝ)⇒(1ℝ): Every non-constant $f(x) \in \mathbb{R}[x]$ has a unique prime factorization in $\mathbb{R}[x]$:

$$f(x) = p_1(x)p_2(x) \cdots p_k(x).$$

By assumption, each prime $p_i(x)$ has degree 1 or 2. If there exists a factor $p_i(x)$ of degree 1, say $p_i(x) = ax + b$ then $f(x)$ has the root $-b/a \in \mathbb{R}$, which is also an element of $\mathbb{C}$. Otherwise, every factor $p_i(x)$ has degree 2. But we know from the quadratic formula that any quadratic polynomial with real coefficients has a root in $\mathbb{C}$. Hence $f(x)$ has a root in $\mathbb{C}$. □

It is more surprising that the real and complex forms of the FTA are also equivalent. To prove this we need another trick.

Lemma for the Equivalence of ℝFTA and ℂFTA

The field extension $\mathbb{C} \supseteq \mathbb{R}$ gives us a ring extension $\mathbb{C}[x] \supseteq \mathbb{R}[x]$. Recall that $\mathbb{R}$ can be viewed as the set of complex numbers $\alpha \in \mathbb{C}$ satisfying $\alpha^* = \alpha$. Now we we will define a similar conjugation operation on polynomials $* : \mathbb{C}[x] \to \mathbb{C}[x]$ such that $\mathbb{R}[x]$ can be viewed as the set of self-conjugate polynomial. To be specific, for any polynomial $f(x) = \sum_k \alpha_k x^k \in \mathbb{C}[x]$ with complex coefficients, we define

$$f^*(x) := \sum_k \alpha_k^* x^k.$$

Then we have the following properties:

(1) For all $f(x) \in \mathbb{C}[x]$ and $\beta \in \mathbb{C}$ we have $f(\beta)^* = f^*(\beta^*)$.

(2) For all $f(x) \in \mathbb{C}[x]$ we have $f(x) \in \mathbb{R}[x]$ if and only if $f^*(x) = f(x)$.

(3) For all $f(x), g(x) \in \mathbb{C}[x]$ we have

$$(f+g)^*(x) = f^*(x) + g^*(x) \quad \text{and} \quad (fg)^*(x) = f^*(x)g^*(x).$$

(4) For all $f(x) \in \mathbb{C}[x]$ we have $f(x) + f^*(x) \in \mathbb{R}[x]$ and $f(x)f^*(x) \in \mathbb{R}[x]$.

Proof of the Lemma. (1): Since $* : \mathbb{C} \to \mathbb{C}$ preserves all ring operations, we have

$$f(\beta)^* = \left(\sum_k \alpha_k \beta^k\right)^* = \sum_k \alpha_k^* (\beta^*)^k = f^*(\beta).$$

(2): Two formal polynomials are equal if and only if their coefficients are equal. The coefficient of x^k in $f(x)$ is α_k and the coefficient of x^k in $f^*(x)$ is α_k^*. If $f^*(x) = f(x)$ then we must have $\alpha_k^* = \alpha_k$, which implies that $\alpha_k \in \mathbb{R}$ for all k. In other words, we must have $f(x) \in \mathbb{R}[x]$.

(3): Let $f(x) = \sum_k \alpha_k x^k$ and $g(x) = \sum_k \beta_k x^k$. The coefficients of $f+g$ are $\alpha_k + \beta_k$, hence the coefficients of $(f+g)^*$ are $(\alpha_k + \beta_k)^* = \alpha_k^* + \beta_k^*$. But these are also the coefficients of $f^* + g^*$, hence $(f+g)(x) = f^*(x) + g^*(x)$. For the second statement, recall that

$$f(x)g(x) = \sum_k \left(\sum_{i+j=k} \alpha_i \beta_j\right) x^k.$$

So the coefficients of $(fg)^*(x)$ are

$$\left(\sum_{i+j=k} \alpha_i \beta_j\right)^* = \left(\sum_{i+j=k} \alpha_i^* \beta_j^*\right).$$

But these are also the coefficients of $f^*(x)g^*(x)$, hence $(fg)^*(x) = f^*(x)g^*(x)$.

(4): As we sometimes do, we will write f instead of $f(x)$ to save space. Let $f(x) \in \mathbb{C}[x]$. Then from part (3) we have

$$(f + f^*)^* = f^* + f^{**} = f^* + f = f + f^*$$

and

$$(ff^*)^* = f^* f^{**} = f^* f = ff^*,$$

hence it follows from part (2) that $f + f^* \in \mathbb{R}[x]$ and $ff^* \in \mathbb{R}[x]$. □

Proof (Equivalence of $\mathbb{R}$FTA and $\mathbb{C}$FTA).

Note that (1$\mathbb{C}$) trivially implies (1$\mathbb{R}$) because $\mathbb{R} \subseteq \mathbb{C}$, hence $\mathbb{C}$FTA implies $\mathbb{R}$FTA. To prove that $\mathbb{R}$FTA implies $\mathbb{C}$FTA, we will show that (1$\mathbb{R}$) also implies (1$\mathbb{C}$). So assume that (1$\mathbb{R}$)

is true and consider some non-constant polynomial $f(x) \in \mathbb{C}[x]$. It follows from Lemma (4) that $g(x) = f(x)f^*(x)$ has real coefficients, so from (1$\mathbb{R}$) there exists some $\alpha \in \mathbb{C}$ satisfying $g(\alpha) = 0$. Then since

$$0 = g(\alpha) = f(\alpha)f^*(\alpha)$$

we must have $f(\alpha) = 0$ or $f^*(\alpha) = 0$. If $f(\alpha) = 0$ then we are done since α is a complex root of $f(x)$. On the other hand, if $f^*(\alpha)$ then Lemma (1) implies

$$f(\alpha^*) = (f^*(\alpha))^* = 0^* = 0$$

so that α^* is a complex root of $f(x)$. □

6.2 Intermediate Value Theorem

In order to prove the FTA we need only prove one of the six equivalent statements from the previous section. In fact, we will prove statement (1$\mathbb{R}$):

Every non-constant $f(x) \in \mathbb{R}[x]$ has a root in $\mathbb{C}$.

And we will do this using a strange sort of induction. For each non-constant $f(x) \in \mathbb{R}[x]$ we can write $\deg(f) = 2^k m$ for some unique integers k, m where $k \geqslant 0$ and m is odd. The idea is to prove by induction on k that $f(x)$ has a root in $\mathbb{C}$. There are two important steps:

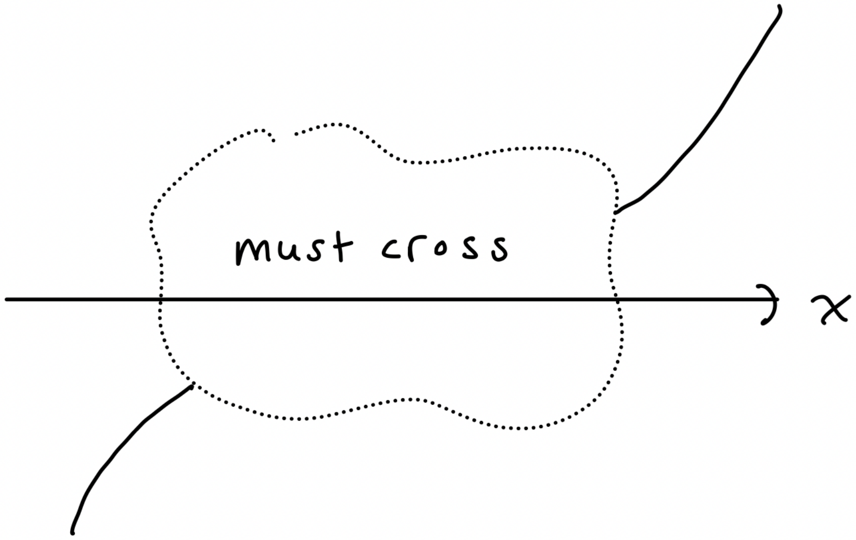
- **The Base Case $k = 0$.** Prove that every polynomial in $\mathbb{R}[x]$ of odd degree has a root in $\mathbb{C}$. In fact, we will show that it has a root in $\mathbb{R}$.
- **The Induction Step.** Assuming that every real polynomial of degree 2^k(odd) has a root in $\mathbb{C}$, prove that every real polynomial of degree 2^{k+1}(odd) has a root in $\mathbb{C}$.

For the induction step we will use a very clever argument of Laplace. Laplace's proof uses some deeper facts about multivariable polynomials so we postpone it until the end of the chapter.

In this section we discuss the base case: polynomials of odd degree. Actually, this case is "obvious". If $f(x) \in \mathbb{R}[x]$ has odd degree then one of the following two cases must hold:

- $f(x) \to +\infty$ as $x \to +\infty$ and $f(x) \to -\infty$ as $x \to -\infty$
- $f(x) \to -\infty$ as $x \to +\infty$ and $f(x) \to +\infty$ as $x \to -\infty$

In either case, the graph of $f(x)$ must cross the x-axis at some point $(c, 0) \in \mathbb{R}^2$, so that $f(x)$ has a real root $f(c) = 0$. Here is a picture of the first case:



This argument was perfectly clear to 18th century mathematicians such as Euler and Lagrange. However, early 19th century mathematicians such as Bolzano and Cauchy became unsatisfied with arguments based on pictures and they looked for a more rigorous proof. And since a rigorous proof must be based on axioms, these efforts forced mathematicians to look for an axiomatic definition of the real numbers.

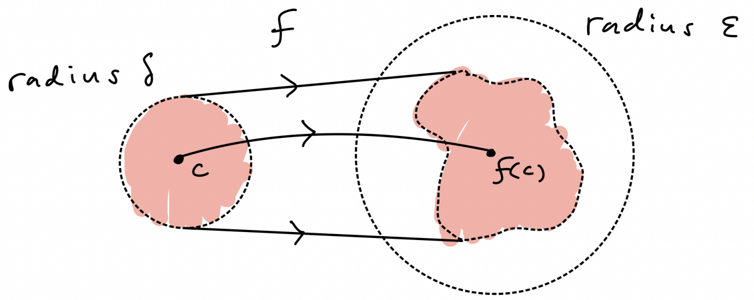
For my own benefit I will give a modern proof based on the standard axiomatization of the real numbers. My algebra students can feel free to ignore this.

Definition of Continuity

Intuition: A function $f : \mathbb{C} \to \mathbb{C}$ is called *continuous at* $c \in \mathbb{C}$ if $f(x) \to f(c)$ as $x \to c$.

Partial Formalization: We can make $f(x)$ as close to $f(c)$ as we please by taking x sufficiently close to c.

Cauchy's Definition: For any real number $\varepsilon > 0$ there exists some real number $\delta > 0$ such that $|x - c| < \delta$ implies $|f(x) - f(c)| < \varepsilon$.



The same definitions apply to real functions $f : \mathbb{R} \to \mathbb{R}$ but the picture is flat.

Cantor's Axiom for Real Numbers

We say that a sequence $c_1, c_2, \ldots \in \mathbb{C}$ *converges to a limit* $c \in \mathbb{C}$ if for any real number $\varepsilon > 0$ there exists an integer N such that $|c_n - c| < \varepsilon$ for all $n \geqslant N$.

We say that $c_1, c_2, c_3, \ldots \in \mathbb{R}$ is a *Cauchy sequence* if for any real number $\varepsilon > 0$ there exists an integer N such that for all $m, n \geqslant N$ we have $|c_n - c_m| < \varepsilon$.

Cantor constructed the real numbers $\mathbb{R}$ from the rational numbers $\mathbb{Q}$ by declaring that

every Cauchy sequence converges to some limit.

The Intermediate Value Theorem (IVT)

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function with $f(a) < 0$ and $f(b) > 0$ for some real numbers $a < b$. Then there exists at least one real number $c \in \mathbb{R}$ satisfying

- $a < c < b$,
- $f(c) = 0$.

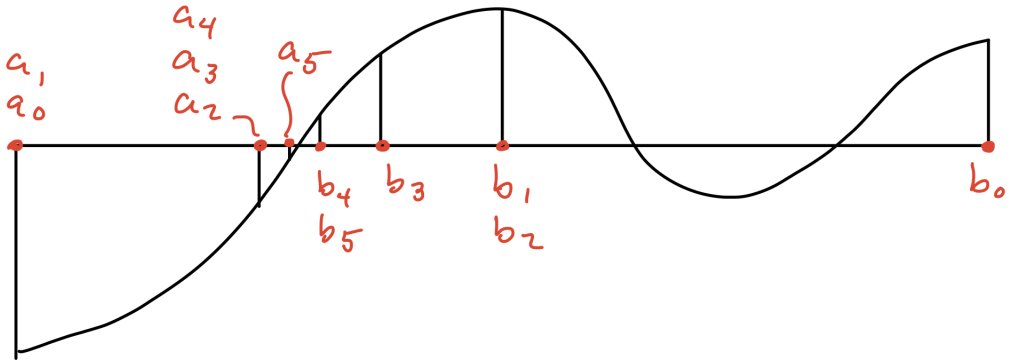
Proof. Define $(a_0, b_0) := (a, b)$ and $m_0 := (a_0 + b_0)/2$. If $f(m_0) = 0$ then we are done. Otherwise, define

$$(a_1, b_1) := \begin{cases} (a_0, m_0) & \text{if } f(m_0) > 0, \\ (m_0, b_0) & \text{if } f(m_0) < 0, \end{cases}$$

and $m_1 =: (a_1 + b_1)/2$. If $f(m_1) = 0$ then we are done. Otherwise we proceed to define

$$(a_{n+1}, b_{n+1}) := \begin{cases} (a_n, m_n) & \text{if } f(m_n) > 0, \\ (m_n, b_n) & \text{if } f(m_n) < 0. \end{cases}$$

At each step we divide the interval in half in such a way that there should still be a root inside the new smaller interval:



If we ever get $f(m_n) = 0$ then we are done. So let us assume that the process never terminates. Then by induction we must have

- $a_0 \leqslant a_1 \leqslant \cdots$,
- $\cdots \leqslant b_1 \leqslant b_0$,
- $a_n < b_n$ for all n,
- $f(a_n) < 0$ for all n,
- $f(b_n) > 0$ for all n.

I claim that $a_0, a_1, \ldots$ is a Cauchy sequence. To see this, let $\varepsilon > 0$ and let N be any integer greater than $\log_2((b-a)/\varepsilon)$. Then for any $m, n \geqslant N$ the numbers a_m, a_n lie in the closed interval $[a_N, b_N]$, so that

$$|a_m - a_n| \leqslant (b_N - a_N) = \frac{1}{2}(b_{N-1} - a_{N-1}) = \cdots = \frac{1}{2^N}(b_0 - a_0) = \frac{1}{2^N}(b-a) < \varepsilon,$$

as desired. It follows from Cantor's axiom that $a_0, a_1, \ldots$ converges to some limit $a' \in \mathbb{R}$. Similarly, the sequence $b_0, b_1, \ldots$ converges to some limit $b' \in \mathbb{R}$.

I claim that $a' = b'$. To see this we must show that $b' < a'$ and $a' < b'$ are impossible. If $b' < a'$ then let $\varepsilon = (a' - b')/2 > 0$ be half the length of the interval $[b', a']$. By definition of convergence we can find some some M, N such that $m \geqslant M$ implies $|a_m - a'| < \varepsilon$ and such that $n \geqslant N$ implies $|b_n - b'| < \varepsilon$. But then for any $\ell \geqslant \max\{M, N\}$ we see that b_ℓ is in the left half of the interval and a_ℓ is in the right half of the interval. This implies that $b_\ell < a_\ell$, which is a contradiction. And if $a' < b'$ then we let $\varepsilon = (b' - a')/3 > 0$ be one third the length of the interval $[a', b']$. Then we can find some integer[54] N so that $n \geqslant N$ implies that $|a_n - a'| < \varepsilon$,

[54] Technically, there exist K, L, M so that $k \geqslant K$ implies $|a_k - a'| < \varepsilon$, $\ell \geqslant L$ implies $|b_\ell - b'| < \varepsilon$ and $m \geqslant M$ implies $|a_m - b_m| < \varepsilon$. The existence of K, L follow from the definition of a', b' as limits. And we can let M be any integer larger than $\log_2((b-a)/\varepsilon)$, as in the earlier argument. Now let $N = \max\{K, L, M\}$.

$|b_n - b'| < \varepsilon$ and $|a_n - b_n| < \varepsilon$. But then we get a contradiction:

$$\begin{aligned} a' - b' &= (a' - a_n) + (a_n - b_n) + (b_n - b') \\ |a' - b'| &\leqslant |a' - a_n| + |a_n - b_n| + |b_n - b'| \\ 3\varepsilon &< \varepsilon + \varepsilon + \varepsilon. \end{aligned}$$

Hence we have shown that $a' = b'$. Call this common limit $c \in \mathbb{R}$.

It remains to show that $f(c) = 0$,[55] and for this we use the assumption that f is a continuous. If $f(c) > 0$ then let $\varepsilon = f(c)$. By continuity of f there exists some δ (depending on ε) so that $|x - c| < \delta$ implies $|f(x) - f(c)| < \varepsilon$. But then since $\lim a_i = c$ there exists some N (depending on δ, hence on ε) such that $n \geqslant N$ implies $|a_n - c| < \delta$, which implies $|f(a_n) - f(c)| < \varepsilon$. In other words, for any $n \geqslant N$ we have

$$f(c) - f(a_n) \leqslant |f(a_n) - f(c)| < f(c).$$

This implies that $f(a_n) > 0$, which contradicts the fact that $f(a_n) < 0$ for all n. Similarly, the assumption $f(c) < 0$ leads to a contradiction. Hence we conclude that $f(c) = 0$ as desired. □

Remark: Rigorous proofs in analysis are not really worth reading because they obscure all the ideas that led to the proof. For each paragraph above I discovered the appropriate bounds by drawing a picture of the number line. As I said, writing it down rigorously was only for my own benefit.

We need one more fact before completing the proof that each odd-degree real polynomial has at least one real root.

Polynomials are Continuous

Any polynomial $f(x) \in \mathbb{C}[x]$ determines a function $f : \mathbb{C} \to \mathbb{C}$ by evaluation. I claim that this function is continuous at every point in $\mathbb{C}$. The same proof will apply to real polynomials $f(x) \in \mathbb{R}[x]$ and real functions $f : \mathbb{R} \to \mathbb{R}$.

Often this result is proved inductively by showing that constant functions are continuous, the function $x \mapsto x$ is continuous, and sums/products of continuous functions are continuous. I prefer a more explicit method, which gives some additional useful information.

Proof. Let $f(x) \in \mathbb{C}[x]$ be non-constant and consider any point $c \in \mathbb{C}$. Applying Descartes' Factor Theorem gives

$$f(x) = (x - c)q(x) + f(c)$$

[55] I guess we also have to show that $a < c < b$. If $c < a$ then since $\lim b_i = c$ we would find some $b_n < a$, which contradicts the fact that $a \leqslant a_n < b_n$.

$$f(x) - f(c) = (x - c)q(x).$$

Thus for any complex number $x \in \mathbb{C}$ we have[56]

$$|f(x) - f(c)| = |x - c||q(x)|.$$

From this it is clear that $|f(x)-f(c)|$ goes to zero as $|x-c|$ goes to zero. To be more rigorous, let's consider the Taylor expansion of $q(x)$ around $x = c$:

$$q(x) = a_n(x - c)^n + \cdots + a_1(x - c) + a_0.$$

Here a_k is the kth derivative of $q(x)$ evaluated at c and divided by $k!$, but it doesn't really matter. The series is finite because the derivatives of a polynomial eventually vanish. If $|x - c| < 1$ then it follows from this that $|x - c|^k < 1$ for all $k \geqslant 1$ and hence

$$\begin{aligned} |q(x)| &= |a_n(x - c)^n + \cdots + a_1(x - c) + a_0| \\ &\leqslant |a_n||x - c|^n + \cdots + |a_1||x - c| + |a_0| \\ &< |a_n| + \cdots + |a_1| + |a_0|. \end{aligned}$$

Finally, let $\varepsilon > 0$ and $\delta = \min\{1, \varepsilon/(|a_n| + \cdots + |a_0|)\}$. Then for all $x \in \mathbb{C}$ satisfying $|x - c| < \delta$ we have

$$|f(x) - f(c)| = |x - c||q(x)| < \frac{\varepsilon}{|a_n| + \cdots + |a_0|} (|a_n| + \cdots + |a_0|) = \varepsilon,$$

as desired. □

Finally, the main result.

Real Polynomials of Odd Degree

If $f(x) \in \mathbb{R}[x]$ has odd degree then there exists at least one $c \in \mathbb{R}$ such that $f(c) = 0$.

Proof. We may suppose without loss of generality that the leading coefficient is positive. Then we can write $f(x) = a_n x^n + \cdots + a_1 x + a_0$ for some $a_0, \ldots, a_n \in \mathbb{R}$ with $a_n > 0$. If $x \geqslant 1$ then for all $1 \leqslant k \leqslant n - 1$ we have $1 \leqslant x^k \leqslant x^{n-1}$, so that

$$-|a_k|x^n \leqslant -|a_k|x^k \leqslant a_k x^k,$$

[56] I am being sloppy here by using the symbol x both for an abstract variable and for a complex number. It doesn't matter because $\mathbb{C}$ is an infinite field, so polynomial expressions and polynomial functions are the same thing.

and hence

$$\begin{aligned} f(x) &= a_n x^n + \sum_{k=0}^{n-1} a_k x^k \\ &\geqslant a_n x^n - \sum_{k=0}^{n-1} |a_k| x^{n-1} \\ &= a_n x^{n-1} \left(x - \sum_{k=0}^{n-1} \frac{|a_k|}{a_n} \right). \end{aligned}$$

If, in addition, we have $x > \sum_{k=0}^{n-1} |a_k|/a_n$ then this implies that $f(x) > 0$. Thus for any $b \in \mathbb{R}$ greater than the maximum of 1 and $\sum_{k=0}^{n-1} |a_k|/a_n$ we have $f(b) > 0$.

So far we have not used the fact that $f(x)$ has odd degree. Consider the polynomial $g(x) := -f(-x)$. If $f(x)$ has odd degree then $g(x)$ has positive leading coefficient, so by the same argument we can find some $b' > 0$ such that

$$-f(-b') = g(b') > 0.$$

But then $a := -b' < 0$ satisfies $f(a) < 0$.

Finally, since polynomials are continuous, it follows from the IVT that there exists at least one real number $a < c < b$ satisfying $f(c) = 0$. □

As a little bonus, the same proof idea gives the following result.

Lagrange's Root Bound

Consider a complex polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{C}[x]$ with $a_n \neq 0$. Then every complex root $f(c) = 0$ satisfies

$$|c| \leqslant \max \left\{ 1, \sum_{k=0}^{n-1} \frac{|a_k|}{|a_n|} \right\}.$$

6.3 Descartes and Euler on Quartic Equations

Welcome back algebra students.

6.4 Multivariable Polynomials

We are interested in the roots of polynomials in a single variable. But the analysis of these roots forces us to consider polynomials in many variables. It is clear how these should be defined, but the notation is more difficult.

Multivariable Polynomials

Let $\mathbb{F}$ be a field and let $x_1, \ldots, x_n$ be some abstract symbols, which we regard as variables. We define the ring of polynomials in $x_1, \ldots, x_n$ by induction:[57]

$$\mathbb{F}[x_1, \ldots, x_n] := (\mathbb{F}[x_1, \ldots, x_{n-1}])\,[x_n]$$

For example, a general element of the ring $\mathbb{F}[x, y] = (\mathbb{F}[x])\,[y]$ has the form $f(x, y) = \sum_{\ell \geqslant 0} g_\ell(x) y^\ell$ for some polynomials $g_1(x), g_2(x), \ldots \in \mathbb{F}[x]$. If we write $g_\ell(x) = \sum_{k \geqslant 0} a_{k\ell} x^k$ for some coefficients $a_{k\ell} \in \mathbb{F}$ then this becomes

$$f(x, y) = \sum_{\ell \geqslant 0} \left(\sum_{k \geqslant 0} a_{k\ell} x^k \right) y^\ell = \sum_{k, \ell \geqslant 0} a_{k\ell} x^k y^\ell,$$

where the sum is taken over all pairs of natural numbers $(k, \ell) \in \mathbb{N}^2$, and we observe that only finitely many of the coefficients $a_{k\ell}$ are nonzero. Similarly, an element of the ring $\mathbb{F}[x_1, \ldots, x_n]$ can be expressed as

$$f(x_1, \ldots, x_n) = \sum_{k_1, k_2, \ldots, k_n \geqslant 0} a_{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n},$$

where the sum is taken over all n-tuples of natural numbers $(k_1, \ldots, k_n) \in \mathbb{N}^n$ and only finitely many of the coefficients $a_{k_1, \ldots, k_n} \in \mathbb{F}$ are nonzero.

Clearly this notation is unworkable, so we make the following abbreviations:

$$\begin{aligned} \mathbf{x} &= (x_1, \ldots, x_n), \\ \mathbf{k} &= (k_1, \ldots, k_n), \\ \mathbf{x}^{\mathbf{k}} &= x_1^{k_1} \cdots x_n^{k_n}. \end{aligned}$$

By convention we write $\mathbf{x}^{\mathbf{0}} = 1 \in \mathbb{F}$, where $\mathbf{0} = (0, \ldots, 0) \in \mathbb{N}^n$. Then a general element of $\mathbb{F}[\mathbf{x}] = \mathbb{F}[x_1, \ldots, x_n]$ looks like

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}^n} a_{\mathbf{k}} \mathbf{x}^{\mathbf{k}},$$

for some coefficients $a_{\mathbf{k}} \in \mathbb{F}$, only finitely many of which are nonzero. This notation allows us to treat multivariable polynomials very much like polynomials in one variable. For example, for any $\mathbf{k}, \boldsymbol{\ell} \in \mathbb{N}^n$ we observe that

$$\begin{aligned} \mathbf{x}^{\mathbf{k}}\mathbf{x}^{\boldsymbol{\ell}} &= (x_1^{k_1} \cdots x_n^{k_n})(x_1^{\ell_1} \cdots x_n^{\ell_n}) \\ &= x_1^{k_1+\ell_1} \cdots x_n^{k_n+\ell_n} \\ &= \mathbf{x}^{\mathbf{k}+\boldsymbol{\ell}}, \end{aligned}$$

where $\mathbf{k} + \boldsymbol{\ell} = (k_1 + \ell_1, \ldots, k_n + \ell_n) \in \mathbb{N}^n$ is the vector sum. Thus the ring operations can be expressed as follows:

$$\begin{aligned} \left(\sum_{\mathbf{k}} a_{\mathbf{k}}\mathbf{x}^{\mathbf{k}}\right) + \left(\sum_{\boldsymbol{\ell}} b_{\boldsymbol{\ell}}\mathbf{x}^{\mathbf{k}}\right) &= \sum_{\mathbf{k}} (a_{\mathbf{k}} + b_{\mathbf{k}})\mathbf{x}^{\mathbf{k}}, \\ \left(\sum_{\mathbf{k}} a_{\mathbf{k}}\mathbf{x}^{\mathbf{k}}\right) \left(\sum_{\boldsymbol{\ell}} b_{\boldsymbol{\ell}}\mathbf{x}^{\boldsymbol{\ell}}\right) &= \sum_{\mathbf{m}} \left(\sum_{\mathbf{k}+\boldsymbol{\ell}=\mathbf{m}} a_{\mathbf{k}} b_{\boldsymbol{\ell}}\right) \mathbf{x}^{\mathbf{m}}. \end{aligned}$$

The only difference from single variable polynomials is that the sums are taken over all elements of $\mathbb{N}^n$ instead of $\mathbb{N}$.

What about the "degree" of a multivariable polynomials? There are many different ways to do this. For our purpose, we need some way to facilitate proofs by induction. So we make the following definition.

Lexicographic Order and Degree

Given $\mathbf{k}, \boldsymbol{\ell} \in \mathbb{N}^n$ we define $\mathbf{k} = \boldsymbol{\ell}$ when $k_i = \ell_i$ for all i. If $\mathbf{k} \neq \boldsymbol{\ell}$ then there exists some minimum $i \geqslant 1$ such that $k_i \neq \ell_i$. We will write $\mathbf{k} <_{\text{lex}} \boldsymbol{\ell}$ when $k_i < \ell_i$. In other words:

$$\mathbf{k} <_{\text{lex}} \boldsymbol{\ell} \quad \Leftrightarrow \quad \mathbf{k} \text{ is smaller than } \boldsymbol{\ell} \text{ in the first place where they differ.}$$

Under this definition, any finite (nonempty) subset of $\mathbb{N}^n$ has a lexicographically maximum element, which we can use to define the degree of a nonzero polynomial:

$$\begin{aligned} f(\mathbf{x}) &= \sum_{\mathbf{k} \in \mathbb{N}^n} a_{\mathbf{k}}\mathbf{x}^{\mathbf{k}}, \\ \deg(f) &:= \max_{\text{lex}}\{\mathbf{k} \in \mathbb{N}^n : a_{\mathbf{k}} \neq 0\}. \end{aligned}$$

[57] Here we are using the fact that $R[x]$ is defined even when the ring R is not a field.

Note that the degree is a vector in $\mathbb{N}^n$. Sometimes this is called a "multi-degree". The degree of the zero polynomial is not defined.

For example, consider the polynomial $f(x,y,z) = 3x^2y - 5x^2y^2z + 6x^2yz$. If we think of $(x,y,z) = (x_1,x_2,x_3)$ then we have

$$\begin{aligned} \deg(x^2y) &= (2,1,0), \\ \deg(x^2y^2z) &= (2,2,1), \\ \deg(x^2yz) &= (2,1,1). \end{aligned}$$

Then since $(2,2,1) >_{\text{lex}} (2,1,1) >_{\text{lex}} (2,1,0)$ we have

$$\deg(f) = \deg(3x^2y - 5x^2y^2z + 6x^2yz) = (2,2,1).$$

We say that $-5x^2y^2z$ is the *leading term* of the polynomial $f(x,y,z)$, and sometimes we write

$$f(x,y,z) = -5x^2y^2z + \text{ lower terms.}$$

The lexicographic degree also has the following nice property.

Degree of a Product

Given two polynomials $f(\mathbf{x}), g(\mathbf{x}) \in \mathbb{F}[x_1,\ldots,x_n]$ with lexicographic degrees

$$\begin{aligned} \deg(f) &= \mathbf{k} = (k_1,\ldots,k_n), \\ \deg(g) &= \boldsymbol{\ell} = (\ell_1,\ldots,\ell_n), \end{aligned}$$

we must have

$$\deg(fg) = \deg(f) + \deg(g) = \mathbf{k} + \boldsymbol{\ell} = (k_1+\ell_1,\ldots,k_n+\ell_n).$$

You will prove this on the homework.

For example, consider the polynomials

$$\begin{aligned} f(x,y,z) &= 3x^2y - 5x^2y^2z + 6x^2yz, \\ g(x,y,z) &= xy^2 - 2xy^2z, \end{aligned}$$

with $\deg(f) = (2,2,1)$ and $\deg(g) = (1,2,1)$. The product is

$$f(x,y,z)g(x,y,z) = (3x^2y - 5x^2y^2z + 6x^2yz)(xy^2 - 2xy^2z)$$

$$\begin{aligned}
&= (3x^2y - 5x^2y^2z + 6x^2yz)(xy^2) + (3x^2y - 5x^2y^2z + 6x^2yz)(-2xy^2z) \\
&= (3x^3y^3 - 5x^3y^4z + \cancel{6x^3y^3z}) + (\cancel{-6x^3y^3z} + 10x^3y^4z^2 - 12x^3y^2z^2) \\
&= 3x^3y^3 - 5x^3y^4z + 10x^3y^4z^2 - 12x^3y^2z^2,
\end{aligned}$$

which has degree $(3,4,2)$ because

$$(3,4,2) >_{\text{lex}} (3,4,1) >_{\text{lex}} (3,3,0) >_{\text{lex}} (3,2,2).$$

Note that $\deg(f) + \deg(g) = (2,2,1) + (1,2,1) = (3,4,1) = \deg(fg)$. We can also write

$$\begin{aligned}
f(x,y,z) &= -5x^2y^2z + \text{lower terms}, \\
f(x,y,z) &= -2xy^2z + \text{lower terms}, \\
f(x,y,z)g(x,y,z) &= 10x^3y^4z^2 + \text{lower terms}.
\end{aligned}$$

Note that the leading term of the product is the product of the leading terms.

Ultimately, we will use the degree for writing proofs by induction. For this purpose we need to know that the lexicographic order on $\mathbb{N}^n$ has the so-called "well-ordering property", also called the "descending chain condition".

Lexicographic Order is a Well-Ordering

Consider the lexicographic ordering on $\mathbb{N}^2$:[58]

$$\begin{aligned}
&(0,0) < (0,1) < (0,2) < \cdots \\
&< (1,0) < (1,1) < (1,2) < \cdots \\
&< (2,0) < (2,1) < (2,2) < \cdots \\
&< \cdots
\end{aligned}$$

This ordering is a bit strange because it contains infinite ascending sequences that are bounded above. For example, the infinite ascending sequence $(0,0) < (0,1) < (0,2) < \cdots$ is bounded above by $(1,0)$.

However, the lexicographic order does not have any infinite **descending** sequences. To be precise, there does not exist an infinite sequence $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \ldots \in \mathbb{N}^n$ satisfying

$$\mathbf{k}_1 > \mathbf{k}_2 > \mathbf{k}_3 > \cdots .$$

Proof. The lexicographic order on $\mathbb{N} = \mathbb{N}^1$ coincides with the usual order. The fact that $\mathbb{N}$ is well-ordered is part of the **definition** of natural numbers. We will prove that $\mathbb{N}^n$ is

[58] When no confusion can I arise I will write $<$ instead of $<_{\text{lex}}$.

well-ordered by induction on n. So assume for induction $\mathbb{N}^{n-1}$ is well-ordered and suppose for contradiction that we have an infinite descending sequence in $\mathbb{N}^n$:

$$\mathbf{k}_1 > \mathbf{k}_2 > \mathbf{k}_3 > \cdots . \qquad (*)$$

Let's write $\mathbf{k}_i = (k_{i1}, k_{i2}, \ldots, k_{in})$. Then by definition of lexicographic order, we must have

$$k_{11} \geqslant k_{21} \geqslant k_{31} \geqslant \cdots .$$

Since $\mathbb{N}$ itself is well-ordered there exists some M such that $k_{m1} = k_{M1}$ for all $m \geqslant M$. The idea is to delete this common first element from each of the vectors to obtain an infinite descending sequence in $\mathbb{N}^{n-1}$. To be precise, let us write

$$\mathbf{k}'_i = (k_{i2}, k_{i3}, \ldots, k_{in}).$$

Then from $(*)$ and the definition of lexicographic order we must have

$$\mathbf{k}'_M > \mathbf{k}'_{M+1} > \mathbf{k}'_{M+2} > \cdots ,$$

which contradicts our assumption that $\mathbb{N}^{n-1}$ is well-ordered. □

Before moving on, we make one final observation about multivariable polynomials, generalizing the observations of Section 2.3.

Polynomial Expressions vs Polynomial Functions

Any polynomial expression $f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ determines a polynomial function $f : \mathbb{F}^n \to \mathbb{F}$ by *evaluation*:

$$\begin{array}{cccc} f : & \mathbb{F}^n & \to & \mathbb{F} \\ & (\alpha_1, \ldots, \alpha_n) & \mapsto & f(\alpha_1, \ldots, \alpha_n). \end{array}$$

If the field $\mathbb{F}$ is infinite then I claim that different polynomial expressions determine different functions. Equivalently, if two polynomial expressions $f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{N}^d} a_{\mathbf{k}}\mathbf{x}^{\mathbf{k}}$ and $g(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{N}^n} b_{\mathbf{k}}\mathbf{x}^{\mathbf{k}}$ determine the same function then they have the same coefficients:[59]

$$f(\boldsymbol{\alpha}) = g(\boldsymbol{\alpha}) \text{ for all } \boldsymbol{\alpha} \in \mathbb{F}^n \quad \Rightarrow \quad a_{\mathbf{k}} = b_{\mathbf{k}} \text{ for all } \mathbf{k} \in \mathbb{N}^d.$$

This fact allows us to be a bit sloppy in our reasoning with multivariable polynomials, at least over infinite fields such as $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

Proof. We prove this by induction on the number of variables.

[59] If the field $\mathbb{F}$ is finite then this result is false.

6.5 Permutations

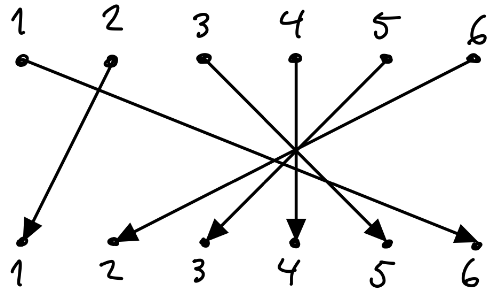
In the previous section we developed the language of multivariable polynomials. Our ultimate goal in this chapter is to prove the FTA. But before this we need another kind of notation, for "permutations". These could have been defined at any point in the course, but now seems appropriate.

Definition of Permutations

A *permutation* is an invertible function from a finite set to itself. Since all sets of the same size are basically equivalent we usually consider the set $\{1, 2, \ldots, n\}$. We denote the set of such permutations by

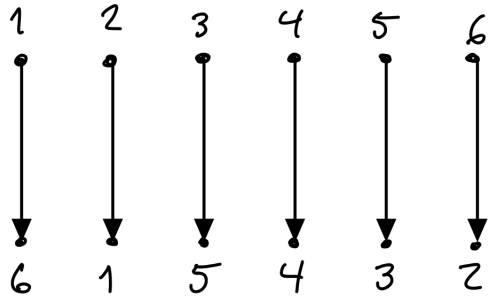
$$S_n = \text{the set of invertible functions } \sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}.$$

For example, consider the permutation $\sigma \in S_6$ defined by the following diagram:



It is inconvenient to draw such diagrams, to we will define two more concise notations.

One-Line Notation. Here we rearrange the arrows so that the number $\sigma(i)$ appears directly under the number i:

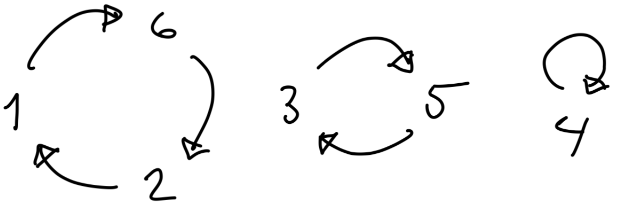


Then we encode σ by listing the numbers in the second row:

$$\sigma = 615432.$$

This notation makes it clear that $\#S_n = n!$, since there are n ways to choose the leftmost number, then $n-1$ ways to choose the next number, etc.

Cycle Notation. Here we only draw the numbers once:



Note that the numbers break up into "oriented cycles". To express these cycles concisely we write them inside parentheses:

$$\sigma = (162)(35)(4).$$

Unfortunately this notation is not unique. For example, we can record a cycle by starting from any number:

$$(162) = (621) = (216) \quad \text{and} \quad (35) = (53).$$

Also, the ordering of the cycles is irrelevant:

$$\sigma = (162)(35)(4) = (4)(216)(35) = (53)(4)(621) = \text{etc.}$$

Another quirk of the notation is that we typically omit the "singleton cycles". In our example this means omitting the (4):

$$\sigma = (162)(35).$$

Nevertheless, this is our preferred notation for permutations since is the most compact and meaningful. A particularly nice property is that the inverse of a permutation is obtained by reversing the orientation of the cycles:

$$\sigma^{-1} = (126)(35).$$

For example, here are all of the $3! = 6$ elements of the set S_3, expressed in one-line notation and in cycle notation:

one-line	123	213	132	321	231	312
cycle	id	(12)	(23)	(13)	(123)	(132)

The important thing about permutations is that they form a "group". Recall from Chapter 4 that a group $(G, *, \varepsilon)$ consists of a set G with a binary operation $* : G \times G \to G$ and a special element ε, satisfying the following axioms:

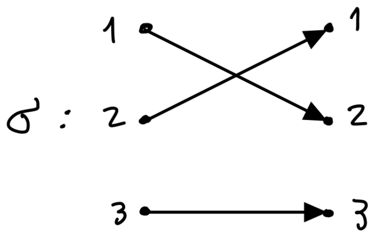
(G1) $\forall a, b, c \in G, a * (b * c) = (a * b) * c$,

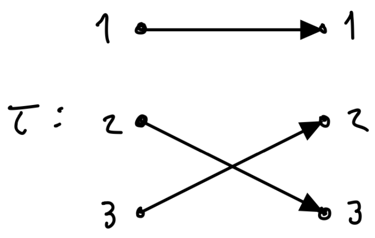
(G2) $\forall a \in G, a * \varepsilon = \varepsilon * a = a$,

(G3) $\forall a \in G, \exists b \in G, a * b = b * a = \varepsilon$.

If $a * b = b * a$ for all $a, b \in G$ then we say that the group is "abelian". Abelian groups can be used to model the properties of addition and multiplication of numbers. Non-abelian groups are used to model the composition of invertible functions. We will show that the structure $(S_n, \circ, \mathrm{id})$ is a group, where $\circ$ is functional composition and $\mathrm{id} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ is the identity function defined by $\mathrm{id}(i) = i$ for all $i \in \{1, 2, \ldots, n\}$.

But before doing this, let me emphasize the composition of permutations is not commutative. For example, consider the permutations $\sigma, \tau \in S_3$ defined by the following diagrams:

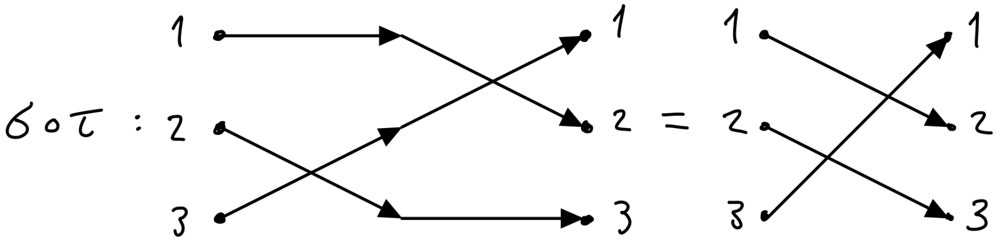




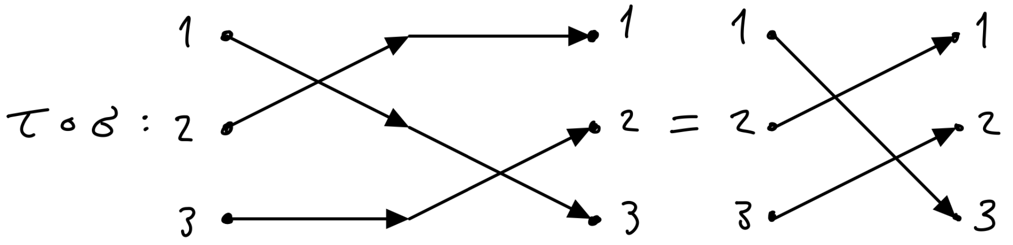
Recall that functional composition is defined as follows:

$$(\sigma \circ \tau)(i) = \sigma(\tau(i)) \quad \text{for all } \sigma, \tau \in S_n \text{ and } i \in \{1, \ldots, n\}. \tag{$*$}$$

Thus we may compose the permutations by juxtaposing the diagrams:



and



Sadly, we read the diagrams from left to right but we read the notation "$\sigma \circ \tau$" from right to left, i.e., "do τ first and then do σ". This is an unavoidable consequence of the notation $(*)$, i.e., the fact that we write the name of a function to the left of its argument.

From the diagrams it is clear that $\sigma \circ \tau$ and $\tau \circ \sigma$ are different functions. Expressing these in cycle notation gives

$$(12) \circ (23) = (123) \quad \text{and} \quad (23) \circ (12) = (132).$$

Now we verify the group properties of permutations:

- Property (G1) is an automatic property of functional composition. Given $\rho, \sigma, \tau \in S_n$ and $i \in \{1, 2, \ldots, n\}$ we have by definition that

  $$[\rho \circ (\sigma \circ \tau)](i) = \rho((\sigma \circ \tau)(i)) = \rho(\sigma(\tau(i)))$$

  and

  $$[(\rho \circ \sigma) \circ \tau](i) = (\rho \circ \sigma)(\tau(i)) = \rho(\sigma(\tau(i))).$$

  Since the functions $\rho \circ (\sigma \circ \tau)$ and $(\rho \circ \sigma) \circ \tau$ do the same thing, they are equal.

- And property (G2) is almost the definition of the identity function. For all $\sigma \in S_n$ and $i \in \{1, \ldots, n\}$ we have

  $$(\sigma \circ \mathrm{id})(i) = \sigma(\mathrm{id}(i)) = \sigma(i)$$

  and

  $$(\mathrm{id} \circ \sigma)(i) = \mathrm{id}(\sigma(i)) = \sigma(i).$$

  Since the functions $\sigma \circ \mathrm{id}$, $\mathrm{id} \circ \sigma$ and σ all do the same thing, they are the same function.

- For property (G3) we will show that for all $\sigma \in S_n$ we have $\sigma^{-1} \in S_n$. If $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ is invertible, then its inverse $\sigma^{-1} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ satisfies $\sigma \circ \sigma^{-1} = \mathrm{id}$ and $\sigma^{-1} \circ \sigma = \mathrm{id}$. These same identities show that σ^{-1} is invertible with $(\sigma^{-1})^{-1} = \sigma$. Hence for any $\sigma \in S_n$ we also have $\sigma^{-1} \in S_n$.

But we forgot something. Given two permutations $\sigma, \tau \in S_n$, is it not quite obvious that the composite function $\sigma \circ \tau : \{1, \ldots, n\} \to \{1, \ldots, n\}$ is a permutation. To see this we must show that $\sigma \circ \tau$ is invertible. In fact, I claim that $(\sigma \circ \tau)^{-1} = \tau^{-1} \circ \sigma^{-1}$.[60] Indeed, by applying properties (G1) and (G2) we obtain

$$(\sigma \circ \tau) \circ (\tau^{-1} \circ \sigma^{-1}) = \sigma \circ (\tau \circ \tau^{-1}) \circ \sigma^{-1} = \sigma \circ \mathrm{id} \circ \sigma^{-1} = \sigma \circ \sigma^{-1} = \mathrm{id}$$

and

$$(\tau^{-1} \circ \sigma^{-1}) \circ (\sigma \circ \tau) = \tau^{-1} \circ (\sigma^{-1} \circ \sigma) \circ \tau = \tau^{-1} \circ \mathrm{id} \circ \tau = \tau^{-1} \circ \tau = \mathrm{id}.$$

Since $\sigma \circ \tau$ is an invertible function from $\{1, \ldots, n\}$ to itself, it is an element of S_n.

Here is a summary.

The Symmetric Group

Let S_n be the set of permutations of the set $\{1, \ldots, n\}$. Then the structure $(S_n, \circ, \mathrm{id})$ is a group, called the *symmetric group on n symbols*.

For example, here is the group table of the symmetric group S_3, where $\sigma \circ \tau$ is the entry in the row corresponding to σ and the column corresponding to τ:

$\circ$	id	(12)	(13)	(23)	(123)	(132)
id	id	(12)	(13)	(23)	(123)	(132)
(12)	(12)	id	(132)	(123)	(23)	(13)
(13)	(13)	(123)	id	(132)	(12)	(23)
(23)	(23)	(132)	(123)	id	(13)	(12)
(123)	(123)	(13)	(23)	(12)	(132)	id
(132)	(132)	(23)	(12)	(13)	id	(123)

This group is not abelian since, for example, we have $(12) \circ (23) = (132)$ and $(23) \circ (12) = (123)$, but $(123) \neq (132)$.

So far we have mostly discussed basic definitions. Before moving on, let's prove a theorem.

Transpositions and the Alternating Group

Permutations of the form $(ij) \in S_n$ with $i \neq j$ are called *transpositions*, or 2-*cycles*. Recall that the function $(ij) : \{1, \ldots, n\} \to \{1, \ldots, n\}$ switches $i \leftrightarrow j$ and sends every other number to itself. Since $(ij) = (ji)$, the number of transpositions in S_n is just the number of pairs of indices $1 \leqslant i < j \leqslant n$, which is $\binom{n}{2} = n(n-1)/2$.

I claim that every element of S_n can be expressed as a composition of transpositions.

[60] You may remember this formula from multiplication of matrices, which is also defined in terms of functional composition.

Assuming this, we let $A_n \subseteq S_n$ denote the set of permutations that can be expressed as a composition of an **even number** of transpositions. This set satisfies the following properties:

- $\text{id} \in A_n$,
- $\sigma, \tau \in A_n \Rightarrow \sigma \circ \tau \in A_n$,
- $\sigma \in A_n \Rightarrow \sigma^{-1} \in A_n$.

In other words, A_n is a **subgroup** of S_n. We call it the *alternating group on n symbols.*

Proof. First we show that every permutation can be expressed as a composition of transpositions. The cycle notation has the property that it can be viewed as a composition of commuting cycles. For example, we have

$$(137)(256)(48) = (137) \circ (256) \circ (48) = (48) \circ (137) \circ (256) = (562) \circ (84) \circ (712) = \text{etc.}$$

Because of this feature, it is common to omit the composition symbol $\circ$ when working with permutations in cycle notation. Next we show that each cycle can be viewed as a composition of (non-commuting) transpositions. For example, we have seen that $(123) = (12) \circ (23)$, we will write as $(123) = (12)(23)$. One can similarly check that

$$\begin{aligned}(1234) &= (12)(23)(34),\\ (12335) &= (12)(23)(34)(45),\end{aligned}$$

and, indeed, for any numbers $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$ we have

$$(i_1 i_2 i_3 \cdots i_{k-1} i_k) = (i_1 i_2)(i_2 i_3) \cdots (i_{k-1} i_k).$$

By combining these two observations, we see that any permutation can be expressed as a composition of (generally non-commuting) cycles.[61] For example,

$$(137)(256)(48) = (13)(37)(25)(56)(48).$$

Next we verify the subgroup axioms:

- By definition we say that id is a composition of no transpositions, which implies that $\text{id} \in A_n$ because zero is even. If you don't like that, observe that for any transposition (ij) we have $\text{id} = (ij)(ij)$. Since 2 is even this implies that $\text{id} \in A_n$.

[61] This expression is not unique. For example, we could also write

$$(i_1 i_2 i_3 \cdots i_{k-1} i_k) = (i_1 i_k)(i_1 i_{k-1}) \cdots (i_1 i_2).$$

- Suppose that $\sigma, \tau \in A_n$. By definition this means we can write

  $$\begin{aligned} \sigma &= s_1 \circ s_2 \circ \cdots \circ s_k, \\ \tau &= t_1 \circ t_2 \circ \cdots \circ t_\ell, \end{aligned}$$

  for some transpositions $s_1, \ldots, s_k, t_1, \ldots, t_k$, where k and ℓ are even. But then $\sigma \circ \tau$ is a composition of $k + \ell$ transpositions:

  $$\sigma \circ \tau = s_1 \circ s_2 \circ \cdots \circ s_k \circ t_1 \circ t_2 \circ \cdots \circ t_\ell.$$

  Since $k + \ell$ is even this implies that $\sigma \circ \tau \in A_n$.

- Let $\sigma \in A_n$ so that $\sigma = s_1 \circ \cdots \circ s_k$ for some transpositions $s_1, \ldots, s_k$, where k is even. Now observe that for any transposition we have $s^{-1} = s$. Furthermore, for any permutations ρ, τ we have $(\rho \circ \tau)^{-1} = \tau^{-1} \circ \rho^{-1}$. It follows that

  $$\sigma^{-1} = (s_1 \circ \cdots \circ s_k)^{-1} = s_k^{-1} \circ \cdots \circ s_1^{-1} = s_k \circ \cdots \circ s_1,$$

  which is a composition of an even number of transpositions. Hence $\sigma^{-1} \in A_n$.

□

It is much harder to prove that some permutation $\sigma \in S_n$ is **not** in the subgroup A_n. We will have a trick for doing this after we discuss the discriminant of a polynomial. We will also prove later that exactly half of the permutations are alternating:

$$\#A_n = \frac{1}{2}\#S_n = n!/2.$$

For example, here are the $6!/2 = 3$ elements of A_3:

$$A_3 = \{\text{id}, (123), (132)\}.$$

And here is the group table:

$\circ$	id	(123)	(132)
id	id	(123)	(132)
(123)	(123)	(132)	id
(132)	(132)	id	(123)

By accident, it happens that this group **is abelian**, and in fact it is isomorphic to the additive group $(\mathbb{Z}/3\mathbb{Z}, +, 0)$. This can be seen by observing that the group tables are "the same" up to renaming of the elements:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

We will show later that **any** two groups of size 3 must be isomorphic.

6.6 Symmetric Polynomials

The concept of symmetric polynomials is intuitive but the notation is difficult. I could present the entire discussion at an intuitive level with examples, but I choose also to develop a rigorous notation. On a first reading you should definitely focus on the examples.

We say that a multivariable polynomial $f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ if the value of corresponding function $f : \mathbb{F}^n \to \mathbb{F}$ is left unchanged by any permutation of its inputs. Sometimes we also call this a "symmetric function". For example, consider the following polynomial in two variables:

$$f(x, y) = x^3 + 5x^2y + 2x + 2y + 5xy^2 + y^3.$$

This polynomial is symmetric because $f(x, y) = f(y, x)$. However the following polynomial is **not** symmetric:

$$g(x, y) = x^3 + 5x^2y + 2x + 2y + 4xy^2 + y^3.$$

To be precise, we have $g(x, y) - g(y, x) = x^2y - xy^2$, which is not zero. The problem here is that the coefficients of the monomials x^2y and xy^2 in g are not the same.

The most basic symmetric polynomials in two variables are just the sum and product. We call these the "elementary symmetric polynomials in x and y":

$$\begin{aligned} e_1(x, y) &= x + y, \\ e_2(x, y) &= xy. \end{aligned}$$

We will prove below that **any** symmetric polynomial can be expressed in terms of *elementary* symmetric polynomials, and the method of proof will be a kind of "division algorithm". Let me show you how this works in the case of the polynomial $f(x, y)$ above. The lexicographic degree of $f(x, y)$ is $(3, 0)$ and the leading term is x^3. Let's write

$$f(x, y) = x^3 + \text{lower terms}.$$

It is easy to find a combination of e_1 and e_2 with the same leading term:

$$\begin{aligned} e_1^3 &= (x + y)^3 \\ &= x^3 + 3x^2y + 3xy^2 + y^3 \\ &= x^3 + \text{lower terms}. \end{aligned}$$

Therefore the difference $f - e_1^3$ has smaller degree:

$$f - e_1^3 = 2x^2y + 2x + 2y + 2xy^2 = 2x^2y + \text{lower terms}.$$

Now we play the same trick again. With a bit of trial-and-error we can find a combination of e_1 and e_2 with the same leading term:

$$\begin{aligned} 2e_1e_2 &= 2(x + y)(xy) \\ &= 2x^2y + 2xy^2 \end{aligned}$$

$$= 2x^2y + \text{lower terms}.$$

Finally, subtracting this from the previous polynomial gives

$$f - e_1^3 - 2e_1^2e_2 = 2x + 2y = 2(x+y) = 2e_1,$$

and we conclude that

$$\begin{aligned} f(x,y) &= e_1(x,y)^3 + 2e_1(x,y)^2e_2(x,y) + 2e_1(x,y), \\ f &= e_1^3 + 2e_1^2e_2 + 2e_1. \end{aligned}$$

The main goal of this chapter is to generalize this algorithm to any number of variables.

But why do we care about symmetric polynomials? Suppose that a quadratic polynomial $x^2 + ax + b$ has coefficients a, b in some field $\mathbb{F}$ and has roots α, β in a larger field $\mathbb{E} \supseteq \mathbb{F}$. By factoring in the ring $\mathbb{E}[x]$ we obtain

$$\begin{aligned} x^2 + ax + b &= (x-\alpha)(x-\beta) \\ &= x^2 - (\alpha+\beta)x + (\alpha\beta). \end{aligned}$$

Then comparing coefficients gives

$$\begin{aligned} a &= -(\alpha+\beta) = -e_1(\alpha,\beta) \\ b &= \alpha\beta = e_2(\alpha,\beta). \end{aligned}$$

Since the coefficients a and b are (up to sign) just the elementary symmetric combinations of the roots α and β, it follows from above algorithm that **any symmetric combination of the roots α, β can be expressed in terms of the coefficents a, b.** For example, we have

$$\begin{aligned} f(\alpha,\beta) &= \alpha^3 + 5\alpha^2\beta + 2\alpha + 2\beta + 5\alpha\beta^2 + \beta^3 \\ &= e_1(\alpha,\beta)^3 + 2e_1(\alpha,\beta)^2e_2(\alpha,\beta) + 2e_1(\alpha,\beta) \\ &= (-a)^3 + 2(-a)^2b + 2(-a) \\ &= -a^3 + 2a^2b - 2a. \end{aligned}$$

We conclude that the number $f(\alpha,\beta)$, which by definition lives in the extension field $\mathbb{E}$, is actually in the base field $\mathbb{F}$.

The general version of this theorem is particularly interesting when $\mathbb{F} = \mathbb{R}$ since it implies that **any symmetric combination of the roots of a real polynomial is real, no matter where the roots live**. This will be a key step in our proof of the FTA.

Definition of Symmetric Polynomials

A permutation $\sigma \in S_n$ "acts on" a polynomial $f(x_1,\ldots,x_n) \in \mathbb{F}[x_1,\ldots,x_n]$ by permuting its inputs. We introduce a special notation for this:

$$(\sigma \cdot f)(x_1,\ldots,x_n) := f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}).$$

That is, for each permutation $\sigma \in S_n$ and each polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ we have another polynomial $\sigma \cdot f \in \mathbb{F}[x_1, \ldots, x_n]$ whose inputs have been permuted according to σ. This is a reasonable notation since for any permutations $\sigma, \tau \in S_n$ and any polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ one can check that

$$(\sigma \circ \tau) \cdot f = \sigma \cdot (\tau \cdot f).$$

One can also check that that the "action" of a permutation σ preserves addition and multiplication of polynomials:

$$\sigma \cdot (f + g) = \sigma \cdot f + \sigma \cdot g \quad \text{and} \quad \sigma \cdot (fg) = (\sigma \cdot f)(\sigma \cdot g).$$

We say that a polynomial $f(x_1, \ldots, x_n)$ is *symmetric* when it is invariant under any permutation of its inputs:[62]

$$f \text{ is symmetric} \quad \Longleftrightarrow \quad \sigma \cdot f = f \text{ for all } \sigma \in S_n.$$

It follows from the properties above that the set of symmetric polynomials is closed under addition and multiplication, hence it is a subring of $\mathbb{F}[x_1, \ldots, x_n]$. We call this the *ring of symmetric polynomials* and we denote it by

$$\mathbb{F}[x_1, \ldots, x_n]^{S_n} \subseteq \mathbb{F}[x_1, \ldots, x_n].$$

Below we will prove the Fundamental Theorem of Symmetric Polynomials (FTSP), which gives a sort of "basis" for the ring of symmetric polynomials. To be precise, there are some "elementary symmetric polynomials" $e_1, \ldots, e_n$ such that every symmetric polynomial f can be expressed in as a polynomial expression in $e_1, \ldots, e_n$:

$$\begin{gathered}\text{any symmetric polynomial} = \text{some polynomial expression in } e_1, \ldots, e_n \\ f = g(e_1, \ldots, e_n).\end{gathered}$$

Elementary Symmetric Polynomials

Consider the ring of polynomials in $n + 1$ variables, which we call $x_1, \ldots, x_{n+1}, y$:

$$\mathbb{F}[x_1, \ldots, x_n, y] = (\mathbb{F}[x_1, \ldots, x_n])\,[y].$$

We define the *elementary symmetric polynomials* $e_1, \ldots, e_n \in \mathbb{F}[x_1, \ldots, x_n]$ as the coefficients of the following polynomial:

$$(y - x_1)(y - x_2) \cdots (y - x_n) = y^n - e_1 y^{n-1} + e_2 y^{n-2} - \cdots + (-1)^n e_n.$$

[62] Actually, since every permutation can be expressed as a composition of transpositions, it is sufficient to check that f is invariant under any transposition of its inputs.

We observe that each coefficient $e_k(x_1, \ldots, x_n)$ is, indeed, a symmetric polynomial in $x_1, \ldots, x_n$ since the expression $(y - x_1) \cdots (y - x_n)$ is invariant under permuting $x_1, \ldots, x_n$. To be explicit, we have

$$\begin{aligned} e_1 &= x_1 + x_2 + \cdots + x_n, \\ e_2 &= x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n = \sum_{1 \leqslant i < j \leqslant n} x_i x_j, \\ &\vdots \\ e_k &= \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} x_{i_1} x_{i_2} \cdots x_{i_k}, \\ &\vdots \\ e_n &= x_1 x_2 \cdots x_n \end{aligned}$$

In essence, the elementary symmetric polynomials just express the relationship between the roots and the coefficients of a single-variable polynomial. To see this, suppose that a polynomial $f(x) \in \mathbb{F}[x]$ has roots $\alpha_1, \ldots, \alpha_n$ in some field $\mathbb{E} \supseteq \mathbb{F}$, so that

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \text{ in } \mathbb{E}[x].$$

Then expanding the right hand side gives

$$f(x) = x^n - e_1(\alpha_1, \ldots, \alpha_n)x^{n-1} + e_2(\alpha_1, \ldots, \alpha_n)x^{n-2} - + \cdots + (-1)^n e_n(\alpha_1, \ldots, \alpha_n).$$

Since the coefficients of $f(x)$ were assumed to be in $\mathbb{F}$ we see that the elementary symmetric combinations of the roots $e_k(\alpha_1, \ldots, \alpha_n)$ are in $\mathbb{F}$. More generally, let $g(x_1, \ldots, x_n)$ be any symmetric polynomial with coefficients from $\mathbb{F}$. It will follow from the FTSP below that there exists some (possibly non-symmetric) polynomial $h \in \mathbb{F}[x_1, \ldots, x_n]$ such that

$$g(x_1, \ldots, x_n) = h(e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n)).$$

Then by substituting α_i for x_i we find that the element $g(\alpha_1, \ldots, \alpha_n)$ of $\mathbb{E}$ is actually in $\mathbb{F}$:

$$g(\alpha_1, \ldots, \alpha_n) = h(e_1(\alpha_1, \ldots, \alpha_n), \ldots, e_n(\alpha_1, \ldots, \alpha_n)) \in \mathbb{F}.$$

In the case $\mathbb{F} = \mathbb{R}$ this shows that **any symmetric combination of the roots of a real polynomial is real**, no matter where the roots live.

To prepare for the FTSP we examine the lexicographic degrees of the elementary symmetric polynomials. Another way to to express the kth elementary symmetric polynomial is as follows:

$$e_k(\mathbf{x}) = \sum_{\boldsymbol{\ell} \in V_{n,k}} \mathbf{x}^{\boldsymbol{\ell}},$$

where $V_{n,k} \subseteq \mathbb{N}^n$ is the set of vectors made from k copies of 1 and $n-k$ copies of 0. For example, when $n = 4$ we have

$$\begin{aligned} e_2(x_1,x_2,x_3,x_4) &= x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4, \\ e_2(\mathbf{x}) &= \mathbf{x}^{(1,1,0,0)} + \mathbf{x}^{(1,0,1,0)} + \mathbf{x}^{(1,0,0,1)} + \mathbf{x}^{(0,1,1,0)} + \mathbf{x}^{(0,1,0,1)} + \mathbf{x}^{(0,0,1,1)}. \end{aligned}$$

The lexicographically largest such vector is $(1,1,\ldots,1,0,0,\ldots,0)$, so we conclude that

$$\begin{aligned} \deg(e_k) &= (1,1,\ldots,1,0,0,\ldots,0), \\ e_k &= x_1x_2\cdots x_k + \text{lower terms}. \end{aligned}$$

Fundamental Theorem of Symmetric Polynomials

For any symmetric polynomial $f(x_1,\ldots,x_n) \in \mathbb{F}[x_1,\ldots,x_n]$ there exists a possibly non-symmetric polynomial $g(x_1,\ldots,x_n) \in \mathbb{F}[x_1,\ldots,x_n]$ such that[63]

$$f(\mathbf{x}) = g(e_1(\mathbf{x}), e_2(\mathbf{x}), \ldots, e_n(\mathbf{x})).$$

The proof of this is an explicit algorithm, similar to division with remainder. The algorithm was first written down in Edward Waring's *Meditationes Arithmeticae* (1770). However, it was probably generally known, going back perhaps to Isaac Newton. In Lagrange's *Treatise on the Theory of Equations* (1770) he used the fact that any symmetric combination of the roots of a real polynomial is real, and he said this was well-known. Maybe he just didn't want to deal with the horrible notation of multivariable polynomial expressions.

Proof. Let $f \in \mathbb{F}[x_1,\ldots,x_n]$ be symmetric with lexicographic degree $\mathbf{k} = (k_1,\ldots,k_n)$ and leading coefficient $c \in \mathbb{F}$:

$$f(\mathbf{x}) = c\mathbf{x}^{\mathbf{k}} + \text{lower terms}.$$

The fact that f is symmetric implies that $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_n$. Indeed, suppose for contradiction that we have $k_i < k_{i+1}$ for some i and let

$$\mathbf{k}' = (k_1,\ldots,k_{i-1},k_{i+1},k_i,k_{i+2},\ldots,k_n).$$

We observe that $\mathbf{k}' > \mathbf{k}$ in lexicographic order, hence by definition of degree the coefficient of $\mathbf{x}^{\mathbf{k}'}$ in f must be zero. On the other hand, since f, being symmetric, is invariant under switching x_i and x_{i+1}, the coefficients of $\mathbf{x}^{\mathbf{k}}$ and $\mathbf{x}^{\mathbf{k}'}$ in f must be equal. Contradiction.

[63] It may be confusing that we write g as an element of $\mathbb{F}[x_1,\ldots,x_n]$. This is an annoying property of our notation for polynomials. We don't have a good way to distinguish between formal polynomial expressions and polynomial functions. Here we are viewing f as a formal polynomial and g as function from $\mathbb{F}[x_1,\ldots,x_n]^n$ to $\mathbb{F}[x_1,\ldots,x_n]$. It might be less confusing to write a formal polynomial as $f(-,-,\ldots,-)$ with empty inputs, but nobody does this.

Now we want a combination of elementary symmetric polynomials with the same leading term as f. I claim that the following polynomial does the job:

$$g(\mathbf{x}) = ce_1(\mathbf{x})^{k_1-k_2}e_2(\mathbf{x})^{k_2-k_3}\cdots e_{n-1}(\mathbf{x})^{k_{n-1}-k_n}e_n(\mathbf{x})^{k_n}.$$

Indeed, by the general properties of degree we have

$$\begin{aligned}\deg(g) &= (k_1-k_2)\deg(e_1) + (k_2-k_3)\deg(e_2) + \cdots + (k_{n-1}-k_n)\deg(e_{n-1}) + k_n\deg(e_n)\\ &= (k_1-k_2)(1,0,\ldots,0)\\ &\quad + (k_2-k_3)(1,1,0,\ldots,0)\\ &\quad\;\vdots\\ &\quad + (k_{n-1}-k_n)(1,\ldots,1,0)\\ &\quad + k_n(1,1,\ldots,1)\\ &= (k_1,k_2,\ldots,k_n) = \mathbf{k}.\end{aligned}$$

Since f and g are symmetric polynomials with the same leading term, it follows that $f-g$ is a symmetric polynomial with degree strictly smaller than $\mathbf{k}$. By induction[64] there exists a polynomial $h \in \mathbb{F}[x_1,\ldots,x_n]$ such that

$$f(\mathbf{x}) - g(\mathbf{x}) = h(e_1(\mathbf{x}), e_2(\mathbf{x}),\ldots,e_n(\mathbf{x})).$$

Finally, we conclude that $f = g(\mathbf{x}) + h(e_1(\mathbf{x}),\ldots,e_n(\mathbf{x}))$ is a combination of elementary symmetric polynomials, as desired. □

When it comes to algorithmic proofs it's usually more instructive to see an example.

Example. Consider a field extension $\mathbb{E} \supseteq \mathbb{F}$ and suppose that

$$x^3 + ax^2 + bx + c = (x-\alpha)(x-\beta)(x-\gamma)$$

for some $a,b,c \in \mathbb{F}$ and $\alpha,\beta,\gamma \in \mathbb{E}$. By expanding the right hand side and equating coefficients we see that

$$\begin{aligned}-a &= e_1(\alpha,\beta,\gamma),\\ b &= e_2(\alpha,\beta,\gamma),\\ -c &= e_3(\alpha,\beta,\gamma).\end{aligned}$$

Our goal is to find some polynomial $x^3 + a'x^2 + b'x + c' \in \mathbb{F}[x]$ whose roots are $\alpha^2,\beta^2,\gamma^2 \in \mathbb{E}$:

$$x^3 + a'x^2 + b'x + c' = (x-\alpha^2)(x-\beta^2)(x-\gamma^2).$$

By expanding the right hand side and equating coefficients we obtain

$$-a' = \alpha^2 + \beta^2 + \gamma^2,$$

[64] This works because the lexicographic order on $\mathbb{N}^n$ has no infinite descending sequences.

$$\begin{aligned} b' &= \alpha^2\beta^2 + \alpha^2\gamma^2 + \beta^2\gamma^2, \\ -c' &= \alpha^2\beta^2\gamma^2. \end{aligned}$$

Since each of these is symmetric in α, β, γ, we know that each of the unknown coefficients a', b', c' can be expressed in terms of the elementary symmetric combinations $e_1(\alpha, \beta, \gamma)$, $e_2(\alpha, \beta, \gamma)$, $e_3(\alpha, \beta, \gamma)$, and hence in terms of the original coefficients a, b, c. We will apply the algorithm three times to obtain these expressions.

We begin with a'. Note that a' and $-e_1^2$ have the same leading term $-\alpha^2$. Expand $-e_1^2$ to get

$$-e_1^2 = -(\alpha + \beta + \gamma)^2 = -\alpha^2 - \beta^2 - \gamma^2 - 2(\alpha\beta + \alpha\gamma + \beta\gamma).$$

Then subtract to get

$$\begin{aligned} a' - (-e_1^2) &= 2(\alpha\beta + \alpha\gamma + \beta\gamma) \\ a' - (-e_1^2) &= 2e_2 \\ a' &= -e_1^2 + 2e_2 \\ &= -(-a)^2 + 2(b) \\ &= 2b - a^2. \end{aligned}$$

Next we compute b'. Observe that b' and e_2^2 have the same leading term $\alpha^2\beta^2$. Expand to get

$$\begin{aligned} e_2^2 &= (\alpha\beta + \alpha\gamma + \beta\gamma)^2 \\ &= \alpha^2\beta^2 + \alpha^2\gamma^2 + \beta^2\gamma^2 + 2\alpha^2\beta\gamma + 2\alpha\beta^2\gamma + 2\alpha\beta\gamma^2. \end{aligned}$$

Then subtract to get

$$\begin{aligned} b' - e_2^2 &= -2(\alpha^2\beta\gamma + \alpha\beta^2\gamma + \alpha\beta\gamma^2) \\ b' - e_2^2 &= -2(\alpha + \beta + \gamma)(\alpha\beta\gamma) \\ b' - e_2^2 &= -2e_1e_3 \\ b' &= e_2^2 - 2e_1e_3 \\ &= (b)^2 - 2(-a)(-c) \\ &= b^2 - 2ac. \end{aligned}$$

Finally, we observe that

$$\begin{aligned} c' &= -\alpha^2\beta^2\gamma^2 \\ &= -(\alpha\beta\gamma)^2 \\ &= -e_3^2 \\ &= -(-c)^2 \\ &= -c^2. \end{aligned}$$

In conclusion, we have

$$x^3 + (2b - a^2)x^2 + (b^2 - 2ac)x - c^2 = (x - \alpha^2)(x - \beta^2)(x - \gamma^2).$$

To check that this makes sense, let's take $(a, b, c) = 1$ so that $(a', b', c') = (2b - a^2, b^2 - 2ac, -c^2) = (1, -1, -1)$. The above formula tells us that the roots of $x^3 + x^2 - x - 1$ are the squares of the roots of $x^3 + x^2 + x + 1$. To verify this, we observe that

$$x^3 + x^2 - x - 1 = (x + 1)(x - 1)^2$$

has roots $-1, 1, 1$, listed with multiplicity. On the other hand, we have the factorization

$$x^4 - 1 = (x - 1)(x^3 + x^2 + x + 1).$$

Since $x^4 - 1$ has roots $\pm 1, \pm i$[65] and $x - 1$ has root $+1$ we see that $x^3 + x^2 + x + 1$ has roots $-1, +i, -i$. And, indeed, squaring these gives $1, -1, -1$.

We end this section with a more interesting example.

The Discriminant of a Polynomial

Suppose that a polynomial $f(x) \in \mathbb{F}[x]$ has roots $\alpha_1, \ldots, \alpha_n$ in a field extension $\mathbb{E} \supseteq \mathbb{F}$:

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).$$

We define the *discriminant of f* as the product of the squares of the differences of the roots:

$$\Delta = \prod_{1 \leqslant i < j \leqslant n} (\alpha_i - \alpha_j)^2 \in \mathbb{E}.$$

The discriminant has two interesting properties:

- We have $\Delta = 0$ if and only if $f(x)$ has a repeated root.
- Since $\Delta(\alpha_1, \ldots, \alpha_n)$ is a symmetric combination of the roots of $f(x)$ it can be expressed in terms of the coefficients of $f(x)$, hence $\Delta \in \mathbb{F}$.

You are certainly familiar with the discriminant of a quadratic polynomial. Suppose that

$$x^2 + ax + b = (x - \alpha)(x - \beta),$$

so that $-a = \alpha + \beta$ and $b = \alpha\beta$. A quick computation shows that

$$\Delta = (\alpha - \beta)^2$$

[65] These are just the 4th roots of unity.

$$
\begin{aligned}
&= \alpha^2 - 2\alpha\beta + \beta^2 \\
&= (\alpha+\beta)^2 - 4(\alpha\beta) \\
&= (-a)^2 - 4b \\
&= a^2 - 4b.
\end{aligned}
$$

We conclude that the polynomial $x^2 + ax + b$ has a repeated root if and only if $a^2 - 4b = 0$. But you are probably not familiar with the formula for the discriminant of a cubic.

The Discriminant of a Cubic Polynomial

Consider a cubic polynomial $f(x) \in \mathbb{F}[x]$ with roots α, β, γ in a field extension $\mathbb{E} \supseteq \mathbb{F}$:

$$f(x) = x^3 + ax^2 + bx + c = (x-\alpha)(x-\beta)(x-\gamma).$$

By applying Waring's algorithm to the discriminant one can show that

$$
\begin{aligned}
\Delta &= (\alpha-\beta)^2(\alpha-\gamma)^2(\beta-\gamma)^2 \\
&= -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2.
\end{aligned}
$$

I do not recommend memorizing this formula. However, we note that the discriminant simplifies quite a bit in the case when $a = 0$:

$$\Delta = -4b^3 - 27c^2.$$

In this case it is more common to write $f(x) = x^3 + px + q$. Then we have the following conclusion:

$$x^3 + px + q \text{ has a repeated root} \quad \Longleftrightarrow \quad -4p^3 - 27q^2 = 0.$$

This strange expression will show up in the next chapter[a] when we discuss the general solution of cubic equations.

[a] Nope. This got posponed until Chapter 9 and then it got cut from the course.

6.7 Laplace's Proof of the FTA

We now have all of the ingredients necessary to discuss Laplace's 1795 proof of the Fundamental Theorem of Algebra. As I will mention below, Laplace's proof has a gap which was filled by Kronecker in 1887. Literally dozens of proofs of the FTA were presented in the late 1700s and early 1800s, and none of them was completely rigorous. The traditional "first correct proof" was given by Gauss in 1799, but it involved topological ideas that were not made rigorous until the twentieth century.

I like Laplace's proof because it is almost completely algebraic. (The only analysis/topology required is the Intermediate Value Theorem for polynomials.) It is also rather short and involves the concepts of symmetric polynomials and Kronecker's Theorem, which I was going to discuss anyway.

Laplace's Proof. We will prove statement (1$\mathbb{R}$) from Section 6.1:

Every non-constant $f(x) \in \mathbb{R}[x]$ has a root in $\mathbb{C}$.

So consider some non-constant $f(x) \in \mathbb{R}[x]$. The proof uses induction on the multiplicity of the prime 2 in the degree of f. Let $\deg(f) = n$ and recall that

$$\nu_2(n) = k \quad \Longleftrightarrow \quad n = 2^k m \text{ for some odd number } m.$$

This multiplicity is well-defined because of the uniqueness of prime factorization in $\mathbb{Z}$. If $\nu_2(n) = 0$ then since $\deg(f) = 2^0 m$ is odd we know from the IVT that $f(x)$ has a real root, hence it has a complex root.

So let us suppose that $\nu_2(n) = k \geqslant 1$ and assume for induction that any polynomial $g(x) \in \mathbb{R}[x]$ with $\nu_2(\deg(g)) = k - 1$ has a root in $\mathbb{C}$. As was traditional in Laplace's time, we will assume the existence of the roots of $f(x)$ and then we will show that at least one of these roots in $\mathbb{C}$. In modern language, we assume the existence of a field extension $\mathbb{E} \supseteq \mathbb{C}$ and elements $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$ such that

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).$$

The modern proof of this was given by Kronecker, and it can be paraphrased as follows: "If you pretend hard enough that $f(x)$ has roots, then it does." We will discuss Kronecker's 1887 proof in Chapter 8. It is not very "difficult" but it is very abstract; much too abstract for the year 1795.

Our goal now is to show that $\alpha_i \in \mathbb{C}$ for some i. Laplace used a very clever trick to do this. For any real number $t \in \mathbb{R}$ and for any pair of indices $1 \leqslant i < j \leqslant n$ we consider the following element of the field $\mathbb{E}$:

$$\beta_{ijt} := \alpha_i + \alpha_j + t\alpha_i\alpha_j \in \mathbb{E}.$$

Then for any real number $t \in \mathbb{R}$ we consider the following polynomial with coefficients in $\mathbb{E}$:

$$g_t(x) := \prod_{1 \leqslant i < j \leqslant n} (x - \beta_{ijt}) \in \mathbb{E}[x].$$

The first surprise is that this polynomial actually has coefficients in $\mathbb{R}$. To see this, think of β_{ijt} as a polynomial expression in the roots of $f(x)$:

$$\beta_{ijt}(\alpha_1, \ldots, \alpha_n) = \alpha_i + \alpha_j + t\alpha_i\alpha_j.$$

Let $\sigma \in S_n$ act on this expression by permuting the inputs:

$$\sigma \cdot \beta_{ijt} := \beta_{ijt}(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(n)}).$$

The transpositions act on the roots of $g_t(x)$ as follows:

$$\begin{aligned} (ij) \cdot \beta_{ijt} &= \beta_{ijt}, & & \\ (jk) \cdot \beta_{ijt} &= \beta_{ikt}, & & k \notin \{i, j\} \\ (k\ell) \cdot \beta_{ijt} &= \beta_{ijt}. & & k, \ell \notin \{i, j\}. \end{aligned}$$

We see that each transposition permutes the roots of $g_t(x)$. Since any permutation is a product of transpositions, this implies that any $\sigma \in S_n$ permutes the roots of $g_t(x)$, hence the coefficients of $g_t(x)$ are symmetric combinations of the roots $\alpha_1, \ldots, \alpha_n$ of $f(x)$. Finally, by the Fundamental Theorem of Symmetric Polynomials, each symmetric combination of the roots of $f(x)$ is a combination of the coefficients of $\mathbb{F}$ (i.e., the elementary symmetric combinations of $\alpha_1, \ldots, \alpha_n$) hence is in $\mathbb{R}$. In other words, $g_t(x) \in \mathbb{R}[x]$.

The second surprise is that the polynomial $g_t(x)$ has degree $2^{k-1}m'$ for some odd number m'. Indeed, we have assumed that $\deg(f) = n = 2^k m$ with m odd. The degree of g_t is the number of pairs of indices $1 \leqslant i < j \leqslant n$, which is

$$\binom{n}{2} = \frac{n(n-1)}{2} = \frac{2^k m(2^k m - 1)}{2} = 2^{k-1}\left[m(2^k m - 1)\right] = 2^{k-1}(\text{some odd number}).$$

Hence by induction we know that $g_t(x)$ has at least one root in $\mathbb{C}$. In other words: For each real number $t \in \mathbb{R}$ there exists at least one pair of indices $1 \leqslant i < j \leqslant n$ such that $\beta_{ijt} \in \mathbb{C}$. Now we apply the so-called Pigeonhole Principle.[66] Let $N = \binom{n}{2}$ and choose $N + 1$ real numbers $t_1, \ldots, t_{N+1} \in \mathbb{R}$, which is always possible because $\mathbb{R}$ is infinite. Now consider the following $N \times (N + 1)$ array of numbers from $\mathbb{E}$:

$\beta_{1,2,t_1}$	$\beta_{1,2,t_2}$	$\cdots$	$\beta_{1,2,t_{N+1}}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
β_{n-1,n,t_1}	β_{n-1,n,t_2}	$\cdots$	$\beta_{n-1,n,t_{N+1}}$

The numbers in the ith column are just the roots of $g_{t_i}(x)$, hence we have shown that each column contains at least one number from $\mathbb{C}$. Since there are more columns then rows, it follows that at least one row contains **at least two** elements from $\mathbb{C}$. We have proved that there exist some indices $1 \leqslant i < j \leqslant n$ and real numbers $s, t \in \mathbb{R}$ with $s \neq t$ such that β_{ijs} and β_{ijt} are both in $\mathbb{C}$:

$$\begin{aligned} \alpha_i + \alpha_j + s\alpha_i\alpha_j &\in \mathbb{C}, \\ \alpha_i + \alpha_j + t\alpha_i\alpha_j &\in \mathbb{C}. \end{aligned}$$

Subtracting these complex numbers shows that $(s - t)\alpha_i\alpha_j \in \mathbb{C}$ and hence $\alpha_i\alpha_j \in \mathbb{C}$. Then since $s \in \mathbb{R}$ we also have $s\alpha_i\alpha_j \in \mathbb{C}$ and hence

$$\alpha_i + \alpha_j = \beta_{ijs} - s\alpha_i\alpha_j \in \mathbb{C}.$$

[66] This is the principle that any function from a larger set to a smaller set must be non-injective. In German this is called Dirichlet's *Schubfachprinzip* ("drawer principle").

We have shown that $f(x)$ has a pair of roots $\alpha_i, \alpha_j \in \mathbb{E}$ satisfying $\alpha_i + \alpha_j \in \mathbb{C}$ and $\alpha_j \alpha_i \in \mathbb{C}$. Moreover, I claim that $\alpha_i \in \mathbb{C}$ or $\alpha_j \in \mathbb{C}$ (actually we will have $\alpha_i \in \mathbb{C}$ **and** $\alpha_j \in \mathbb{C}$, but we don't care). Indeed, we observe that α_i and α_j are the roots of a quadratic polynomial with complex coefficients:

$$(x - \alpha_i)(x - \alpha_j) = x^2 - (\alpha_i + \alpha_j)x + \alpha_i \alpha_j \in \mathbb{C}[x].$$

But every quadratic polynomial with complex coefficients has a complex root. (This follows from the quadratic formula and the fact that every nonzero complex number has a complex square root.) Suppose that our quadratic polynomial has a root $\gamma \in \mathbb{C}$. Then we have

$$(\gamma - \alpha_i)(\gamma - \alpha_j) = 0,$$

which implies that $\alpha_i = \gamma \in \mathbb{C}$ or $\alpha_j = \gamma \in \mathbb{C}$. □

And that is the shortest proof of the FTA that I know.

6.8 The Missing Piece: Kronecker's Theorem

As I said in the previous section, the missing piece in Laplace's proof of the FTA is the abstract existence of roots of polynomials. In Laplace's time this was usually assumed without proof because they had no idea how to make it rigorous.

The first glimpse of the proof came from Cauchy, who used Gauss' idea of "congruence mod n" to construct the complex numbers from the real numbers. Recall from Chapter 1 that the complex numbers are originally just abstract symbols:

$$\mathbb{C} := \{a + bi : a, b \in \mathbb{R}\},$$

where i is an abstract symbol satisfying the abstract formula "$i^2 = -1$". We did a significant amount of work to show that these abstract symbols can be treated as "numbers" with all of the obvious properties, including the fact that $\mathbb{C}$ is a field, which we proved by "rationalizing the denominator":

$$\frac{1}{a+bi} = \frac{1}{a+bi} \cdot \frac{1}{a-bi} = \left(\frac{a}{a^2+b^2}\right) + \left(\frac{-b}{a^2+b^2}\right) i.$$

Cauchy showed that all of this can be explained more simply by doing "modular arithmetic" in the ring of polynomials $\mathbb{R}[x]$. To be precise, we define an equivalence relation on $\mathbb{R}[x]$ called "congruence mod $x^2 + 1$" by setting

$$f(x) \equiv g(x) \bmod x^2 + 1 \quad \Longleftrightarrow \quad f(x) - g(x) = (x^2 + 1)h(x) \text{ for some } h(x) \in \mathbb{R}[x].$$

The proof that this is an equivalence relation is "exactly the same" as the proof that "congruence mod n" is an equivalence relation on $\mathbb{Z}$. Furthermore, we can show that

$$\left\{ \begin{array}{rcl} f_1(x) & \equiv & f_2(x) \bmod x^2+1 \\ g_1(x) & \equiv & g_2(x) \bmod x^2+1 \end{array} \right\} \quad \Rightarrow \quad \left\{ \begin{array}{rcl} f_1(x)f_2(x) & \equiv & g_1(x)g_2(x) \bmod x^2+1 \\ f_1(x)+g_1(x) & \equiv & f_2(x)+g_2(x) \bmod x^2+1 \end{array} \right\}$$

Again, the proof is “the same” as the proof for integers. Thus we obtain a ring of “congruence classes of polynomials modulo $x^2 + 1$”:

$$\big(\mathbb{R}[x]/(x^2+1)\mathbb{R}[x], +, \cdot, 0, 1\big)\,.$$

But what do the elements of this ring look like? In Chapter 4 we used division with remainder to prove that each element of $\mathbb{Z}/n\mathbb{Z}$ has a unique representation as r mod n for some $0 \leqslant r \leqslant n-1$, so the ring $\mathbb{Z}/n\mathbb{Z}$ has n elements. We have a similar property here.

Congruence Classes of Polynomials Modulo $x^2 + 1$

For any $f(x) \in \mathbb{R}[x]$, there exist unique real numbers $a, b \in \mathbb{R}$ satisfying

$$f(x) \equiv a + bx \bmod x^2 + 1.$$

- **Existence.** Divide $f(x)$ by x^2+1 in the ring $\mathbb{R}[x]$ to obtain $f(x) = (x^2+1)q(x) + r(x)$ for some $r(x) \in \mathbb{R}[x]$ satisfying $\deg(r) < \deg(x^2+1)$ or $r(x) = 0$. It follows from this that $r(x) = a + bx$ for some $a, b \in \mathbb{R}$. Hence we have

  $$f(x) \equiv (x^2+1)q(x) + r(x) \equiv 0q(x) + r(x) \equiv r(x) \equiv a + bx \bmod x^2 + 1.$$

- **Uniqueness.** This follows from the uniqueness of remainders in $\mathbb{R}[x]$.

Let us examine how addition and multiplication work using these standard representatives. If $f(x) \equiv a + bx$ and $g(x) \equiv c + dx$ mod $x^2 + 1$ then we have

$$f(x) + g(x) \equiv (a+bx) + (c+dx) \equiv (a+c) + (b+d)x \bmod x^2 + 1$$

and

$$\begin{aligned} f(x)g(x) &\equiv (a+bx)(c+dx) \\ &\equiv ac + (ad+bc)x + bdx^2 \\ &\equiv ac + (ad+bc)x + bc(-1) \\ &\equiv (ad+bc) + (ac-bd)x, & \bmod x^2 + 1 \end{aligned}$$

because $x^2 \equiv -1 \bmod x^2 + 1$. This shows that elements of the ring $\mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$ behave just like complex numbers, where the congruence class “x mod $x^2 + 1$” plays the role of $i \in \mathbb{C}$.

And what about the fact that $\mathbb{C}$ is a field? Instead of “rationalizing the denominator” we can use the Euclidean Algorithm. Recall our proof that the ring $\mathbb{Z}/p\mathbb{Z}$ is a field when p is prime.

Proof that $\mathbb{Z}/p\mathbb{Z}$ is a field. Consider any nonzero element $a \in \mathbb{Z}/p\mathbb{Z}$. By definition this means that $a \not\equiv 0 \bmod p$, and hence $p \nmid a$. Since p is prime and $p \nmid a$ we have $\gcd(a,p) = 1$ and it follows from the Extended Euclidean Algorithm that there exist some $x, y \in \mathbb{Z}$ satisfying

$$\begin{aligned} ax + py &= 1 \\ ax + 0y &\equiv 1 \\ ax &\equiv 1 \qquad\qquad\qquad \bmod p. \end{aligned}$$

In other words, the element $a \in \mathbb{Z}/p\mathbb{Z}$ is invertible with $a^{-1} \equiv x \bmod p$.

The same proof shows that the ring $\mathbb{R}[x]/p(x)\mathbb{R}[x]$ is a field whenever $p(x) \in \mathbb{R}[x]$ is a prime polynomial.

Proof that $\mathbb{R}[x]/p(x)\mathbb{R}[x]$ is a field. Consider any nonzero element $f(x) \in \mathbb{R}[x]/p(x)\mathbb{R}[x]$. By definition this means that $f(x) \not\equiv 0 \bmod p(x)$, and hence $p(x) \nmid f(x)$. Since $p(x)$ is prime and $p(x) \nmid f(x)$ we have $\gcd(f,p) = 1$ and it follows from the Extended Euclidean Algorithm that there exist some $a(x), b(x) \in \mathbb{R}[x]$ satisfying

$$\begin{aligned} f(x)a(x) + p(x)b(x) &= 1 \\ f(x)a(x) + 0b(x) &\equiv 1 \\ f(x)a(x) &\equiv 1 \qquad\qquad\qquad \bmod p(x). \end{aligned}$$

In other words, the element $f(x) \in \mathbb{R}[x]/p(x)\mathbb{R}[x]$ is invertible with $f(x)^{-1} \equiv a(x) \bmod p(x)$.

Finally, we observe that the polynomial $x^2 + 1$ is a prime element of $\mathbb{R}[x]$.

Proof that x^2+1 is prime.[67] Suppose for contradiction that x^2+1 is not a prime element of $\mathbb{R}[x]$. This means we can write $x^2 + 1 = f(x)g(x)$ where $f(x), g(x) \in \mathbb{R}[x]$ are both non-units, i.e., where $\deg(f) \geqslant 1$ and $\deg(g) \geqslant 1$. Since $\deg(f) + \deg(g) = \deg(fg) = \deg(x^2+1) = 2$, this implies that $\deg(f) = \deg(g) = 1$. In other words, we must have $f(x) = a + bx$ and $g(x) = c + dx$ for some $a, b, c, d \in \mathbb{R}$ with $b \neq 0$ and $d \neq 0$. But this implies that

$$\begin{aligned} (-a/b)^2 + 1 &= f(-a/b)g(-a/b) \\ (-a/b)^2 + 1 &= 0g(-a/b) \\ (-a/b)^2 + 1 &= 0 \\ (-a/b)^2 &= -1 \\ (-a/b)^2 &< 0, \end{aligned}$$

for some real number $-a/b \in \mathbb{R}$, which is impossible because $\alpha^2 \geqslant 0$ for all $\alpha \in \mathbb{R}$.

[67] Warning to sophisticated readers: Here I use the words "prime" and "irreducible" interchangeably, which is okay because $\mathbb{R}[x]$ is a Euclidean domain.

Observe that this proof is significantly different from "rationalizing the denominator" because it doesn't give us an explicit formula for the inverse, only an algorithm. This is because in fields $\mathbb{R}[x]/p(x)\mathbb{R}[x]$ corresponding to more complicated prime polynomials $p(x)$ the formula for the inverse can very intricate. See the example at the end of section 8.2.

Finally, let me give a sketch of Kronecker's Theorem. We will fill in all the details later.

Preview of Kronecker's Theorem

Let $\mathbb{F}$ be a field and consider a polynomial $f(x) \in \mathbb{F}[x]$ of degree n. Let $p(x)$ be any prime factor of $f(x)$, with $f(x) = p(x)g(x)$ for some $g(x) \in \mathbb{R}[x]$. Now consider the field of congruence classes modulo $p(x)$:

$$\mathbb{E} := \mathbb{F}[x]/p(x)\mathbb{F}[x].$$

We can think of $\mathbb{F}$ as a subfield of $\mathbb{E}$ via the injective homomorphism that sends $a \in \mathbb{F}$ to the congruence class of the constant polynomial a modulo $p(x)$. We also observe that the field $\mathbb{E}$ contains a root of $f(x)$. Indeed, we have

$$f(x) \equiv p(x)g(x) \equiv 0g(x) \equiv 0 \text{ mod } p(x),$$

so that the element x mod $p(x)$ of $\mathbb{E}$ is a root of $f(x)$. (This is what I mean by "pretending hard enough" that $f(x)$ has a root.) Let's denote the congruence class of x by $\alpha \in \mathbb{E}$. Then by Descartes' Theorem we have[68]

$$f(x) = (x - \alpha)h(x) \text{ for some } h(x) \in \mathbb{E}[x] \text{ of degree } n - 1.$$

By induction on degree we may assume that there exists a field $\mathbb{E}' \supseteq \mathbb{E} \supseteq \mathbb{F}$ and elements $\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{E}'$ such that

$$h(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_{n-1}).$$

Finally, we have

$$f(x) = (x - \alpha)(x - \alpha_1) \cdots (x - \alpha_{n-1}),$$

with elements $\alpha, \alpha_1, \ldots, \alpha_{n-1}$ in the extension field $\mathbb{E}' \supseteq \mathbb{F}$.

What did you think of that?

[68] Now you might object that I am using the letter x for two different purposes. I apologize, but I think that any other symbol would make the proof **less** understandable. Again, this is the same difficulty that we don't have a notation to distinguish between formal polynomials and evaluations of polynomials.

7 Some Group Theory

7.1 Congruence Modulo a Subgroup

As we have seen in the previous chapter, the theory of polynomial equations can be quite intricate. Laplace's 1795 proof of the Fundamental Theorem of Algebra is quite slick but it hides some deep ideas involving permutations and multivariable polynomials. Laplace's proof was built on the hard work of Euler and Lagrange, who had trouble finishing the proof due to difficulty of the notation. In 1781, Lagrange despaired that the subject had perhaps become too difficult to merit further investigation:

> *I begin to notice how my inner resistance increases little by little, and I cannot say whether I will still be doing geometry*[69] *ten years from now. It also seems to me that the mine has maybe already become too deep and unless one finds new veins it might have to be abandoned.*
>
> *Physics and chemistry now offer a much more glowing richness and much easier exploitation. Also, the general taste has turned entirely in this direction, and it is not impossible that the place of Geometry in the Academies will someday become what the role of the Chairs of Arabic at the universities is now.*

The next generation of mathematicians were only able to make progress by abandoning the old language in favor of a new, abstract point of view. The young mathematician Évariste Galois around 1830 made a brilliant breakthroughs by inventing the concept of a "group". Unfortunately, he died at the age of 21 and it took several decades for his work to be appreciated. Today the concept of an abstract group is probably the most important definition in algebra. In this chapter we will explore the abstract theory of groups before returning in further chapters to its applications in the theory of polynomial equations.

First let me remind you of the definition.

Concept of a Group

A *group* is a structure $(G, *, \varepsilon)$ consisting of a set G, a binary operation $* : G \times G \to G$, and a special element ε, satisfying the following three axioms:

- For all $a, b, c \in G$ we have $a * (b * c) = (a * b) * c$.
- For all $a \in G$ we have $a * \varepsilon = \varepsilon * a = a$.
- For all $a \in G$ there exists some $b \in G$ such that $a * b = b * a = \varepsilon$.

The element b whose existence is guaranteed by the third axiom is actually unique. Indeed, if we have $a * b = b * a = \varepsilon$ and $a * c = c * a = \varepsilon$ for some $b, c \in G$ then it follows from the first two axioms that

$$b = b * \varepsilon = b * (a * c) = (b * a) * c = \varepsilon * c = c.$$

[69] Geometry was the general 18th century term for mathematics.

This unique element is called *the inverse of* $a \in G$ and we write it as a^{-1}.

In Chapter 4 we studied the group of units $(\mathbb{Z}/n\mathbb{Z})^\times$, whose group operation is multiplication modulo n. This group satisfies the extra property that it is *abelian*, meaning that the group operation is commutative. In this chapter we want to develop a general theory that applies also to non-abelian groups, such as a the symmetric group $(S_n, \circ, \text{id})$.

The first general theorem of "group theory" involves the notion of "congruence modulo a subgroup". This is a vast generalization of modular arithmetic.

Concept of a Subgroup

Let $(G, *, \varepsilon)$ be a group and let $H \subseteq G$ be a subset. We say that H is a *subgroup* when it satisfies the following three properties:

(1) We have $\varepsilon \in H$.

(2) For all $a \in H$ we have $a^{-1} \in H$.

(3) For all $a, b \in H$ we have $a * b \in H$.

That is, a subgroup (1) contains the identity, (2) is closed under inversion, and (3) is closed under the group operation. It follows from this that the structure $(H, *, \varepsilon)$ is itself a group. You will prove on the homework that the three defining properties of a subgroup can be summarized by the following single property:

$$a, b \in H \quad \Longrightarrow \quad a^{-1} * b \in H.$$

The whole reason for defining subgroups is so we can generalize Gauss' concept of congruence.

Congruence Modulo a Subgroup

Let $(G, *, \varepsilon)$ be a group and let $H \subseteq G$ be a subgroup. Then for all $a, b \in G$ we define the following relation:[70]

$$a \equiv b \text{ mod } H \quad \Longleftrightarrow \quad a^{-1} * b \in H.$$

The properties (1), (2) and (3) of subgroups are defined precisely so that this relation is an equivalence.

Reflexive. From (1) we have $a^{-1} * a = \varepsilon \in H$ and hence $a \equiv a$ mod H for all $a \in G$.

Symmetric. For all $a, b \in G$ we have

$$\begin{aligned} a \equiv b \bmod H &\implies a^{-1} * b \in H \\ &\implies (a^{-1} * b)^{-1} \in H && \text{from (2)} \\ &\implies b^{-1} * (a^{-1})^{-1} \in H \\ &\implies b^{-1} * a \in H \\ &\implies b \equiv a \bmod H. \end{aligned}$$

Transitive. For all $a, b, c \in G$ we have

$$\begin{aligned} a \equiv b \text{ and } b \equiv c \bmod H &\implies a^{-1} * b \in H \text{ and } b^{-1} * c \in H \\ &\implies (a^{-1} * b) * (b^{-1} * c) \in H && \text{from (3)} \\ &\implies a^{-1} * (b * b^{-1}) * c \in H \\ &\implies a^{-1} * \varepsilon * c \in H \\ &\implies a^{-1} * c \in H \\ &\implies a \equiv c \bmod H. \end{aligned}$$

Let's see how this concept connects with Gauss' concept of modular arithmetic.

Subgroups of $(\mathbb{Z}, +, 0)$

Consider the additive group of integers $(\mathbb{Z}, +, 0)$ and let $H \subseteq \mathbb{Z}$ be a subgroup. In this case I claim that there exists an integer $n \geqslant 0$ such that H is just the multiples of n:

$$H = n\mathbb{Z} := \{nk : k \in \mathbb{Z}\}.$$

Before proving this, we note that the set $n\mathbb{Z}$ is, indeed, a subgroup of $(\mathbb{Z}, +, 0)$:

(1) We have $0 \in n\mathbb{Z}$ because $0 = n0$.

(2) For all $nk \in \mathbb{Z}$ we have $-(nk) = n(-k) \in n\mathbb{Z}$.

(3) For all $nk, n\ell \in n\mathbb{Z}$ we have $nk + n\ell = n(k + \ell) \in n\mathbb{Z}$.

The case $n = 0$ corresponds to the "trivial subgroup" $0\mathbb{Z} = \{0\}$, which has just one element, and the case $n = 1$ corresponds to the "full subgroup" $1\mathbb{Z} = \mathbb{Z}$.

[70] If H is not abelian then we can also define a notion of congruence by saying that $a \equiv b \bmod H$ if and only if $a * b^{-1} \in H$. In general these two relations are not the same, as we will see in the next section.

Proof. Let H be a subgroup of $(\mathbb{Z}, +, 0)$. If $H = \{0\}$ then we are done because $H = 0\mathbb{Z}$. So suppose that $H \neq \{0\}$. Since H is closed under taking negatives, it must contain a strictly positive integer. Let $n \in H$ be the smallest positive integer in H. In this case I claim that $H = n\mathbb{Z}$.

First we show that $n\mathbb{Z} \subseteq H$. Indeed, by property (1) we have $0 \in H$ and since $n \in H$ we have by property (3) that $n + n + \cdots + n \in H$ for any number of summands. Hence $nk \in H$ for all non-negative $k \geqslant 0$. Finally, by property (2) we have $n(-k) = -(nk) \in H$, so that $n\ell \in \mathbb{Z}$ for all $\ell = -k \leqslant 0$. We conclude that $nk \in H$ for all $k \in \mathbb{Z}$ and hence $n\mathbb{Z} \subseteq H$.

On the other hand, we will show that $H \subseteq n\mathbb{Z}$. To do this, consider any element $m \in H$ and divide by n to obtain

$$\begin{cases} m = nq + r, \\ 0 \leqslant r < n. \end{cases}$$

We observe that $r = m - nq$ is an element of H because $m \in H$ and $nq \in H$ (from the argument in the previous paragraph). If $r \neq 0$ then the condition $0 < r < n$ contradicts the fact that n is the **smallest positive element of** H. Hence we must have $r = 0$ and it follows that $m = nq \in n\mathbb{Z}$. Since every $m \in H$ is contained in $n\mathbb{Z}$ we conclude that $H \subseteq n\mathbb{Z}$, as desired. □

So we have seen that every subgroup of $(\mathbb{Z}, +, 0)$ has the form $n\mathbb{Z}$. Moreover, we observe that "congruence modulo the subgroup $n\mathbb{Z}$" is just the same as "congruence modulo n":[71]

$$\begin{aligned} a \equiv b \bmod n\mathbb{Z} &\iff -a + b \in n\mathbb{Z} \\ &\iff -a + b = nk \text{ for some } k \in \mathbb{Z} \\ &\iff n|(b-a) \\ &\iff n|(a-b) \\ &\iff a \equiv b \bmod n. \end{aligned}$$

Thus the concept of congruence modulo a subgroup is a generalization of modular arithmetic. It turns out that it is quite a vast generalization, which can be applied to the theory of polynomial equations and also to geometry.

7.2 Cosets and Lagrange's Theorem

In this section we will prove a theorem that is a direct generalization of Fermat's Little Theorem and Euler's Totient Theorem from Chapter 4. The key is to investigate the "shape" of congruence classes modulo a subgroup.

(Left) Cosets of a Subgroup

Let $(G, *, \varepsilon)$ be a group and let $H \subseteq G$ be a subgroup. For any element $a \in G$ we define

[71] When the group operation is addition then the expression $a^{-1} * b$ becomes $-a + b$.

the *(left) coset of H generated by a*:

$$a * H := \{a * h : h \in H\} \subseteq G.$$

I claim that these sets are precisely the equivalence classes for congruence modulo H. That is, for all $a, b \in G$ I claim that

$$a \equiv b \bmod H \qquad \Longleftrightarrow \qquad a * H = b * H.$$

We will denote the set of all cosets of H by G/H and we read this as "G mod H". The reason for the notation "G/H" is explained by Lagrange's Theorem below.

Remark: There is a corresponding notion of *right cosets* $H * a = \{h * a : h \in H\}$ which are the equivalence classes for the relation of *right congruence*, where $a \equiv b$ if and only if $a * b^{-1} \in H$. In the case of abelian groups there is no difference. In the case of non-abelian groups the difference is quite important. See the next section.

Proof. I will repeat the proof from the homework solutions. First let us suppose that $a * H = b * H$. Since $\varepsilon \in H$ we have $b = b * \varepsilon \in H$. Then since $b * H = a * H$ we have $b \in a * H$, hence $b = a * h$ for some $h \in H$. It follows that $a^{-1} * b = a^{-1} * a * h = h \in H$ and hence $a \equiv b \bmod H$.

Conversely, let us suppose that $a \equiv b \bmod H$, so that $a^{-1} * b = h$ for some $h \in H$. Applying a on the left gives $b = a * h$ and then applying h^{-1} on the right gives $a = b * h^{-1}$. Our goal is to prove that $a * H = b * H$ and for this we must prove two inclusions: $a * H \subseteq b * H$ and $b * H \subseteq a * H$. For the first inclusion, consider an arbitrary element $a * h' \in a * H$. Then since $h^{-1}, h \in H$ we have $h^{-1} * h \in H$ and hence

$$a * h' = (b * h^{-1}) * h' = b * (h^{-1} * h') \in b * H.$$

For the second inclusion, consider an arbitrary element $b * h'' \in b * H$. Then since $h, h'' \in H$ we have $h * h'' \in H$ and hence

$$b * h'' = (a * h) * h'' = a * (h * h'') \in a * H.$$

□

Since the (left) cosets of H are the equivalence classes for congruence mod H, it follows that G is a disjoint union of these cosets. Here are a few examples:

Examples of Cosets.

- Consider the group $(\mathbb{Z}, +, 0)$ and the subgroup $n\mathbb{Z} \subseteq \mathbb{Z}$. For any integer $a \in \mathbb{Z}$ the coset of $n\mathbb{Z}$ generated by a is the set

  $$\begin{aligned} a + n\mathbb{Z} &= \{a + b : b \in n\mathbb{Z}\} \\ &= \{a + nk : k \in \mathbb{Z}\} \\ &= \{a, a \pm n, a \pm 2n, a \pm 3n, \ldots\}. \end{aligned}$$

  We have seen that each coset of $n\mathbb{Z}$ can be expressed as $r + n\mathbb{Z}$ for some unique integer $0 \leqslant r < n$. Thus the set of cosets is

  $$\mathbb{Z}/n\mathbb{Z} = \{0 + n\mathbb{Z}, 1 + n\mathbb{Z}, 2 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}\}.$$

  Here we have partitioned the set $\mathbb{Z}$ into n pieces, where $\mathbb{Z}/n\mathbb{Z}$ is the set of pieces. Of course we know that $\mathbb{Z}/n\mathbb{Z}$ is much more interesting than just a set; it is a ring. In the next section we will explain why $\mathbb{Z}/n\mathbb{Z}$ is an additive group and in the next chapter we will explain why $\mathbb{Z}/n\mathbb{Z}$ is a ring.

- Consider the group $(\mathbb{R}^2, +, \mathbf{0})$ of points in the Euclidean plane under vector addition. As with any group, we always have the trivial subgroup $\{\mathbf{0}\}$ and the full subgroup $\mathbb{R}^2$. Apart from these, the most interesting subgroups are *lines through the origin*:[72]

  $$L = \{t\mathbf{v} : t \in \mathbb{R}\} \subseteq \mathbb{R}^2 \quad \text{for some “direction vector” } \mathbf{v} \in \mathbb{R}^2.$$

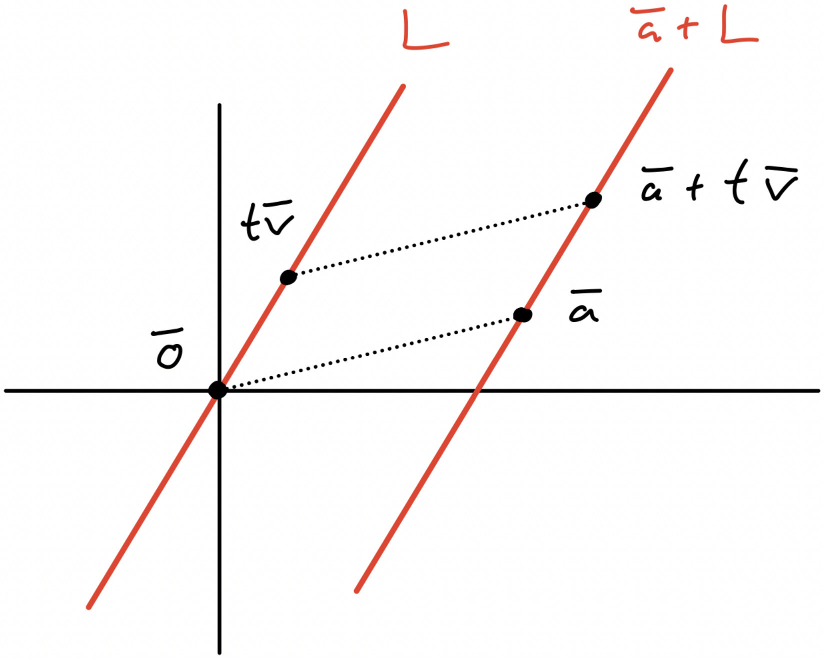
  Let’s verify that L is, indeed, a subgroup. First we observe that $\mathbf{0} = 0\mathbf{v} \in L$. Then for any $t\mathbf{v} \in L$ we observe that $-t\mathbf{v} = (-t)\mathbf{v} \in L$. Finally, we observe for any $s\mathbf{v}, t\mathbf{v} \in L$ that $s\mathbf{v} + t\mathbf{v} = (s+t)\mathbf{v} \in L$.

  And what about the cosets of the subgroup $L \subseteq \mathbb{R}^2$? I claim that the cosets are the lines parallel to L. (Apart from L itself, these lines do not pass through the origin, hence they are not subgroups.) To see this, we observe that

  $$\mathbf{a} + L = \{\mathbf{a} + t\mathbf{v} : t \in \mathbb{R}\},$$

  which is the line that is parallel to L and passes through the point $\mathbf{a}$. Here is a picture:

[72] Actually these are the only reasonable subgroups. The other subgroups come from bizarre properties of the real numbers, which you don’t want to hear about. Another name for “reasonable subgroups of $(\mathbb{R}^2, +, \mathbf{0})$” are “vector subspaces”.



- Consider the group $(\mathbb{C}^\times, \times, 1)$ of nonzero complex numbers under multiplication. I claim that the unit circle is a subgroup, called the *circle group*:[73]

  $$U(1) = \{\alpha \in \mathbb{C} : |\alpha| = 1\}.$$

  This is a subgroup because of the fact that $|\alpha\beta| = |\alpha||\beta|$ for all complex numbers $\alpha, \beta \in \mathbb{C}$. Indeed:

  - Since $|1| = 1$ we have $1 \in U(1)$.
  - Given $\alpha, \beta \in U(1)$ we have $|\alpha\beta| = |\alpha||\beta| = 1 \cdot 1$, and hence $\alpha\beta \in U(1)$.
  - Given $\alpha \in U(1)$ we have

    $$1 = |1| = |\alpha\alpha^{-1}| = |\alpha||\alpha^{-1}| = 1 \cdot |\alpha^{-1}| = |\alpha^{-1}|$$

    and hence $\alpha^{-1} \in U(1)$.

  Using polar form we can also express the circle group as follows:

  $$U(1) = \{e^{i\theta} : \theta \in \mathbb{R}\}.$$

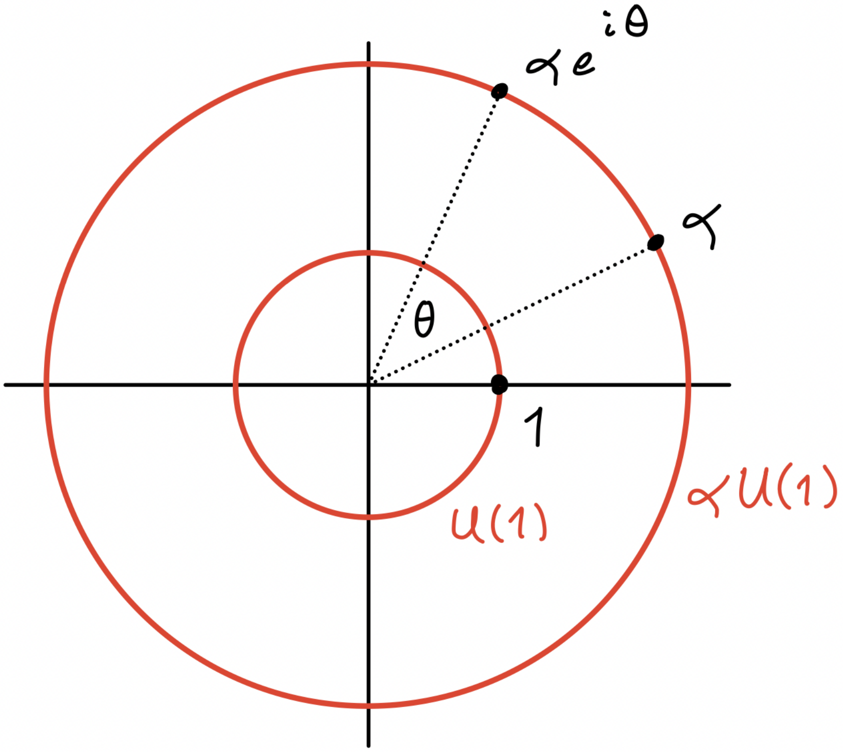
  Then the group operation becomes addition of angles: $e^{i\eta}e^{i\theta} = e^{i(\eta+\theta)}$. The cosets of $U(1)$ are the *circles centered at the origin.* (Apart from $U(1)$ itself, these circles do not

[73] There are several other notations for the circle group, such as $\mathbb{T}$, $\mathbb{S}^1$ and $SO(2)$. The notation $U(1)$ comes from the *unitary group* $U(n)$, which is the group of $n \times n$ matrices A with complex entries that satisfy the relation $AA^* = I$, where A^* is the conjugate-transpose matrix. The 1×1 unitary matrices are just numbers $\alpha \in \mathbb{C}$ satisfying $\alpha\alpha^* = 1$, hence $|\alpha| = 1$.

pass through 1, hence are not subgroups.) To see this, we observe for any $\alpha \in \mathbb{C}^\times$ that

$$\alpha U(1) = \{\alpha e^{i\theta} : \theta \in \mathbb{R}\},$$

which is the circle centered at 0 and passing through α, as in the following picture:



Each of these examples was abelian. We will see in the next section that the situation is more interesting/complicated for non-abelian groups. Now we present one of the fundamental results of group theory. The proof is quite easy since we have developed the right technology.

Lagrange's Theorem

Let $(G, *, \varepsilon)$ be a group and let $H \subseteq G$ be a subgroup. If G is finite then the size of H divides the size of G:

$$\#H \,\Big|\, \#G.$$

More specifically, if G/H is the set of cosets of H then we have

$$\#G = \#(G/H) \cdot \#H.$$

This motivates the notation "G/H" for the set of cosets, since it implies that

$$\#(G/H) = \#G/\#H.$$

Remark: This theorem can be viewed as a vast generalization of the Euler-Fermat Theorem from Chapter 4. See the section below on Cyclic Groups.

Proof. For any element $a \in G$ we have a natural function from H to $a * H$:

$$\begin{array}{rccc} \varphi : & H & \to & a * H \\ & h & \mapsto & a * h. \end{array}$$

This function is surjective by definition and it is injective because

$$\begin{aligned} \varphi(h_1) = \varphi(h_2) &\Longrightarrow a * h_1 = a * h_2 \\ &\Longrightarrow a^{-1} * (a * h_1) = a^{-1} * (a * h_2) \\ &\Longrightarrow (a^{-1} * a) * h_1 = (a^{-1} * a) * h_2 \\ &\Longrightarrow \varepsilon * h_1 = \varepsilon * h_2 \\ &\Longrightarrow h_1 = h_2. \end{aligned}$$

Hence φ is bijective. If G (and hence H) is finite, it follows that any two cosets of H have the same size; namely, $\#H$. Finally, if G/H is the set of cosets of H then since G is the disjoint union of these cosets we conclude that

$$\#G = (\#\text{ of cosets}) \cdot (\text{size of each coset}) = \#(G/H) \cdot \#H.$$

□

As the name suggests, Lagrange's Theorem has something to do with Lagrange, but he only stated a very special case. In his study of the roots of polynomials, Lagrange considered the set of permutations that leave a given polynomial invariant. Given $f(x_1, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]$, he considered the following set[74]

$$H := \{\sigma \in S_n : \sigma \cdot f = f\} \subseteq S_n.$$

It is easy to check that this set $H \subseteq S_n$ is a subgroup, hence it follows from Lagrange's Theorem that the size of H divides the size of S_n:

$$\#H \,\Big|\, n!.$$

The set H is also called the *stabilizer of f* under the *action of S_n on the set of polynomials.* We discuss the general context in the next section.

[74] Note that this set H is equal to the full group S_n if and only if f is a symmetric polynomial.

7.3 The Orbit-Stabilizer Theorem

The abstract concept of a group emerged at the end of the 19th century as a way to systematize certain ideas that are common to the following three subjects:

- **Number Theory.** As we have seen in Chapter 4, Euler's Totient Theorem $a^{\phi(n)} \equiv 1$ mod n is an example group-theoretical thinking. The Chinese Remainder Theorem can also be viewed as an isomorphism of groups $\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m \times \mathbb{Z}/n$.
- **Classical Algebra.** By this I mean the theory of solutions of polynomial equations. Indeed, this was the context that inspired the definition of groups. Évariste Galois invented the group concept in order to be able to talk more precisely about permutations of the roots of a polynomial.
- **Geometry and Physics.** Many new concepts of "geometry" emerged in the 19th century, including projective and hyperbolic geometry. In his *Erlangen program* (1872), Felix Klein proposed to organize all of these new geometries in terms of their groups of transformations, which can often be viewed as groups of matrices. This group-theoretic language became fundamental to physics in the 20th century.

Thus we have three types of groups:

- Additive groups and multiplicative groups of numbers, which are abelian.
- Groups of permutations.
- Groups of matrices.

The second and third types are based on functional composition, and are in general not abelian. These types of groups can also be viewed as "acting on" certain structures, such as polynomials or points in space. This concept of "action" is also an important part of the abstract theory of groups.

Definition of Group Action

Let $(G, *, \varepsilon)$ be a group and let X be a set. Suppose we have a function $G \times X \to X$, which we denote by $(g, x) \mapsto g \cdot x$. We call this function an *action of G on X* when the following two properties are satisfied:

(i) For all $x \in X$ we have $\varepsilon \cdot x = x$.

(ii) For all $a, b \in G$ and $x \in X$ we have $a \cdot (b \cdot x) = (a * b) \cdot x$.

Having such an action allows us to think of each group element $a \in G$ as a function $X \to X$ defined by $x \mapsto a \cdot x$.[75] Axiom (i) says that the identity element $\varepsilon \in G$ corresponds to the identity function $X \to X$ and axiom (ii) tells us that the group operation in G corresponds to the composition of functions.

Here are the two fundamental examples.

- **Permuting the Inputs of a Function.** Recall from the previous chapter that for any permutation $\sigma \in S_n$ and any polynomial $f(x_1, \ldots, x_n)$ we have another polynomial obtained by permuting the inputs of f according to σ:

  $$(\sigma \cdot f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

  It is easy to check that this satisfies the axioms of group action. Namely, the group S_n acts on the ring of polynomials $\mathbb{F}[x_1, \ldots, x_n]$ over any field $\mathbb{F}$. We can also apply this construction to more general kinds of functions with several inputs.

- **Matrices Acting on Vector Spaces.** Let $GL_n(\mathbb{F})$ denote the group of invertible $n \times n$ matrices with entries from a field $\mathbb{F}$, which is called a *general linear group*.[76] Here the group operation is matrix multiplication and the identity element is the identity matrix. Let $\mathbb{F}^n$ denote the vector space of $n \times 1$ column vectors. Then for each invertible matrix $A \in GL_n(\mathbb{F})$ we obtain a function $\mathbb{F}^n \to \mathbb{F}^n$, also defined by matrix multiplication:

  $$A \in GL_n(\mathbb{F}) \text{ and } \mathbf{v} \in \mathbb{F}^n \quad \Longrightarrow \quad A\mathbf{v} \in \mathbb{F}^n.$$

  As we saw in Chapter 1, matrix multiplication is **defined** so that the matrix AB corresponds to the composition of functions $A \circ B$. Hence this is a group action.

The following theorem could also be called the “fundamental theorem of group actions”. It is closely related to Lagrange’s Theorem from the previous section. The theorem looks quite abstract at first, but it turns out to be quite useful.

The Orbit-Stabilizer Theorem

Consider an action of a group $(G, *, \varepsilon)$ on a set X. For each element $x \in X$, its *orbit* is the set of elements of X that can be obtained from x by the action of G:[77]

$$\text{Orb}(x) = \{a \cdot x : a \in G\} \subset X.$$

For each element $x \in X$, its *stabilizer* is the set of elements of G that act trivially on x:

$$\text{Stab}(x) = \{a \in G : a \cdot x = x\} \subseteq G.$$

I claim that $\text{Stab}(x) \subseteq G$ is a subgroup, and, furthermore, that the assignment $a \cdot x \mapsto a * \text{Stab}(x)$ defines a bijection from elements of the orbit to (left) cosets of the stabilizer:

$$\begin{array}{rccc} \varphi : & \text{Orb}(x) & \to & G/\,\text{Stab}(x) \\ & a \cdot x & \mapsto & a * \text{Stab}(x). \end{array}$$

[75] The only subtlety is that two different group elements $a, b \in G$ might correspond to the same function $X \times X$. That is, we might have $a \cdot x = b \cdot x$ for all $x \in X$.

[76] Equivalently, $GL_n(\mathbb{F})$ is the set of matrices with nonzero determinant.

If the group G is finite, it then follows from Lagrange's Theorem that the sizes of G, $\text{Orb}(x)$ and $\text{Stab}(x)$ are related as follows:

$$\#G = \#\text{Orb}(x) \cdot \#\,\text{Stab}(x).$$

Proof. First we show that $\text{Stab}(x) \subseteq G$ is a subgroup.

- From (i) we have $\varepsilon \cdot x = x$ and hence $\varepsilon \in \text{Stab}(x)$.
- Suppose that $a, b \in \text{Stab}(x)$ so that $a \cdot x = x$ and $b \cdot x = x$. Then from (ii) we have

  $$(a * b) \cdot x = a \cdot (b \cdot x) = a \cdot x = x,$$

  and hence $a * b \in \text{Stab}(x)$.
- Suppose that $a \in \text{Stab}(x)$ so that $a \cdot x = x$. Then from (i) and (ii) we have

  $$\begin{aligned} a \cdot x &= x \\ a^{-1} \cdot (a \cdot x) &= a^{-1} \cdot x \\ (a^{-1} * a) \cdot x &= a^{-1} \cdot x \\ \varepsilon \cdot x &= a^{-1} \cdot x \\ x &= a^{-1} \cdot x, \end{aligned}$$

  and hence $a^{-1} \in \text{Stab}(x)$.

Next we show that the function $\varphi(a \cdot x) = a * \text{Stab}(x)$ is a bijection from the set $\text{Orb}(x)$ to the set of cosets $G/\text{Stab}(x)$. It is clearly surjective because any coset has the form $a * \text{Stab}(x) = \varphi(a \cdot x)$. The following two-way sequence of implications shows that φ is well-defined and injective:

$$\begin{aligned} a \cdot x = b \cdot x &\iff a^{-1} \cdot (a \cdot x) = a^{-1} \cdot (b \cdot x) \\ &\iff (a^{-1} * a) \cdot x = (a^{-1} * b) \cdot x \\ &\iff \varepsilon \cdot x = (a^{-1} * b) \cdot x \\ &\iff x = (a^{-1} * b) \cdot x \\ &\iff a^{-1} * b \in \text{Stab}(x) \\ &\iff a * \text{Stab}(x) = b * \text{Stab}(x). \end{aligned}$$

The last step follows from the theorem on cosets proved in the previous section.

[77] In the formalism of Hamiltonian mechanics, the evolution of a physical system can be viewed as an infinite, continuous group acting on a phase space of possible configurations. For example, the evolution of our solar system under gravity can be viewed this way, in which case the planetary orbits are literally orbits under this group action.

Finally, we apply Lagrange's Theorem. If G is finite then the subgroup $\operatorname{Stab}(x)$ is finite and the number of cosets satisfies

$$\#(G/\operatorname{Stab}(x)) = \#G/\#\operatorname{Stab}(x).$$

Then since φ is a bijection, the orbit $\operatorname{Orb}(x)$ has the same size as $G/\operatorname{Stab}(x)$, hence

$$\begin{aligned}\#(G/\operatorname{Stab}(x)) &= \#\operatorname{Orb}(x)\\ \#G/\#\operatorname{Stab}(x) &= \#\operatorname{Orb}(x)\\ \#G &= \#\operatorname{Orb}(x)\cdot\#\operatorname{Stab}(x).\end{aligned}$$

□

Are you starting to get a feel for these abstract algebra proofs? The key is to work with the symbols literally and not try to interpret them. David Hilbert was one of the leading mathematicians in the late 1800s and early 1900s, and was instrumental in raising the level of rigor in the foundations of mathematics. When it comes to rigorous proofs in geometry, he apparently said the following:

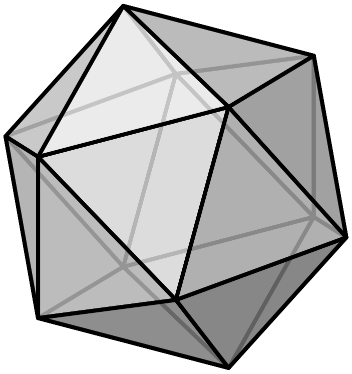
> *One must be able to say at all times — instead of points, straight lines, and planes — tables, beer mugs, and chairs.*

The interpretation comes **after** the theorem is proved.

So let's see some interpretations.

The Icosahedral Group

Let I be the group of rotational symmetries of a regular icosahedron:



This is one of the five *Platonic solids*, which are the polyhedra with maximal symmetry.[78] Suppose that the icosahedron is centered at the origin in $\mathbb{R}^3$. By definition each element of $a \in I$ is a rotational function $a : \mathbb{R}^3 \to \mathbb{R}^3$ that leaves the icosahedron "looking the same".[79] I claim that

$$\#I = 60.$$

Proof. Let V be the set of vertices of the icosahedron, so that $\#V = 12$. By definition, elements of the group I send vertices to vertices, hence the group I acts on the set V. Fix some vertex $v \in V$ and consider the orbit $\mathrm{Orb}(v) \subseteq V$ and the stabilizer $\mathrm{Stab}(v) \subseteq I$. The word "regular" in "regular icosahedron" means that any two vertices look the same with respect to some symmetry. To be precise, for any two vertices $u, v \in V$ there exists some $a \in I$ such that $u = a(v)$. In other words, we have[80]

$$\mathrm{Orb}(v) = \{a(v) : a \in I\} = V.$$

Now let's consider the stabilizer $\mathrm{Stab}(v) = \{a \in I : a(v) = v\}$. Since each element of I is a rotation, $\mathrm{Stab}(v)$ consists of rotations that do not move the vertex v. Since 5 triangles meet at each vertex, we observe that $\mathrm{Stab}(v)$ consists of the 5 rotations around v by angles $2\pi k/n5$, where $k = 0, 1, 2, 3, 4$, hence $\#\,\mathrm{Stab}(v) = 5$. Finally, we conclude from the Orbit-Stabilizer Theorem that

$$\#I = \#\mathrm{Orb}(v) \cdot \#\,\mathrm{Stab}(v) = \#V \cdot \#\,\mathrm{Stab}(v) = 12 \cdot 5 = 60.$$

□

In the next example we compute the size of the alternating group A_n.[81]

The Alternating Group, Part 2

Recall from the previous chapter that every permutation $\sigma \in S_n$ can be expressed (in many ways) as a composition of transpositions $(ij) \in S_n$. We defined $A_n \subseteq S_n$ as the set of permutations that can be expressed as a composition of an even number of transpositions:

$$A_n = \{\sigma \in S_n : \text{there exist transpositions } t_1, \ldots, t_{2k} \text{ with } \sigma = t_1 \circ \cdots \circ t_{2k}\}$$

And we showed that $A_n \subseteq S_n$ is a subgroup. Based on this definition, it is not immediately clear that $A_n \neq S_n$. That is, it is not immediately clear that there exists any permutation that is not alternating. Now we will use the Orbit-Stabilizer Theorem to

[78] The other four are the regular tetrahedron, cube, octahedron and dodecahedron.

[79] It is actually a bit tricky to show that this is a group. The identity and inverse axioms are easy, but it is difficult to show that a composition of two rotations of $\mathbb{R}^3$ is also a rotation of $\mathbb{R}^3$. See the homework.

[80] In this case we say that I acts *transitively* on the set V.

[81] This example is not unrelated to the previous. It is a surprising fact that the icosahedral group I is isomorphic to the alternating group A_5. Maybe we will prove this later.

prove that exactly half of the permutations are not alternating:

$$\#A_n = \frac{1}{2}\#S_n = \frac{n!}{2} = 2 \cdot 3 \cdots (n-1) \cdot n.$$

I will phrase the proof as a discussion. The key idea is to view A_n as the stabilizer of a very specific polynomial in n variables. Recall that S_n acts on polynomials by

$$(\sigma \cdot f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Now consider the polynomial[82]

$$\delta(x_1, \ldots, x_n) = \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j).$$

Observe what happens when a permutation $\sigma \in S_n$ acts on δ:

$$\sigma \cdot \delta = \prod_{1 \leqslant i < j \leqslant n} (x_{\sigma(i)} - x_{\sigma(j)}).$$

If $\sigma(i) < \sigma(j)$ then the factor $x_{\sigma(i)} - x_{\sigma(j)}$ also appears in δ. But if $\sigma(i) > \sigma(j)$ then the negative factor $-(x_{\sigma(i)} - x_{\sigma(j)})$ appears in δ. In other words, we must have $\sigma \cdot \delta = \pm\delta$ and the sign is determined by the number of pairs (i, j) satisfying $i < j$ and $\sigma(i) > \sigma(j)$. We give these a special name.

Inversions of a Permutation

Consider the set of pairs $T = \{(i, j) : 1 \leqslant i < j \leqslant n\}$, of size $\#T = \binom{n}{2} = n(n-1)/2$. Any permutation $\sigma \in S_n$ breaks this set into two pieces, called *inversions* and *non-inversions* of σ. The set of inversions is defined as follows:

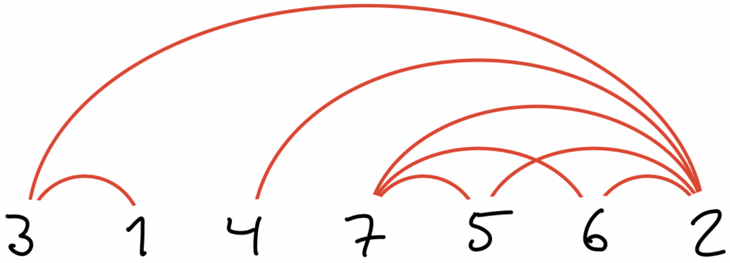
$$\mathrm{Inv}(\sigma) = \{(i, j) : 1 \leqslant i < j \leqslant n \text{ and } \sigma(i) > \sigma(j)\} \subseteq T.$$

Then from the above discussion we see that the action of σ on δ is determined by the number of inversions:

$$\sigma \cdot \delta = (-1)^{\#\mathrm{Inv}(\sigma)}\delta.$$

Inversions can be computed graphically using the one-line notation for σ. They correspond to pairs where the larger number appears to the left. For example, the following diagram shows the inversions of the permutation $\sigma = 3147562$:

[82] Note that δ is a square root of the discriminant Δ. This fact plays an important role in solvability of polynomial equations.



We can read the set of inversions from the diagram:

$$\mathrm{Inv}(\sigma) = \{(1,2),(1,7),(3,7),(4,5),(4,6),(4,7),(6,7)\}.$$

The key fact about inversions is that the *adjacent transpositions* have one inversion each. Let us define the adjacent transpositions $s_1, s_2, \ldots, s_{n-1} \in S_n$ by

$$s_1 = (12), s_2 = (23), \ldots s_{n-1} = (n, n-1).$$

Then one can see by drawing the diagram that $\mathrm{Inv}(s_i) = \{(i, i+1)\}$. It follows from this that

$$s_i \cdot \delta = (-1)^{\#\mathrm{Inv}(s_i)}\delta = (-1)^1\delta = -\delta.$$

In fact, I claim that for **any** transposition $t \in S_n$ we have $t \cdot \delta = -\delta$. To see this,[83] we first observe that the transposition (37) can be expressed as follows:

$$(37) = (34)(45)(56)(67)(56)(45)(34).$$

More generally, any transposition $t = (ij)$ with $i < j$ can be expressed as a composition of an odd number of adjacent transpositions:

$$t = (ij) = s_i \circ s_{i+1} \circ \cdots \circ s_{j-2} \circ s_{j-1} \circ s_{j-2} \circ \cdots \circ s_{i+1} \circ s_i.$$

By grouping in pairs we see that $\#\mathrm{Inv}(t) = 2(j-i-1)+1$, which is odd, and hence

$$t \cdot \delta = (-1)^{\mathrm{odd}}\delta = -\delta.$$

Now we are ready to prove that $A_n = \mathrm{Stab}(\delta)$. First, suppose that $\sigma \in A_n$ so there exist transpositions $t_1, \ldots, t_{2k}$ satisfying $\sigma = t_1 \circ \cdots \circ t_{2k}$, so that

$$\sigma \cdot \delta = t_1 \cdot (t_2 \cdot (t_3 \cdot \cdots t_{2k} \cdot \delta) \cdots) = (-1)^{2k}\delta = \delta.$$

It follows that $\sigma \in \mathrm{Stab}(\delta)$. Conversely, suppose that $\sigma \in \mathrm{Stab}(\delta)$ so that $\sigma \cdot \delta = \delta$, and assume for contradiction that $\sigma \notin A_n$. Any permutation is a composition of transpositions.

[83] This can also be seen by drawing the diagram in one-line notation.

Since σ is not a composition of an even number of transpositions, there exist an odd number of transpositions $t_1, \ldots, t_{2k+1}$ such that

$$\sigma = t_1 \circ \cdots \circ t_{2k+1},$$

and hence

$$\sigma \cdot \delta = t_1 \cdot (t_2 \cdot (t_3 \cdot \cdots t_{2k+1} \cdot \delta) \cdots) = (-1)^{2k+1}\delta = -\delta.$$

Since we also have $\sigma \cdot \delta = \delta$, this implies that $\delta = -\delta$. But then $2\delta = 0$ implies $\delta = 0$, which is a contradiction.

Finally, we observe that the orbit of δ under the action of S_n is just the two element set $\{\delta, -\delta\}$. Indeed, if σ is a composition of k transpositions then $\sigma \cdot \delta = (-1)^k\delta = \pm\delta$, hence $\mathrm{Orb}(\delta) \subseteq \{\delta, -\delta\}$. And we know that both possibilities occur because $\mathrm{id} \cdot \delta = \delta$ and $t \cdot \delta = -\delta$ for any transposition t. It follows from the Orbit-Stabilizer Theorem that

$$\begin{aligned} \#\mathrm{Orb}(\delta) \cdot \#\,\mathrm{Stab}(\delta) &= \#S_n \\ 2 \cdot \#A_n &= \#S_n \\ \#A_n &= \frac{1}{2}\#S_n. \end{aligned}$$

□

Remark: This proof is trickier than you might have expected. Secretly, we are developing some of the properties of determinants of square matrices. Indeed, the polynomial δ can be viewed as a determinant:

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j).$$

This is called *Vandermonde's determinant.*

7.4 Quotient Groups

For any subgroup $H \subseteq G$ we have studied the set G/H of (left) cosets. The definitions were inspired by our previous experience with modular arithmetic:

- Every subgroup of $(\mathbb{Z}, +, 0)$ has the form $n\mathbb{Z}$ for some $n \geqslant 0$.
- Congruence mod $n\mathbb{Z}$ is the same as congruence mod n.
- The set of cosets $\mathbb{Z}/n\mathbb{Z}$ has n elements:

$$\mathbb{Z}/n\mathbb{Z} = \{0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}\}.$$

But $\mathbb{Z}/n\mathbb{Z}$ is not only a set; it is a ring. In Chapter 4 we proved (using slightly different language) that the following operations on cosets are well-defined:

$$\begin{aligned}(a+n\mathbb{Z})+(b+n\mathbb{Z}) &:= (a+b)+n\mathbb{Z},\\ (a+n\mathbb{Z})(b+n\mathbb{Z}) &:= (ab)+n\mathbb{Z}.\end{aligned}$$

Then one can check that these operations make the set $\mathbb{Z}/n\mathbb{Z}$ into a ring with additive identity $0+n\mathbb{Z}$ and multiplicative identity $1+n\mathbb{Z}$.

We will generalize this construction to arbitrary rings in the next chapter. For now we will focus on groups, which have only a single operation. So let $(G,*,\varepsilon)$ be a group and let $H \subseteq G$ be a subgroup. Then there is a natural candidate for a group operation on the set G/H:[84]

$$(a*H)*(b*H) := (a*b)*H.$$

However, this operation might not be well-defined. Let's try to imitate the proof for $\mathbb{Z}/n\mathbb{Z}$ and see where it goes wrong. Our goal is to prove that

$$a_1*H = a_2*H \text{ and } b_1*H = b_2*H \quad\Longrightarrow\quad (a_1*b_1)*H = (a_2*b_2)*H.$$

By the results in the previous section this is equivalent to the following statement:

$$a_1^{-1}*a_2 \in H \text{ and } b_1^{-1}*b_2 \in H \quad\Longrightarrow\quad (a_1*b_1)^{-1}*(a_2*b_2) \in H.$$

So let us assume that $a_1^{-1}*a_2 = h_1$ and $b_1^{-1}*b_2 = h_2$ for some elements $h_1,h_2 \in H$. In this case we want to show that $(a_1*b_1)^{-1}*(a_2*b_2) \in H$. We begin by observing that

$$\begin{aligned}(a_1*b_1)^{-1}*(a_2*b_2) &= b_1^{-1}*a_1^{-1}*a_2*b_2\\ &= b_1^{-1}*h_1*b_2,\end{aligned}$$

but then we are stuck. If G is abelian then $b_1^{-1}*h_1*b_2 = h_1*b_1^{-1}*b_2 = h_1*h_2 \in H$, but in general $b_1^{-1}*h_1*b_2$ need not be an element of H.

The following concept will seem unmotivated at first. It was introduced by Galois (1830) in his study of polynomials. I will describe Galois' motivation at the end of this section. In the next section I will describe the modern point of view, which makes the definition seem less random.

Concept of a Normal Subgroup

Let $(G,*,\varepsilon)$ be a group and let $H \subseteq G$. Then I claim that the following two conditions are equivalent:

(N1) For all $g \in G$ and $h \in H$ we have $g*h*g^{-1} \in H$.

(N2) For all $g \in G$ the right and left cosets are equal: $g*H = H*g$.

When these conditions hold we say that H is a *normal subgroup* of G.

[84] Don't take the notation too literally. It is not necessarily true that $(a*H)*(b*H)$ is the set of elements of the form g_1*g_2 where g_1 is in the set $a*H$ and g_2 is in the set $b*H$.

Remark: Every subgroup of an abelian group is normal. This concept is only interesting for non-abelian groups such as the symmetric group or general linear groups.

Proof. (N2)⇒(N1): Suppose that (N2) is true. In order to prove (N1), consider any $g \in G$ and $h \in H$. Our goal is to show that $g * h * g^{-1} \in H$. Since $g * h \in g * H$ and since $g * H = H * g$ by (N2), we must have $g * h \in H * g$ and hence $g * h = h' * g$ for some $h' \in H$. Finally, we have

$$g * h * g^{-1} = h' \in H.$$

(N1)⇒(N2): Suppose that (N1) is true. In order to prove (N2), consider any $g \in G$. Our goal is to prove the following inclusions:

(i) $g * H \subseteq H * g$

(ii) $H * g \subseteq g * H$

To prove (i), consider any element $a \in g * H$, which must have the form $a = g * h$ for some $h \in H$. Then by (N1) we have $g * h * g^{-1} = h'$ for some $h' \in H$ and it follows that

$$a = g * h = h' * g \in H * g.$$

The proof of (ii) is similar. □

Before moving on, it is good to see at least one example of a **non-normal subgroup**. As mentioned, every subgroup of an abelian group is normal so we must begin with a non-abelian group. The smallest such group is the symmetric group of size $3! = 6$:

$$S_3 = \{\mathrm{id}, (12), (23), (13), (123), (132)\}.$$

I claim that the subset $H = \{\mathrm{id}, (12)\}$ is an example of a non-normal subgroup. Indeed, it is a subgroup because $(12) \circ (12) = \mathrm{id}$ and $(12)^{-1} = (12)$. To see that it is non-normal we observe that property (N2) fails:

$$\begin{aligned}(23) \circ H &= \{(23) \circ \mathrm{id}, (23) \circ (12)\} = \{(23), (132)\},\\ H \circ (23) &= \{\mathrm{id} \circ (23), (12) \circ (23)\} = \{(23), (123)\}.\end{aligned}$$

These two cosets are not equal because the permutations (123) and (132) are not equal.

The concept of a normal subgroup allows us to construct quotient groups.

Concept of a Quotient Group

Consider a group $(G, *, \varepsilon)$ and a **normal** subgroup $H \subseteq G$. Then the following operation is well-defined and makes the set of cosets G/H into a group:

$$(a * H) * (b * H) := (a * b) * H.$$

Proof. Suppose that $a_1 * H = a_2 * H$ and $b_1 * H = b_2 * H$ for some $a_1, a_2, b_1, b_2 \in G$. In this case we want to prove that $(a_1 * b_1) * H = (a_2 * b_2) * H$. Equivalently, we want to show that $a_1^{-1} * a_2 \in H$ and $b_1^{-1} * b_2 \in H$ implies $(a_1 * b_1)^{-1} * (a_2 * b_2) \in H$. So let us suppose that $a_1^{-1} * a_2 = h_1$ and $b_1^{-1} * b_2 = h_2$ for some $h_1, h_2 \in H$. Then we have

$$\begin{aligned}(a_1 * b_1)^{-1} * (a_2 * b_2) &= b_1^{-1} * a_1^{-1} * a_2 * b_2 \\ &= b_1^{-1} * h_1 * b_2,\end{aligned}$$

Now we will use the fact that H is **normal**. In particular, we will use the fact that $b_2 * H = H * b_2$. Since $h_1 * b_2$ is an element of $H * b_2$, it must also be an element of $b_2 * H$, so that $h_1 * b_2 = b_2 * h_3$ for some element $h_3 \in H$. Then we have

$$b_1^{-1} * h_1 * b_2 = b_1^{-1} * b_2 * h_3 = h_2 * h_3 \in H.$$

as desired. Hence the operation is well-defined.

Next we check the groups axioms:

- The coset $\varepsilon * H = H$ plays the role of the identity element. Indeed, for any element $a \in G$ we have

$$(a * H) * (\varepsilon * H) = (a * \varepsilon) * H = a * H = (\varepsilon * a) * H = (\varepsilon * H) * (a * H).$$

- Given a coset $a \in H$, the coset $a^{-1} * H$ plays the role of the inverse:

$$(a * H) * (a^{-1} * H) = (a * a^{-1}) * H = \varepsilon * H = (a * a^{-1}) * H = (a * H) * (a^{-1} * H).$$

- Finally, the associative property follows from the associative property in G:

$$\begin{aligned}(a * H) * [(b * H) * (c * H)] &= (a * H) * [(b * c) * H] \\ &= (a * [b * c]) * H \\ &= ([a * b] * c) * H \\ &= [(a * b) * H] * (c * H) \\ &= [(a * H) * (b * H)] * (c * H).\end{aligned}$$

□

That is a lot of abstraction, so there had better be some good applications. Galois originally applied this concept to the problem of solvability of polynomials. Recall the quadratic formula:

$$x^2+ax+b=0 \quad\Longleftrightarrow\quad x=\frac{-a\pm\sqrt{a^2-4b}}{2}.$$

In Chapter 1 we discussed a similar formula for cubic polynomails, called *Cardano's formula*:[85]

$$x^3+ax^2+bx+c=0 \quad\Longleftrightarrow\quad \sqrt[3]{-q+\sqrt{q^2+p^3}}+\sqrt[3]{-q-\sqrt{q^2+p^3}}+\frac{a}{3},$$

where p and q are given in terms of a,b,c by

$$p=\frac{3b-a^2}{9} \quad\text{and}\quad q=\frac{27c-9ab+2a^3}{54}.$$

As we saw, there is some difficulty to interpret this formula, but at least it gives a precise algebraic algorithm to find the roots of the polynomial x^3+ax^2+bx+c in terms of the coefficients, the field operations $+,-,\times,\div$, square roots $\sqrt{\ }$ and cube roots $\sqrt[3]{\ }$. Cardano's student Ferrari gave a similar formula for equations of degree 4. After this, the central problem of algebra was to find formulas for polynomials of higher degree.

The Central Problem of Classical Algebra

Consider the general polynomial equation of degree n:

$$x^n+a_1x^{n-1}+a_2x^{n-2}+\cdots+a_{n-1}x+a_n=0.$$

Find a precise formula to express the solutions of this equation in terms of the coefficients, the field operations $+,-,\times,\div$, and the root operations $\sqrt{\ }, \sqrt[3]{\ }, \ldots, \sqrt[n]{\ }$. If this can be done then we say that the equation is *solvable by radicals*.

The cubic and quartic formulas were discovered in the early 1500s and published by Cardano in the *Ars Magna* (1545). After this, progress stalled on the quintic equation. After the efforts of many generations of "geometers" (the word "algebraist" did not yet exist), Lagrange summarized the state of the art in his *Treatise on the solution of equations in all degrees* (1770). He suggested that the general quintic is likely unsolvable but he could not find a way to prove it. In fact, he suggested that the subject of algebra had become too complicated to be interesting.

[85] We have not yet explained where this formula comes from. We will do this in Chapter 9.

Two more generations later, Niels Henrik Abel gave the first proof of impossibility, which was extremely complicated, as expected by Lagrange. At the same time, Évariste Galois approached the problem from a completely different point of view. His idea was to ignore the details and to concentrate on the "symmetries" of the roots. Here is Galois' fundamental theorem, written in modern language.

Galois' Solvability Theorem

The general polynomial equation of degree n is solvable by radicals if and only if there exists a chain of subgroups in the symmetric group

$$S_n = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_r = \{\text{id}\},$$

satisfying the following two properties:

- For each i, G_{i+1} is a **normal** subgroup of G_i.
- Each quotient group G_i/G_{i+1} is abelian.

Abel died in 1829 at age 26 from tuberculosis and Galois died in 1831 at age 21 in a duel. After this the chain of transmission was broken and it took several decades for others to pick up on Galois' fundamental ideas. The details of algebraic computations slowly faded away and were replaced by the theory of permutations (called "substitutions"). The next major progress came with Camille Jordan's *Treatise on permutations* (1870). After this, even the concept of permutations slowly faded away and was replaced by abstract "group theory".

I will give an introduction to Galois theory in Chapter 9, but we do not have time in this course to present a full proof of Galois' theorem.

7.5 The First Isomorphism Theorem

This section is the most abstract one in the course. Here we will learn the modern language that is used to discuss normal subgroups and quotient groups. The key is to focus on the "maps between groups" instead of just the groups in themselves. This point of view was advocated by Emmy Noether in the 1920s and became standard when her ideas were published in the textbook *Modern Algebra* (1930) by van der Waerden.

Concept of Group Homomorphism

Consider two groups $(G, *, \varepsilon)$ and $(G', \bullet, \delta)$. A function $\varphi : G \to G'$ is called a *group* *homomorphism* when the following property is satisfied for all $a, b \in G$:

$$\varphi(a * b) = \varphi(a) \bullet \varphi(b).$$

This definition satisfies the following basic properties:

(1) $\varphi(\varepsilon) = \delta$

(2) $\varphi(a^{-1}) = \varphi(a)^{-1}$ for all $a \in G$.

(3) If the inverse function $\varphi^{-1} : G' \to G$ exists then it is a group homomorphism.

Proof. (1): First we observe that

$$\begin{aligned} \varepsilon * \varepsilon &= \varepsilon \\ \varphi(\varepsilon * \varepsilon) &= \varphi(\varepsilon) \\ \varphi(\varepsilon) \bullet \varphi(\varepsilon) &= \varphi(\varepsilon) \\ \varphi(\varepsilon)^{-1} \bullet \varphi(\varepsilon) \bullet \varphi(\varepsilon) &= \varphi(\varepsilon)^{-1} \bullet \varphi(\varepsilon) \\ \varphi(\varepsilon) &= \delta. \end{aligned}$$

(2): Then for any element $a \in G$ we observe that

$$\begin{aligned} a * a^{-1} &= \varepsilon \\ \varphi(a * a^{-1}) &= \varphi(\varepsilon) \\ \varphi(a) \bullet \varphi(a^{-1}) &= \delta && \text{from (1)} \\ \varphi(a)^{-1} \bullet \varphi(a) \bullet \varphi(a^{-1}) &= \varphi(a)^{-1} \bullet \delta \\ \varphi(a^{-1}) &= \varphi(a)^{-1}. \end{aligned}$$

(3): Finally, we observe for all $a', b' \in G'$ that

$$\begin{aligned} \varphi\left(\varphi^{-1}(a') * \varphi^{-1}(b')\right) &= \varphi\left(\varphi^{-1}(a')\right) \bullet \varphi\left(\varphi^{-1}(b')\right) \\ &= a' \bullet b'. \end{aligned}$$

Then applying φ^{-1} to both sides gives the desired result:

$$\varphi^{-1}(a' \bullet b') = \varphi^{-1}\left(\varphi\left(\varphi^{-1}(a') * \varphi^{-1}(b')\right)\right) = \varphi^{-1}(a') * \varphi^{-1}(b').$$

□

Concept of Group Isomorphism

By an *isomorphism of groups* we mean a bijective group homomorphism $\varphi : G \to G'$ whose inverse function $\varphi^{-1} : G' \to G$ is also a group homomorphism. As we saw in the previous proof, this second condition is redundant.

When such an isomorphism exists[86] we say that G and G' are *isomorphic*[87] and we write

$$G \cong G'.$$

We can think of an isomorphism $\varphi : G \to G'$ are a “relabeling” of the elements of a group, leaving the relationships between these elements the same.

For example, we observed in the previous chapter that the groups $(\mathbb{Z}/3\mathbb{Z}, +, 0)$ and $(A_3, \circ, \mathrm{id})$ have the same group table, up to relabeling:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\circ$	id	(123)	(132)
id	id	(123)	(132)
(123)	(123)	(132)	id
(132)	(132)	id	(123)

In other words, the function $\varphi : \mathbb{Z}/3\mathbb{Z} \to A_3$ defined by $\varphi(0) = \mathrm{id}$, $\varphi(1) = (123)$ and $\varphi(2) = (132)$ is a group isomorphism.[88]

In general, we need a between way to construct isomorphisms beyond just staring at the group tables. The following theorem can be viewed as the “fundamental theorem of group isomorphisms”.

The First Isomorphism Theorem (FIT)

Let $\varphi : (G, *, \varepsilon) \to (G', \bullet, \delta)$ be a group homomorphism. The *kernel* and *image* of φ are the following subsets of G and G', respectively:

$$\begin{aligned} \ker\varphi &:= \{a \in G : \varphi(a) = \delta\} \subseteq G, \\ \operatorname{im}\varphi &:= \{a' \in G' : \exists a \in G, \varphi(a) = a'\} \subseteq G'. \end{aligned}$$

I claim that $\operatorname{im}\varphi \subseteq G'$ is a subgroup and that $\ker\varphi \subseteq G$ is a **normal** subgroup. Furthermore, I claim that the following is a well-defined group isomorphism:

$$\begin{array}{rccc} \tilde{\varphi} : & G/\ker\varphi & \to & \operatorname{im}\varphi \\ & a * \ker\varphi & \mapsto & \varphi(a). \end{array}$$

[86] This isomorphism need not be unique. In general, a given pair of isomorphic groups will have many different isomorphisms between them.

[87] The word “isomorphism” is also used for other algebraic structures, such as rings and vector spaces. If the distinction needs to be made we will say that G and G' are *isomorphic as groups*.

[88] But it is not unique because the function $\mu(0) = \mathrm{id}$, $\mu(1) = (132)$ and $\mu(2) = (123)$ is also a group isomorphism.

Our proof uses properties (1) and (2) of homomorphisms from the previous theorem.

Proof. First we show that $\ker\varphi \subseteq G$ is a subgroup:

- **Identity.** By (1) we have $\varphi(\varepsilon) = \delta$ and hence $\varepsilon \in \ker\varphi$.
- **Inversion.** Suppose that $a \in \ker\varphi$, so that $\varphi(a) = \delta$. Then from (2) we have

  $$\varphi(a^{-1}) = \varphi(a)^{-1} = \delta^{-1} = \delta,$$

  so that $a^{-1} \in \ker\varphi$.
- **Closure under group operation.** Suppose that $a, b \in \ker\varphi$ so that $\varphi(a) = \delta$ and $\varphi(b) = \delta$. Then from the definition of group homomorphism we have

  $$\varphi(a * b) = \varphi(a) \bullet \varphi(b) = \delta \bullet \delta = \delta,$$

  so that $a * b \in \ker\varphi$.

Next we show that $\ker\varphi \subseteq G$ is normal. To do this, consider any $g \in G$ and $h \in \ker\varphi$, so that $\varphi(h) = \delta$. Then from the definition of homomorphism and property (2) we have

$$\begin{aligned}\varphi(g * h * g^{-1}) &= \varphi(g) \bullet \varphi(h) \bullet \varphi(g)^{-1}\\ &= \varphi(g) \bullet \delta \bullet \varphi(g)^{-1}\\ &= \varphi(g) \bullet \varphi(g)^{-1}\\ &= \delta.\end{aligned}$$

It follows that $g * h * g^{-1} \in \ker\varphi$, hence $\ker\varphi$ is normal by property (N1).

Next we verify that $\operatorname{im}\varphi \subseteq G'$ satisfies the subgroup axioms:

- **Identity.** By (1) we have $\delta = \varphi(\varepsilon) \in \operatorname{im}\varphi$.
- **Inversion.** Let $a' \in \operatorname{im}\varphi$, so that $a' = \varphi(a)$ for some $a \in G$. Then from (2) we have

  $$(a')^{-1} = \varphi(a)^{-1} = \varphi(a^{-1}) \in \operatorname{im}\varphi.$$

- **Closure under group operation.** Suppose that $a', b' \in \operatorname{im}\varphi$ so that $a' = \varphi(a)$ and $b' = \varphi(b)$ for some $a, b \in G$. Then from the definition of group homomorphism we have

  $$a' \bullet b' = \varphi(a) \bullet \varphi(b) = \varphi(a * b) \in \operatorname{im}\varphi.$$

Finally, we show that $\tilde{\varphi}$ is a well-defined group isomorphism. If the function $\tilde{\varphi}$ is well-defined then then it is certainly surjective. To see that it is well-defined and injective, we observe for all $a, b \in G$ that

$$a * \ker\varphi = b * \ker\varphi \Longleftrightarrow a^{-1} * b \in \ker\varphi$$

$$
\begin{aligned}
&\iff \varphi(a^{-1} * b) = \delta \\
&\iff \varphi(a)^{-1} \bullet \varphi(b) = \delta \\
&\iff \varphi(a) = \varphi(b).
\end{aligned}
$$

□

The following condition (N3) can be taken as the modern definition of normal subgroups. I believe this is the correct definition because the concept of homomorphism is so natural.

Modern Definition of Normal Subgroups

Let $(G, *, \varepsilon)$ be a group and let $H \subseteq G$. Then I claim that the following two conditions are equivalent:

(N1) For all $g \in G$ and $h \in H$ we have $g * h * g^{-1} \in H$.

(N2) For all $g \in G$ the right and left cosets are equal: $g * H = H * g$.

(N3) There exists a group G' and a group homomorphism $\varphi : G \to G'$ such that

$$\ker \varphi = H.$$

Proof. We have already seen that (N1) and (N2) are equivalent. We will show that (N3) is equivalent to both of these.

First suppose that we have a group homomorphism $\varphi : G \to G'$. Then we saw in the proof of the FIT that $\ker \varphi$ is a normal subgroup in the sense of (N1) and (N2). Conversely, let $H \subseteq G$ be a normal subgroup in the sense of (N1) and (N2). In this case, we showed in the previous section that the set of cosets G/H is a group with operation

$$(a * H) * (b * H) = (a * b) * H.$$

In fact, this definition says that the following *quotient map* is a group homomorphism:

$$
\begin{array}{rccc}
\varphi : & G & \to & G/H \\
& a & \mapsto & a * H.
\end{array}
$$

Finally, since H is the identity element of the group G/H, we observe that the kernel is H:

$$
\begin{aligned}
a \in \ker \varphi &\iff \varphi(a) = H \\
&\iff a * H = H \\
&\iff a \in H.
\end{aligned}
$$

□

The First Isomorphism Theorem is the most abstract result that we will prove in this course. The rest is applications and examples.

Some Examples.

- **The Circle Group.** Consider the multiplicative circle group

$$U(1) = \{\alpha \in \mathbb{C} : |\alpha| = 1\} = \{e^{i\theta} : \theta \in \mathbb{R}\}.$$

We have the following natural homomorphism from the additive group of real numbers:

$$\begin{array}{rccc} \varphi : & (\mathbb{R}, +, 0) & \to & (U(1), \times, 1) \\ & t & \mapsto & e^{2\pi i t}. \end{array}$$

Check that this is a homomorphism:

$$\varphi(t_1 + t_2) = e^{2\pi i(t_1+t_2)} = e^{2\pi i t_1} e^{2\pi i t_2} = \varphi(t_1)\varphi(t_2).$$

It is clearly surjective and its kernel is the additive group of integers:

$$\begin{aligned} t \in \ker\varphi &\Longleftrightarrow \varphi(t) = 1 \\ &\Longleftrightarrow e^{2\pi i t} = 1 \\ &\Longleftrightarrow t \in \mathbb{Z}. \end{aligned}$$

Hence we have an isomorphism:

$$\mathbb{R}/\mathbb{Z} = \mathbb{R}/\ker\varphi \cong \operatorname{im}\varphi = U(1).$$

The operation on the left is "addition of real numbers, modulo whole numbers". This is supposed to represent the set of angles under addition.[89]

- **Roots of Unity.** Let $\omega = e^{2\pi i/n}$ and let Ω_n denote the nth roots of unity:

$$\Omega_n = \{1, \omega, \omega^2, \ldots, \omega^{n-1}\}.$$

This is a group under multiplication: $(\Omega_n, \times, 1)$. Now consider the following surjective group homomorphism from the additive group of integers:

$$\begin{array}{rcc} \varphi : (\mathbb{Z}, +, 0) & \to & (\Omega_n, \times, 1) \\ k & \mapsto & \omega^k. \end{array}$$

The kernel is the subgroup $n\mathbb{Z} \subseteq \mathbb{Z}$:

$$k \in \ker\varphi \quad \Longleftrightarrow \quad \omega^k = 1 \quad \Longleftrightarrow \quad n|k.$$

[89] We could also have defined $\varphi(t) = e^{it}$ with kernel $2\pi\mathbb{Z} = \{2\pi k : k \in \mathbb{Z}\} \subseteq \mathbb{R}$, so that $U(1) \cong \mathbb{R}/2\pi\mathbb{Z}$. I chose to put the 2π in the homomorphism rather than in the kernel.

Hence we obtain a group isomorphism:

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\ker\varphi \cong \operatorname{im}\varphi = \Omega_n.$$

Let me emphasize the special case $n = 2$. Here, the additive group consisting of $\{0,1\}$ is isomorphic to the multiplicative group consisting of $\{1,-1\}$:

$$\begin{array}{ccc} (\{0,1\}, + \bmod 2, 0) & \cong & (\{1,-1\}, \times, 1) \\ 0 & \leftrightarrow & 1 \\ 1 & \leftrightarrow & -1. \end{array}$$

• **The Alternating Group.** In the section on the Orbit-Stabilizer Theorem we studied the action of the symmetric group S_n on the *Vandermonde polynomial*

$$\delta(x_1, \ldots, x_n) = \prod_{1\leqslant i<j\leqslant n} (x_i - x_j).$$

Let us define the *sign of the permutation* $\sigma \in S_n$ as the number $\operatorname{sgn}(\sigma) \in \{\pm 1\}$ such that

$$\sigma \cdot \delta = \operatorname{sgn}(\sigma)\delta.$$

Since the action of S_n on polynomials is "linear" (i.e., respects addition and scalar multiplication) we observe that

$$\begin{aligned} \operatorname{sgn}(\sigma\circ\tau)\delta &= (\sigma\circ\tau)\cdot\delta \\ &= \sigma\cdot(\tau\cdot\delta) \\ &= \sigma\cdot[\operatorname{sgn}(\tau)\delta] \\ &= \operatorname{sgn}(\tau)\,[\sigma\cdot\delta] && \text{scalar comes outside} \\ &= \operatorname{sgn}(\tau)\,[\operatorname{sgn}(\sigma)\delta] \\ &= [\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)]\,\delta, \end{aligned}$$

and hence $\operatorname{sgn}(\sigma\circ\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$. In other words, the sign of a permutation is a group homomorphism:

$$\operatorname{sgn} : S_n \to (\{\pm 1\}, \times, 1).$$

The kernel of this homomorphism is the same as the stabilizer of δ, hence it follows from our result in the Orbit-Stabilizer section that

$$\ker(\operatorname{sgn}) = A_n.$$

This implies that $A_n \subseteq S_n$ is a normal subgroup, and its quotient group is given by the First Isomorphism Theorem:

$$S_n/A_n = S_n/\ker(\operatorname{sgn}) \cong \operatorname{im}(\operatorname{sgn}) = \{\pm 1\}.$$

- **Euler's Isomorphism.** Recall Euler's formula:

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

From a more abstract point of view we can see this as a group isomorphism:

$$U(1) \cong SO(2).$$

To explain the notation, we define the *orthogonal groups* $O(n)$ and the *unitary groups* $U(n)$. These consist of invertible $n \times n$ matrices with real (resp. complex) entries whose inverse is equal to its transpose (resp. conjugate transpose):

$$\begin{aligned} O(n) &= \{A \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : A^T A = I\}, \\ U(n) &= \{A \in \mathrm{Mat}_{n\times n}(\mathbb{C}) : A^* A = I\}. \end{aligned}$$

These are complicated infinite groups. However, they can be described explicitly for small values of n. Note that the group $O(1)$ consists of invertible 1×1 matrices with real entries (i.e., just nonzero real numbers α) whose inverse is equal to their transpose (i.e., such that $\alpha^2 = \alpha\alpha = 1$). Hence this group only has two elements:

$$O(1) = \{\alpha \in \mathbb{R} : \alpha^2 = 1\} = \{\pm 1\}.$$

The group $U(1)$ consists of invertible 1×1 matrices with complex entries (i.e., just nonzero complex numbers) whose inverse is equal to their conjugate transpose (i.e., just their complex conjugate). In other words, $U(1)$ is just the circle group:[90]

$$U(1) = \{\alpha \in \mathbb{C} : |\alpha| = \alpha^*\alpha = 1\}.$$

The group $O(2)$ is more interesting. Consider any 2×2 matrix A with real entries. We can write this as

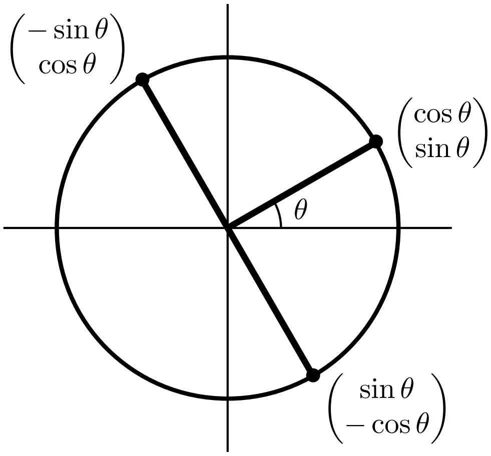
$$A = \begin{pmatrix} | & | \\ \mathbf{u} & \mathbf{v} \\ | & | \end{pmatrix},$$

for some 2×1 column vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$. If $A \in O(2)$ they we must have

$$\begin{aligned} A^T A &= I \\ \begin{pmatrix} - & \mathbf{u}^T & - \\ - & \mathbf{v}^T & - \end{pmatrix} \begin{pmatrix} | & | \\ \mathbf{u} & \mathbf{v} \\ | & | \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \begin{pmatrix} \mathbf{u}^T\mathbf{u} & \mathbf{u}^T\mathbf{v} \\ \mathbf{v}^T\mathbf{u} & \mathbf{v}^T\mathbf{v} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \begin{pmatrix} \|\mathbf{u}\|^2 & \mathbf{u} \bullet \mathbf{v} \\ \mathbf{u} \bullet \mathbf{v} & \|\mathbf{v}\|^2 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

which implies that $\mathbf{u}$ and $\mathbf{v}$ are perpendicular unit vectors. If we let $\mathbf{u} = (\cos\theta, \sin\theta)$ then this gives two possible choices for $\mathbf{v}$:

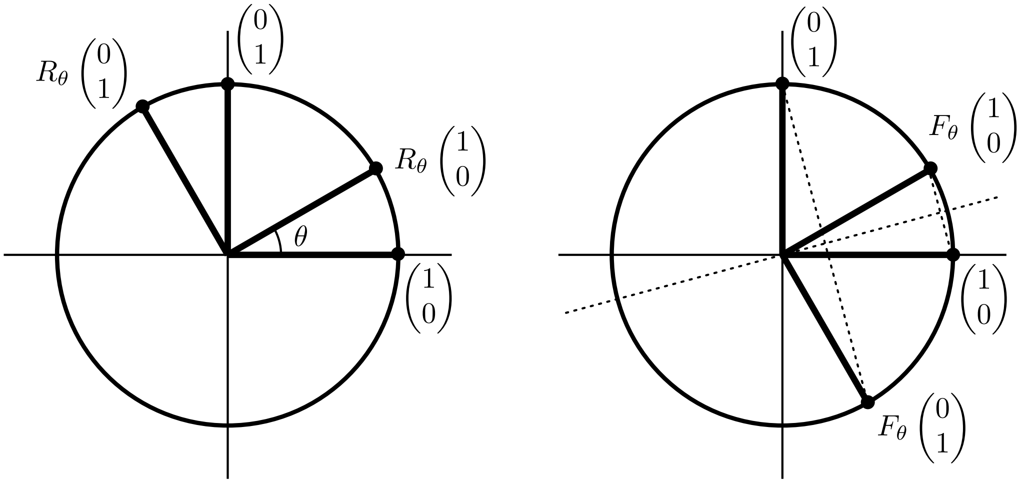
[90] And this is the explanation for the notation $U(1)$. The U stands for "unitary group".


$\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$
$\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$
θ
$\begin{pmatrix}\sin\theta\\ -\cos\theta\end{pmatrix}$

In other words, every element of the group $O(2)$ looks like one of the following matrices:

$$R_\theta = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} \quad \text{or} \quad F_\theta = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$$

If we view these as linear functions $\mathbb{R}^2 \to \mathbb{R}^2$ then the following diagram shows that R_θ is the **rotation** by angle θ and F_θ is the **reflection** across the line having angle $\theta/2$ from the positive x-axis:[91]


$R_\theta\begin{pmatrix}0\\1\end{pmatrix}$
$\begin{pmatrix}0\\1\end{pmatrix}$
$R_\theta\begin{pmatrix}1\\0\end{pmatrix}$
θ
$\begin{pmatrix}1\\0\end{pmatrix}$
$\begin{pmatrix}0\\1\end{pmatrix}$
$F_\theta\begin{pmatrix}1\\0\end{pmatrix}$
$\begin{pmatrix}1\\0\end{pmatrix}$
$F_\theta\begin{pmatrix}0\\1\end{pmatrix}$

Observe that rotations have determinant 1 while reflections have determinant -1:

$$\det R_\theta = \cos^2\theta + \sin^2\theta \quad \text{and} \quad \det F_\theta = -\cos^2\theta - \sin^2\theta = -1.$$

Furthermore, one can check the following identities:

- $R_\alpha R_\beta = R_{\alpha+\beta}$,
- $F_\alpha F_\beta = R_{\alpha-\beta}$,

[91] R is for Rotation and F is for reFlection (or Flip).

- $R_\alpha F_\beta = F_\beta R_{-\alpha} = F_{\alpha+\beta}$.

These identities show that the set of rotation matrices is a subgroup of $O(2)$, while the set of reflection matrices is not a subgroup. Indeed, the second identity above shows that the composition of two reflections is a rotation. We use the following notation for the group of rotations:[92]

$$SO(2) = \{R_\theta : \theta \in \mathbb{R}\}.$$

Finally, the classical Euler's formula implies that the following function is a group isomorphism:[93]

$$\begin{array}{rccc} \varphi : & U(1) & \to & SO(2) \\ & e^{i\theta} & \mapsto & R_\theta. \end{array}$$

• **"Special" Matrix Groups.** You have probably seen a non-rigorous treatment of matrix determinants, including the fact that

$$\det(AB) = \det(A)\det(B).$$

From a higher point of view we can see the determinant as a group homomorphism from the general linear group $GL_n(\mathbb{C})$ to the group of nonzero complex numbers:[94]

$$\det : GL_n(\mathbb{C}) \to (\mathbb{C}^\times, \times, 1).$$

The kernel of this homomorphism is the *special linear group*:

$$SL_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) : \det(A) = 1\}.$$

There is also a real version $SL_n(\mathbb{R}) \subseteq GL_n(\mathbb{R})$, corresponding to real invertible matrices with determinant 1.

Since the determinant of a transpose satisfies $\det(A^T) = \det(A)$, we observe that the determinant of an orthogonal matrix can only be 1 or -1. Indeed, if A is a real matrix then $\det(A)$ is real, and we must have

$$\begin{aligned} A^T A &= I \\ \det(A^T)\det(A) &= \det(I) \\ \det(A)\det(A) &= 1 \\ \det(A)^2 &= 1. \end{aligned}$$

Hence we obtain a group homomorphism:

$$\det : O(n) \to O(1) = \{\pm 1\}.$$

[92] See the next bullet point for an explanation of this notation.

[93] This is closely related to the results of Section 1.5.

[94] It would take us too far afield to give a rigorous definition and proof. The determinant is powerful precisely because it is hard to study.

The kernel is called the *special orthogonal group*:

$$SO(n) = \{A \in O(n) : \det(A) = 1\}.$$

Finally, since the determinant of a conjugate transpose satisfies $\det(A^*) = \det(A)^*$, we find that the determinant of a unitary matrix has length 1:

$$\begin{aligned} A^*A &= I \\ \det(A^*)\det(A) &= \det(I) \\ \det(A)^*\det(A) &= 1 \\ |\det(A)| &= 1. \end{aligned}$$

Hence we obtain a group homomorphism:

$$\det : U(n) \to U(1).$$

The kernel is called the *special unitary group*:

$$SU(n) = \{A \in U(n) : \det(A) = 1\}.$$

These groups are important in quantum physics. For example, there is a certain group homomorphism from $SU(2)$ to $SO(3)$ that is responsible for quantum spin:

$$\text{spin} : SU(2) \to SO(3).$$

- **The Alternating Group Again.** There is an important relationship between groups of permutations and groups of matrices. For any permutation $\sigma \in S_n$ we let $[\sigma]$ denote the $n \times n$ matrix whose i, j entry is 1 if $i = \sigma(j)$ and 0 otherwise. Essentially, the matrix $[\sigma]$ is obtained from the identity matrix by permuting its columns. For example, in the group S_3 we have

$$[(12)] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad [(123)] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

One can check that this assignment sends composition of permutations to multiplication of matrices:

$$[\sigma \circ \tau] = [\sigma][\tau].$$

Furthermore, it sends the inverse to the transpose and the sign to the determinant:

$$[\sigma^{-1}] = [\sigma]^T \quad \text{and} \quad [\text{sgn}(\sigma)] = \det[\sigma].$$

The first of these identities shows that permutation matrices are orthogonal. In other words, the function $\sigma \mapsto [\sigma]$ is an injective group homomorphism from S_n to $O(n)$:

$$\begin{array}{rccc} [-] : & S_n & \to & O(n) \\ & \sigma & \mapsto & [\sigma]. \end{array}$$

The second identity says that this homomorphism restricts to a homomorphism from the alternating group A_n into the special orthogonal group $SO(n)$:

$$\begin{array}{rccc} [-]: & A_n & \to & SO(n) \\ & \sigma & \mapsto & [\sigma]. \end{array}$$

In this sense, the alternating group is the "special" subgroup of permutations.

I included these examples of matrix groups for context and cultural exposure. You will not be tested on this material because there is simply no time to treat it in detail.

7.6 Cyclic Groups

The concept of *isomorphism* is an equivalence relation on the "set of all groups":[95]

- **Reflexive.** The identity function $G \to G$ is an isomorphism.
- **Symmetric.** The inverse of an isomorphism $G \to H$ is an isomorphism $H \to G$.
- **Transitive.** The composition of two homomorphisms is a homomorphism and the composition of two bijections is a bijection.

This leads to the following problem.

The Problem of Classification

For any $n \geqslant 1$, describe all groups of size n up to isomorphism.

There is always at least one group of size n; namely, $(\mathbb{Z}/n\mathbb{Z}, +, 0)$. And if p is prime then we will show that $(\mathbb{Z}/p\mathbb{Z}, +, 0)$ is the only group of size p. However, there are many groups of size 2^k and it is impossible to describe them in any coherent way. Here is a list showing the number of groups of small order, up to isomorphism:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
# of groups of size n	1	1	1	2	1	2	1	5	2	2	1	5	1	2	1	14	1

Actually, we have seen most of these groups already. The two groups of size 4 are $\mathbb{Z}/4\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, where the *direct product* is defined as the set of ordered pairs

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} = \{(a, b) : a \in \mathbb{Z}/2\mathbb{Z}, b \in \mathbb{Z}/2\mathbb{Z}\},$$

[95] There are some logical difficulties in thinking of the collection of all groups as a set. Similarly, Russell showed that there can be no such thing as the "set of all sets". If there were then we would could define

$$S = \text{the set of all sets that are not members of themselves.}$$

But this definition leads to a logical contradiction because $S \in S$ if and only if $S \notin S$.

with the componentwise group operation

$$(a,b)+(a',b')=(a+a' \bmod 2, b+b' \bmod 2).$$

This group is **not** isomorphic to $\mathbb{Z}/4\mathbb{Z}$ because, for example, every element of $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ when added to itself gives the identity element:[96]

$$(a,b)+(a,b)=(2a \bmod 2, 2b \bmod 2)=(0,0).$$

But the elements 1 and 3 in $\mathbb{Z}/4\mathbb{Z}$ do not have this property:

$$1+1\not\equiv 0 \bmod 4 \quad \text{and} \quad 3+3\not\equiv 0 \bmod 4$$

The problem of classification is deep and challenging. In this course we will only take the first step: the classification of "cyclic groups". This study begins with the concept of the order of an element.

The Order of an Element

Consider a group $(G,*,\varepsilon)$ and an arbitrary element $a\in G$. Then for any integer k we define the exponential notation

$$a^k := \begin{cases} \overbrace{a*a*\cdots*a}^{k \text{ times}} & \text{if } k\geqslant 1,\\ \varepsilon & \text{if } k=0,\\ \underbrace{a^{-1}*a^{-1}*\cdots*a^{-1}}_{-k\text{ times}} & \text{if } k\leqslant -1.\end{cases}$$

By a tedious case-by-case check,[97] one can show that

$$a^{k+\ell}=a^k*a^\ell.$$

In other words, the function $\varphi:(\mathbb{Z},+,0)\to(G,*,\varepsilon)$ defined by $\varphi(k)=a^k$ is a group homomorphism. It follows from this that the set of all powers of a is a subgroup of G, called the *cyclic subgroup generated by a*. We use the notation

$$\langle a\rangle=\{a^k : k\in\mathbb{Z}\}=\operatorname{im}\varphi\subseteq G.$$

The kernel of this homomorphism is a subgroup of $(\mathbb{Z},+,0)$, hence it must have the form $n\mathbb{Z}$ for some $n\geqslant 0$, and it follows from the FIT that

$$\langle a\rangle=\operatorname{im}\varphi\cong\mathbb{Z}/\ker\varphi=\mathbb{Z}/n\mathbb{Z}.$$

[96] If m and n are coprime then we recall from the Chinese Remainder Theorem that the group $\mathbb{Z}/m\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z}$ **is** isomorphic to $\mathbb{Z}/mn\mathbb{Z}$.

To be explicit, this isomorphism says that

$$a^k = a^\ell \quad \Longleftrightarrow \quad k \equiv \ell \bmod n.$$

We define the *order of a as an element of G as follows*:

$$\operatorname{ord}_G(a) = \#\langle a \rangle = \#(\mathbb{Z}/n\mathbb{Z}) = \begin{cases} n & \text{if } n \geqslant 1, \\ \infty & \text{if } n = 0. \end{cases}$$

In the case of finite order we have

$$\langle a \rangle = \{\varepsilon, a, a^2, \ldots, a^{n-1}\}$$

and in the case of infinite order there is no repetition among the powers of a:

$$\langle a \rangle = \{\ldots, a^{-2}, a^{-1}, \varepsilon, a, a^2, \ldots\}.$$

As a corollary we obtain a generalization of the Euler-Fermat Theorem from Chapter 4.

Generalized Euler-Fermat Theorem

Let G be a finite group. Then for any element $a \in G$ we have

$$a^{\#G} = \varepsilon.$$

In Chapter 4 we presented a proof due to Euler that holds for abelian groups. Now we give the proof for non-abelian groups.

Proof. Let $n = \operatorname{ord}_G(a) = \#\langle a \rangle$. Since G is finite we must have $n < \infty$, and by Lagrange's Theorem we must have

$$n = \#\langle a \rangle \mid \#G,$$

so that $\#G = nk$ for some $k \in \mathbb{Z}$. It follows that

$$a^{\#G} = a^{nk} = (a^n)^k = \varepsilon^k = \varepsilon.$$

□

Here are some more examples.

[97] There are 9 cases.

Examples.

- **A Matrix of Infinite Order.** Let $\mathbb{F}$ be any field and consider the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in GL_2(\mathbb{F}).$$

I claim that this matrix has infinite order. In fact, I claim that for any integer $k \in \mathbb{Z}$ we have

$$A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

from which it follows that A^k is the identity if and only if $k = 0$. To prove this, we first note that $A^0 = I$ by definition. Now assume for induction that the statement is true for A^k. In this case the statement is also true for $k+1$ because

$$A^{k+1} = AA^k = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & k+1 \\ 0 & 1 \end{pmatrix}.$$

Finally, we observe that

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which implies that

$$A^{-k} = (A^k)^{-1} = \begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix}.$$

- **Rotation Matrices.** Recall the rotation matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in GL_2(\mathbb{R}),$$

and recall the identities

(a) $R_\alpha R_\beta = R_{\alpha+\beta}$ for all $\alpha, \beta \in \mathbb{R}$.

(b) $R_\alpha = R_\beta$ if and only if $\alpha - \beta = 2\pi k$ for some $k \in \mathbb{Z}$.

If $\theta = 2\pi k/n$ for some $k, n \in \mathbb{Z}$ with $n \geqslant 1$ then it follows from (a) and (b) that

$$R_\theta^n = R_{n\theta} = R_{2\pi k} = I,$$

so that R_θ has order dividing n. The precise order is $n/\gcd(k,n)$, which we will prove below.

If θ is an irrational multiple of 2π, say $\theta = 2\pi\alpha$, then I claim that R_θ has infinite order. To see this, suppose for contradiction that $R_\theta^n = I$ for some $n \geqslant 1$. Then from (a) we have

$$R_{n\theta} = R_\theta^n = I,$$

and from (b) we conclude that $n\theta = 2\pi\alpha$ is an integer multiple of 2π, say $n\theta = 2\pi k$. But then we have

$$\begin{aligned} n2\pi\alpha &= 2\pi k \\ \alpha &= k/n, \end{aligned}$$

which contradicts the fact that α is irrational.

• **Primitive nth Roots of Unity.** Consider the group $(\mathbb{C}^\times, \times, 1)$ and let $\zeta \in \mathbb{C}^\times$ be an element of finite order n so that ζ has n distinct powers:

$$\langle\zeta\rangle = \{1, \zeta, \zeta^2, \ldots, \zeta^{n-1}\}.$$

Since $\zeta^n = 1$ we see that ζ is an nth root of unity. Furthermore, we observe that every power of ζ is an nth root of unity:

$$(\zeta^k)^n = (\zeta^n)^k = 1.$$

It follows that $\langle\zeta\rangle$ is the full group of nth roots of unity:

$$\Omega_n = \langle\zeta\rangle = \{1, \zeta, \zeta^2, \ldots, \zeta^{n-1}\}.$$

Any element $\zeta \in \mathbb{C}^\times$ of finite order n is called a *primitive nth root of unity.* For example, the usual $\omega = e^{2\pi i/n}$ is a primitive nth root or unity. But there are others.

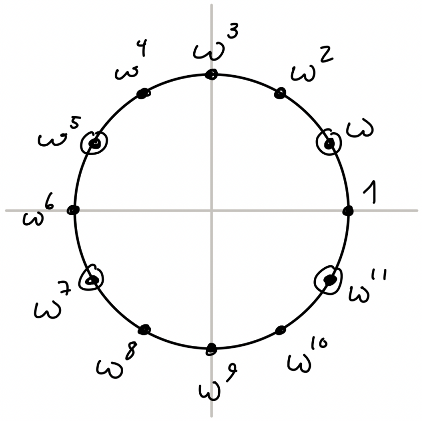
Below we will prove that there are exactly $\phi(n)$ primitive nth roots of unity, where $\phi(n)$ is Euler's totient function

$$\phi(n) = \{k \in \mathbb{Z} : 1 \leqslant k \leqslant n \text{ and } \gcd(k, n) = 1\}$$

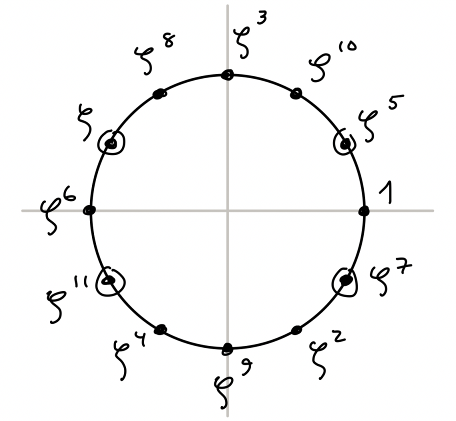
In fact, if ζ is any primitive nth root of unity then we will prove that the full set of primitive roots is $\{\zeta^k : 1 \leqslant k \leqslant n \text{ and } \gcd(k, n) = 1\}$. For example, consider the primitive 12th root $\omega = e^{2\pi i/12}$. The numbers below 12 that are coprime to 12 are $1, 5, 7, 11$. Hence there are exactly four primitive 12th roots of unity:

$$\omega^1, \omega^5, \omega^7, \omega^{11}.$$

Here is a picture showing the primitive roots:



To verify that $\zeta = \omega^5$ is a primitive root, the following diagram shows that the every 12th root of unity is a power of ζ. Note that multiplying by ζ moves 5 steps counterclockwise around the circle:



Moreover, this picture shows that the $1, 5, 7, 11$th powers of ζ are the full set of primitive roots, as expected. From an algebraic point of view the primitive 12th roots of unity are indistinguishable, i.e., they satisfy all of the same algebraic identities.[98]

Messing around with intricate computations such as these eventually gave rise to the abstract definition of a cyclic group.

Concept of a Cyclic Group

We say that a group $(G, *, \varepsilon)$ is *cyclic* if there exists an element $a \in G$ such that

$$G = \langle a \rangle = \{a^k : k \in \mathbb{Z}\}.$$

In this case we say that a is a *generator* for G. If G is cyclic then it follows from the above discussion that

$$G \cong \mathbb{Z} \quad \text{or} \quad G \cong \mathbb{Z}/n\mathbb{Z} \;\text{ for some } n \geqslant 1.$$

In particular, this implies that any two cyclic groups of the same size are isomorphic.

It turns out that any group of prime size is cyclic, which implies that there is only one group of size p up to isomorphism.

[98] This comment will be made precise later. It follows from the fact that the primitive nth roots of unity are the roots of an irreducible polynomial $\Phi_n(x)$ over $\mathbb{Q}$, called the nth cyclotomic polynomial.

Groups of Prime Order

Let $p \geqslant 2$ be prime. Then any group of size p is isomorphic to $\mathbb{Z}/p\mathbb{Z}$.

Proof. Let G be a group of prime size $p \geqslant 2$ and let $a \in G$ be a non-identity element, so that $\#\langle a\rangle \neq 1$. By Lagrange's Theorem we must have $\#\langle a\rangle | p$. However, since p is prime and $\#\langle a\rangle \neq 1$ this implies that $\#\langle a\rangle = p$, and hence $G = \langle a\rangle$. Finally, the group homomorphism $\varphi : \mathbb{Z} \to G$ defined by $\varphi(k) = a^k$ is surjective with kernel $p\mathbb{Z}$, hence

$$G = \langle a\rangle = \operatorname{im}\varphi \cong \mathbb{Z}/\ker\varphi = \mathbb{Z}/p\mathbb{Z}.$$

□

As I said at the beginning of this section, the structure of theory of groups is deep and challenging. In this course we will only consider the structure theory of cyclic groups. If a group G has size n then Lagrange's Theorem says that any subgroup has size d for some positive divisor $d|n$. However, for a given divisor $d|n$ we are not guaranteed that a subgroup of size d exists. For example, the alternating group A_4 has size 12 but one can check that it does not have a subgroup of size 6. It is also possible that for a given divisor $d|n$ there exist many subgroups of size d.

The following theorem says that cyclic groups satisfy a sort of converse to Lagrange's Theorem. We will state and prove the theorem in its abstract form, and then we will apply it to the original example, which is the group of nth roots of unity.

The Fundamental Theorem of Cyclic Groups (FTCG)

Let $G = \langle a\rangle$ be a cyclic group of finite size n. Then for any divisor $d|n$ there exists a unique subgroup of size d; namely,

$$\langle a^{n/d}\rangle \subseteq G.$$

In particular, this says that every subgroup of a cyclic group is itself cyclic.

The proof will require the following lemma.

Lemma: The Order of a Power

Let $(G, *, \varepsilon)$ be a group and let $a \in G$ be an element of finite order n. Then

(i) For any $k \in \mathbb{Z}$ we have $\langle a^k \rangle = \langle a^{\gcd(k,n)} \rangle$.

(ii) For any positive divisor $d|n$ we have $\#\langle a^d \rangle = n/d$.

(ii) For any $k \in \mathbb{Z}$ we have $\#\langle a^k \rangle = n/\gcd(k,n)$.

Proof of the Lemma. (i): Let $d = \gcd(k,n)$ with $k = dk'$. Our goal is to show that $\langle a^k \rangle = \langle a^d \rangle$. To prove $\langle a^k \rangle \subseteq \langle a^d \rangle$, consider any power of a^k, say $(a^k)^m = a^{km}$. Then we have

$$a^{km} = a^{dk'm} = (a^d)^{k'm} \in \langle a^d \rangle.$$

To prove $\langle a^d \rangle \subseteq \langle a^k \rangle$, consider any power of a^d, say $(a^d)^m = a^{dm}$. Since $d = \gcd(k,n)$ we know from Bézout's Identity that $d = kx + ny$ for some $x, y \in \mathbb{Z}$. Hence we have

$$a^{dm} = a^{(kx+ny)m} = (a^k)^{xm} * (a^n)^{ym} = (a^k)^{xm} * (\varepsilon)^{ym} = (a^k)^{xm} \in \langle a^k \rangle.$$

(ii): Let $d|n$ with $n = dd'$. Our goal is to prove that the first d' powers of a^d are distinct:

$$\varepsilon, a^d, (a^d)^2, \ldots, (a^d)^{d'-1}.$$

Suppose for contradiction that we have $0 \leqslant k < \ell < d'$ with $(a^d)^k = (a^d)^\ell$, so that

$$\begin{aligned}(a^d)^\ell &= (a^d)^k \\ a^{d\ell} &= a^{dk} \\ a^{d(\ell-k)} &= \varepsilon.\end{aligned}$$

Since $0 \leqslant k < \ell < d'$ we have $0 < \ell - k < d'$ and hence $0 < d(\ell - k) < dd' = n$. But since a has order n, the identity $a^{d(k-\ell)} = \varepsilon$ implies that $d(k-\ell)$ is a multiple of n. Contradiction.

(iii): Since $\gcd(k,n)$ is a divisor of n, it follows from (i) and (ii) that

$$\#\langle a^k \rangle = \#\langle a^{\gcd(k,n)} \rangle = n/\gcd(k,n).$$

□

Proof of the FTCG. Let $G = \langle a \rangle$ be cyclic of size n and consider a divisor $d|n$ with $n = dd'$. We will prove that $\langle a^{d'} \rangle \subseteq G$ is the unique subgroup of size d, in three steps:

(a) The subgroup $\langle a^{d'} \rangle$ has size d.

(b) Any cyclic subgroup $H \subseteq G$ of size d is equal to $\langle a^{d'} \rangle$.

(c) Any subgroup of G is cyclic.

(a): From the Lemma (ii) we have

$$\#\langle a^{d'}\rangle = n/d' = d.$$

(b): Consider any cyclic subgroup $H = \langle b\rangle \subseteq G$ of size d. Since $G = \langle a\rangle$ we know that $b = a^k$ for some $k \in \mathbb{Z}$. From the Lemma (i) and (ii) we have

$$H = \langle a^k\rangle = \langle a^{\gcd(k,n)}\rangle$$

and

$$\#H = \#\langle a^{\gcd(k,n)}\rangle = n/\gcd(k,n).$$

Since we have assumed that $\#H = d$ this implies that $\gcd(k,n) = d'$ and hence $H = \langle a^{d'}\rangle$.

(c): Consider any subgroup $H \subseteq G$. If H has size 1 then it is cyclic: $H = \langle \varepsilon\rangle$. So assume that $\#H \geqslant 2$, which means that $a^k \in H$ for some $0 < k \leqslant n$. Let $m > 0$ be the smallest positive integer such that $a^m \in H$. In this case we will show that $H = \langle a^m\rangle$, and hence H is cyclic.

To prove this, we first observe that any power of a^m is in H because H is a subgroup. Hence $\langle a^m\rangle \subseteq H$. On the other hand, we will show that any element of H is a power of a^m. So consider any element $b \in H$. Since $G = \langle a\rangle$ we can write $b = a^k$ for some $k \in \mathbb{Z}$. Divide k by m to obtain

$$\begin{cases} k = mq + r, \\ 0 \leqslant r < m. \end{cases}$$

We observe that $a^r \in H$ because $a^{-m} \in H$ and hence

$$a^r = a^{k-mq} = a^k(a^{-m})^q \in H.$$

But if $r \neq 0$ then this contradicts the definition of m. It follows that $r = 0$ and hence $b = a^k = a^{mq} = (a^m)^q$ is a power of a^m, as desired. □

Remark: The proof of part (c) recalls our proof that every subgroup of $(\mathbb{Z}, +, 0)$ has the form $n\mathbb{Z}$. In fact, there is a way to prove the FTCG by comparing subgroups of $\mathbb{Z}/n\mathbb{Z}$ with subgroups of $\mathbb{Z}$. To be specific, one can show that any homomorphism $\varphi : G \to G'$ induces a bijection between subgroups of $\operatorname{im}\varphi$ and subgroups of G that contain $\ker\varphi$. This is called the *correspondence theorem*. Then one can prove the FTCG by considering a surjective homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ with $\ker\varphi = n\mathbb{Z}$. This proof is more conceptual, but ultimately it would have taken longer to write out all of the details.

Now we discuss the application of the FTCG to roots of unity. At the end of Chapter 3 we discussed the problem of factoring the polynomial $x^n - 1$ in the ring $\mathbb{Z}[x]$ and we observed some strange behavior. Now we are able to discuss this factorization in full detail. First we prove a theorem on roots of unity. This was first worked out by Gauss in the final chapter of his *Disquisitiones Arithmeticae* (1798), written when he was just 21 years old.

Primitive nth Roots of Unity

For all $n \geqslant 1$ let $(\Omega_n, \times, 1)$ denote the group of nth roots of unity. Recall that $\Omega_n = \langle \omega \rangle$ for $\omega = e^{2\pi i/n}$. More generally, we say that $\zeta \in \Omega_n$ is a *primitive* n*th root of unity* if it generates the whole group. We denote the set[99] of primitive roots by

$$\Omega_n' = \{\zeta \in \Omega_n : \langle \zeta \rangle = \Omega_n\}.$$

Then we have the following:

(1) The subgroups of Ω_n are just Ω_d for positive divisors $d|n$.

(2) For any fixed primitive root $\zeta \in \Omega_n'$ I claim that we have

$$\Omega_n' = \{\zeta^k : 1 \leqslant k \leqslant n \text{ and } \gcd(k,n) = 1\},$$

and hence the number of primitive roots is given by Euler's totient function $\phi(n)$.

(3) The set of nth roots of unity can be expressed as the disjoint union of primitive dth roots of unity for positive divisors $d|n$:

$$\Omega_n = \coprod_{d|n} \Omega_d'.$$

Then it follows from part (2) that

$$n = \#\Omega_n = \sum_{d|n} \#\Omega_d' = \sum_{d|n} \phi(d).$$

(4) More precisely, for any fixed primitive root $\zeta \in \Omega_n'$ I claim that

$$\Omega_d' = \{\zeta^k : 1 \leqslant k \leqslant n \text{ and } \gcd(k,n) = n/d\}.$$

Proof. (1): For each divisor $d|n$ we recall from the FTCG that $\Omega_n = \langle \omega \rangle$ has a unique subgroup of size d; namely $\langle \omega^{n/d} \rangle$. I claim that

$$\langle \omega^{n/d} \rangle = \Omega_d.$$

Indeed, we know that $\#\langle \omega^{n/d} \rangle = d$, so we will be done if we can show that $\langle \omega^{n/d} \rangle \subseteq \Omega_d$. In other words, we want to show that every power of $\omega^{n/d}$ is a dth root of unity. And this is straightforward:

$$\left(\left(\omega^{n/d}\right)^k\right)^d = \omega^{nk} = (\omega^n)^k = 1^k = 1.$$

[99] This is **not** a subgroup of Ω_n.

(2): Fix some primitive root $\zeta \in \Omega_n'$[100] so that $\Omega_n = \{1, \zeta, \zeta^2, \ldots, \zeta^{n-1}\}$. I claim that ζ^k is a primitive root if and only if $\gcd(k,n) = 1$. To see this, we recall from Lemma (iii) that

$$\#\langle \zeta^k \rangle = n/\gcd(k,n).$$

It follows that

$$\langle \zeta^k \rangle = \Omega_n \quad \Leftrightarrow \quad \#\langle \zeta^k \rangle = n \quad \Leftrightarrow \quad \gcd(k,n) = 1.$$

(3): Every nth root of unity $\zeta \in \Omega_n$ generates a cyclic subgroup $\langle \zeta \rangle \subseteq \Omega_n$, which from part (1) must be $\langle \zeta \rangle = \Omega_d$ for some divisor $d|n$. Thus we can express Ω_n as a disjoint union:

$$\Omega_n = \coprod_{d|n} \{\zeta \in \Omega_n : \langle \zeta \rangle = \Omega_d\} = \coprod_{d|n} \Omega_d'.$$

(4): Finally, let $\zeta \in \Omega_n'$ so that $\langle \zeta \rangle = \Omega_n$. Then for any $1 \leqslant k \leqslant n$ we have

$$\begin{aligned} \zeta^k \in \Omega_d' &\Longleftrightarrow \langle \zeta^k \rangle = \Omega_d \\ &\Longleftrightarrow \#\langle \zeta^k \rangle = d \\ &\Longleftrightarrow n/\gcd(k,n) = d \\ &\Longleftrightarrow \gcd(k,n) = n/d. \end{aligned}$$

□

We can see the identity $\Omega_n = \coprod_{d|n} \Omega_d'$ more clearly by reducing each of the fractions $\{k/n : 1 \leqslant k \leqslant n\}$ to lowest terms. If the reduced form of the fraction k/n has denominator d then ω^k is a primitive dth root of unity. For example, we have

$$\left(\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}, \frac{6}{6}\right) \xrightarrow{\text{reduce}} \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, \frac{1}{1}\right)$$

If $\omega = e^{2\pi i/6}$ (or any primitive 6th root of unity) then the primitive dth roots are

$$\begin{aligned} \Omega_1' &= \{\omega^6\}, \\ \Omega_2' &= \{\omega^3\}, \\ \Omega_3' &= \{\omega^2, \omega^4\}, \\ \Omega_6' &= \{\omega^1, \omega^5\}. \end{aligned}$$

Somewhat miraculously, this decomposition of the 6th roots of unity tells us how to factor the polynomial $x^6 - 1$ over the integers:

$$\begin{aligned} x^6 - 1 &= (x-\omega^1)(x-\omega^2)(x-\omega^3)(x-\omega^4)(x-\omega^5)(x-\omega^6) \\ &= \left[(x-\omega^6)\right]\left[(x-\omega^3)\right]\left[(x-\omega^2)(x-\omega^4)\right]\left[(x-\omega^1)(x-\omega^5)\right] \\ &= \left[x-1\right]\left[x+1\right]\left[x^2+x+1\right]\left[x^2-x+1\right]. \end{aligned}$$

Here is the general theorem.

[100] For example, we could take $\zeta = \omega = e^{2\pi i/n}$.

Cyclotomic Polynomials

For all $n \geqslant 1$, we define the *nth cyclotomic polynomial* $\Phi_n(x)$, whose roots are the primitive nth roots of unity:

$$\Phi_n(x) = \prod_{\zeta \in \Omega'_n} (x - \zeta) \in \mathbb{C}[x].$$

From the definition we see that the coefficients of $\Phi_n(x)$ are in $\mathbb{C}$, but I claim that in fact the coefficients are in $\mathbb{Z}$. Furthermore, I claim that the prime factorization of $x^n - 1$ in the ring $\mathbb{Z}[x]$ is given by the product of cyclotomic polynomials over the divisors of n:

$$x^n - 1 = \prod_{d|n} \Phi_d(x),$$

Note that the polynomial $\Phi_n(x)$ has degree $\phi(n)$. Thus taking degrees on both sides gives

$$n = \sum_{d|n} \phi(d).$$

Partial Proof. From the identity $\Omega_n = \coprod_{d|n} \Omega'_d$ we have

$$\prod_{d|n} \Phi_d(x) = \prod_{d|n} \prod_{\zeta \in \Omega'_d} (x - \zeta) = \prod_{\zeta \in \Omega_n} (x - \zeta) = x^n - 1.$$

Now we will use this to prove by induction that $\Phi_n(x) \in \mathbb{Z}[x]$ for all $n \geqslant 1$. The base case is true because $\Phi_1(x) = x - 1$ has integer coefficients. Now suppose for induction that $n \geqslant 2$ and that $\Phi_k(x)$ has integer coefficients for all $1 \leqslant k < n$. From the previous identity we have

$$x^n - 1 = \Phi_n(x) f(x),$$

where $f(x)$ is the product of $\Phi_d(x)$ over all divisors $d|n$ except $d = n$. By induction, this $f(x)$ is a product of polynomials with integer coefficients, hence $f(x)$ itself has integer coefficients.

Next we observe that $f(x)$ has leading coefficient 1 since it is a product of polynomials $\Phi_d(x)$, each with leading coefficient 1. This means that we can perform long division[101] in the ring $\mathbb{Z}[x]$ to obtain polynomials $q(x), r(x) \in \mathbb{Z}[x]$ satisfying

$$\begin{cases} x^n - 1 = q(x)f(x) + r(x), \\ r(x) = 0 \text{ or } \deg(r) < \deg(f). \end{cases}$$

[101] In Chapter 2 we only discussed long division over a field. It turns out that long division can be performed over any ring, as long as the leading coefficient of the divisor is a unit. Since 1 is a unit and $\mathbb{Z}$, we can divide by any polynomial with leading coefficient 1.

On the other hand, we have the identity $x^n - 1 = \Phi_n(x)f(x) + 0$ in the ring $\mathbb{C}[x]$. It follows from uniqueness of quotients in $\mathbb{C}[x]$ that

$$\Phi_n(x) = q(x) \in \mathbb{Z}[x].$$

□

The only thing remaining is to prove that each cyclotomic polynomial is irreducible over the ring $\mathbb{Z}$.[102] This is quite tricky. Gauss gave a complicated proof that $\Phi_p(x)$ is irreducible for prime p, and this proof was later simplified by Gotthold Eisenstein.[103] As far as I am aware, Gauss did not prove that $\Phi_n(x)$ is irreducible for all n. The first proofs were given 50 years later by Kronecker and Dedekind, and they are too complicated for us.

The polynomials $\Phi_n(x)$ have surprisingly random behavior. There is no closed formula for their coefficients, but they can be computed recursively using the identity $x^n - 1 = \prod_{d|n} \Phi_n(x)$. Here are the first twelve:

n	$\Phi_n(x)$
1	$x - 1$
2	$x + 1$
3	$x^2 + x + 1$
4	$x^2 + 1$
5	$x^5 + x^4 + x^3 + x^2 + x + 1$
6	$x^2 - x + 1$
7	$x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$
8	$x^4 + 1$
9	$x^6 + x^3 + 1$
10	$x^4 - x^3 + x^2 - x + 1$
11	$x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$
12	$x^4 - x^2 + 1$

You might see some patterns here. For example, for any prime p we observe that

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1.$$

This is easy to prove. Since the only divisors of p are 1 and p we must have

$$\begin{aligned}
\Phi_1(x)\Phi_p(x) &= x^p - 1 \\
(x-1)\Phi_p(x) &= x^p - 1 \\
\Phi_p(x) &= (x^p - 1)/(x-1) \\
\Phi_p(x) &= x^{p-1} + x^{p-2} + \cdots + x + 1.
\end{aligned}$$

[102] This is equivalent to being irreducible over the field $\mathbb{Q}$. The equivalence is called *Gauss' Lemma*. It is not that tricky to prove but we have run out of time.

[103] It uses the so-called *Eisenstein criterion* for irreducibility.

Another pattern you might observe is that the nonzero coefficients are all either 1 or -1, and you might conjecture that this is always the case. But actually this is false. The first cyclotomic polynomial with a coefficient larger than 1 is

$$\begin{aligned}\Phi_{105}(x) = x^{48}+x^{47}+x^{46}-x^{43}-x^{42}-\mathbf{2}x^{41}-x^{40}-x^{39}+x^{36}+x^{35}+x^{34}+x^{33}+x^{32}+x^{31}-x^{28} \\ -x^{26}-x^{24}-x^{22}-x^{20}+x^{17}+x^{16}+x^{15}+x^{14}+x^{13}+x^{12}-x^{9}-x^{8}-\mathbf{2}x^{7}-x^{6}-x^{5}+x^{2}+x+1.\end{aligned}$$

And it is known that arbitrarily large coefficients can occur.

We end this chapter by completing our discussion of the 12th roots of unity.

Example: 12th Roots of Unity. Let $\omega = e^{2\pi i/12}$ so that

$$\Omega_{12} = \{\omega^1, \omega^2, \omega^3, \omega^4, \omega^5, \omega^6, \omega^7, \omega^8, \omega^9, \omega^{10}, \omega^{11}, \omega^{12}\}.$$

The subgroups of Ω_{12} are

$$\begin{aligned}
\Omega_1 &= \langle \omega^{12} \rangle = \{1\}, \\
\Omega_2 &= \langle \omega^{6} \rangle = \{\omega^6, \omega^{12}\} = \{-1, 1\}, \\
\Omega_3 &= \langle \omega^{4} \rangle = \{\omega^4, \omega^8, \omega^{12}\} = \left\{(-1+i\sqrt{3})/2, (-1-i\sqrt{3})/2, 1\right\}, \\
\Omega_4 &= \langle \omega^{3} \rangle = \{\omega^3, \omega^6, \omega^9, \omega^{12}\} = \{i, -1, -i, 1\}, \\
\Omega_6 &= \langle \omega^{2} \rangle \\
&= \{\omega^2, \omega^4, \omega^6, \omega^8, \omega^{10}, \omega^{12}\} \\
&= \{(1+i\sqrt{3})/2, (-1+i\sqrt{3})/2, -1, (-1-i\sqrt{3})/2, (1-i\sqrt{3})/2, 1\},
\end{aligned}$$

and $\Omega_{12} = \langle \omega^1 \rangle$ itself. The sets of primitive roots are

$$\begin{aligned}
\Omega_1' &= \{\omega^{12}\}, \\
\Omega_2' &= \{\omega^{6}\}, \\
\Omega_3' &= \{\omega^4, \omega^{8}\}, \\
\Omega_4' &= \{\omega^3, \omega^{9}\}, \\
\Omega_6' &= \{\omega^2, \omega^{10}\}, \\
\Omega_{12}' &= \{\omega, \omega^5, \omega^7, \omega^{11}\},
\end{aligned}$$

which correspond to the cyclotomic polynomials

$$\begin{aligned}
\Phi_1(x) &= (x - \omega^{12}) = x - 1, \\
\Phi_2(x) &= (x - \omega^{6}) = x + 1, \\
\Phi_3(x) &= (x - \omega^{4})(x - \omega^{8}) = x^2 + x + 1, \\
\Phi_4(x) &= (x - \omega^{3})(x - \omega^{9}) = x^2 + 1,
\end{aligned}$$

$$\begin{aligned}\Phi_6(x) &= (x-\omega^2)(x-\omega^{10}) = x^2 - x + 1,\\ \Phi_{12}(x) &= (x-\omega)(x-\omega^5)(x-\omega^7)(x-\omega^{11}) = x^4 - x^2 + 1.\end{aligned}$$

Finally, we obtain the irreducible factorization of the polynomial $x^{12}-1$ over the integers:

$$\begin{aligned}x^{12}-1 &= (x-\omega^1)(x-\omega^2)\cdots(x-\omega^{12})\\ &= \big[(x-\omega^{12})\big]\big[(x-\omega^6)\big]\cdots\big[(x-\omega)(x-\omega^5)(x-\omega^7)(x-\omega^{11})\big]\\ &= \Phi_1(x)\Phi_2(x)\Phi_3(x)\Phi_4(x)\Phi_6(x)\Phi_{12}(x)\\ &= (x-1)(x+1)(x^2+x+1)(x^2+1)(x^2-x+1)(x^4-x^2+1).\end{aligned}$$

8 Field Extensions

8.1 Some Ring Theory

Ring theory is dangerous. Like group theory, the abstract theory of rings is extremely deep. With groups it was easy for me to omit various definitions without telling you because the whole concept of groups seems unfamiliar. Rings seem superficially familiar because they are based on “numbers” and “polynomials”. However, the abstract theory is quite wild and leads quickly away from intuition. My goal in this section is to say just enough, without veering into unnecessary abstraction.[104]

We are guided by the example of modular arithmetic. In the previous chapter we showed that the subgroups of $(\mathbb{Z}, +0)$ are precisely $n\mathbb{Z}$ for integers $n \geqslant 0$. We constructed the set of cosets

$$\mathbb{Z}/n\mathbb{Z} = \{a + n\mathbb{Z} : a \in \mathbb{Z}\},$$

and defined on this the following group operation:

$$(a + n\mathbb{Z}) + (b + n\mathbb{Z}) = (a + b) + n\mathbb{Z}.$$

We called $(\mathbb{Z}/n\mathbb{Z}, +, 0 + n\mathbb{Z})$ a quotient group. Furthermore, we had a quotient homomorphism

$$\begin{array}{cccc}\varphi : & \mathbb{Z} & \to & \mathbb{Z}/n\mathbb{Z}\\ & a & \mapsto & a + n\mathbb{Z},\end{array}$$

with kernel $n\mathbb{Z}$. But we saw in Chapter 4 that $\mathbb{Z}/n\mathbb{Z}$ is not just an additive group; it also has a multiplication operation, making it into a ring. This multiplication is defined on cosets as follows:

$$(a + n\mathbb{Z})(b + n\mathbb{Z}) = (ab) + n\mathbb{Z}.$$

If we can show that this is well-defined then the ring properties will follow immediately. So let me recall the proof that it is well-defined. Assume that $a + n\mathbb{Z} = a' + n\mathbb{Z}$ and $b + n\mathbb{Z} = b' + n\mathbb{Z}$

[104] This was also a challenge in Chapter 3 on Unique Prime Factorization. In some sense it would be more efficient to prove the unique factorization theorem in the context of Principal Ideal Domains. On the other hand, I believe that approach is too abstract for students learning the material for the first time.

so that $a - a' \in n\mathbb{Z}$ and $b - b' \in n\mathbb{Z}$. In this case we wish to show that $ab - a'b' \in n\mathbb{Z}$, so that $(ab) + n\mathbb{Z} = (a'b') + n\mathbb{Z}$ as sets. In Chapter 4 we did this by first naming integers $k, \ell \in \mathbb{Z}$ such that $a - a' = nk$ and $b - b' = \ell n\ell$, and then we expressed $ab - a'b'$ as n(something). Today I will avoid doing this because this method doesn't lend itself to generalization. Instead I will refer to the following abstract properties of the set $n\mathbb{Z}$:

- For all $c, d \in n\mathbb{Z}$ we have $c \pm d \in n\mathbb{Z}$.
- For all $c \in \mathbb{Z}$ and $d \in n\mathbb{Z}$ we have $cd \in n\mathbb{Z}$.

Then since $a - a'$ and $b - b'$ are in $n\mathbb{Z}$ we immediately have

$$ab - a'b' = ab - a'b + a'b - a'b' = (a - a')b + a'(b - b') \in n\mathbb{Z}.$$

This example inspires the following definition.

Ideals and Quotient Rings

Consider a ring $(R, +, \cdot, 0, 1)$ and a subset $I \subseteq R$. We say that I is an *ideal of* R when the following two properties are satisfied:

- For all $c, d \in I$ we have $c \pm d \in I$. Equivalently, $(I, +, 0)$ is a subgroup of $(R, +, 0)$.
- For all $c \in R$ and $d \in I$ we have $cd \in I$.

Since $(I, +, 0)$ is a subgroup of $(R, +, 0)$ we may construct the quotient group R/I with operation

$$(a + I) + (b + I) = (a + b) + I.$$

The second property of ideals guarantees that the following multiplication operation is also well-defined:

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

Then it is an easy and boring exercise to check that R/I is a ring with additive identity $0 + I$ and multiplicative identity $1 + I$.

Proof. Suppose that $a + I = a + I$ and $b + I = b' + I$, so that $a - a' \in I$ and $b - b' \in I$. Then since I is closed under multiplication by elements of R we have

$$ab - a'b' = ab - a'b + a'b - a'b' = (a - a')b + a'(b - b') \in I,$$

and hence $(ab) + I = (a'b') + I$. □

The general theory of ideals is quite elaborate. In this class we are only interested in the following special cases.

A Field is a Ring With Exactly Two Ideals

Every ring R has a *zero ideal* $\{0\} \subseteq R$ and a *unit ideal* $R \subseteq R$.[105] We call it the "unit ideal" because of the following fact: For any ideal $I \subseteq R$ we have

$$I = R \quad \Longleftrightarrow \quad I \text{ contains a unit.}$$

It follows from this that R is a field if and only if it has exactly two ideals.

Proof. First suppose that $I = R$. In this case I contains every unit of R. In particular, $1 \in I$. Conversely, suppose that I contains a unit, say $u \in I$. Since I is and ideal with $u \in I$ and $u^{-1} \in R$ this implies that $1 = uu^{-1} \in I$. Finally, for any $a \in R$ we have $a = 1a \in I$, which implies that $I = R$.

Now we will show that R is a field if and only if it has exactly two ideals. For one direction, suppose that R is a field. Then any nonzero ideal I contains a nonzero element of R. Since every nonzero element of a field is a unit, this implies that I contains a unit, hence $I = R$. For the other direction, suppose that R has exactly two ideals $\{0\}$ and R. For any element $a \in R$, the following set is an ideal:

$$aR = \{ab : b \in R\}.$$

Indeed, for any $ab, ac \in aR$ and $d \in R$ we have $ab \pm ac = a(b \pm c) \in aR$ and $(ab)d = a(bc) \in aR$. If $a \neq 0$ then we have $aR \neq \{0\}$ and hence $aR = R$, since R has only two ideals. It follows that $1 \in aR$ and hence $1 = ab$ for some $b \in R$. In other words, R is a field. □

Thus fields are the "simplest" rings from the point of view of ideal theory. The next simplest kind of rings are the the so-called "principal ideal domains". The most important class of these are the Euclidean domains, which we studied in Chapter 3.

Quotients of Euclidean Domains

Let (R, N) be a Euclidean domain. Then:

(i) Every ideal has the form $aR \subseteq R$ for some element $a \in R$. An ideal of the form $aR \subseteq R$ is called the *principal ideal generated by a* and any domain having only principal ideals is called a *principal ideal domain* (PID).

(ii) There is a bijection between ideals and association classes of elements:

$$aR = bR \quad \Longleftrightarrow a \sim b.$$

[105] Technically: We do allow the case where $\{0\} = R$. This is called the *zero ring*. However, it is an axiom of fields that $0 \neq 1$, so there is no such thing as the *zero field*.

(iii) R/pR is a field if and only if $p \in R$ is prime.[106]

Proof. (i): Let $I \subseteq R$ be an ideal. The zero ideal is principal: $\{0\} = 0R$. So let us assume that $I \neq \{0\}$ and let $a \in I$ be a nonzero element with smallest possible size $N(a)$. In this case I claim that $I = aR$. Indeed, since I is an ideal we have $ab \in I$ for any $b \in R$, hence $aR \subseteq I$. On the other hand we may divide any element $c \in I$ by a to obtain

$$\begin{cases} c = aq + r, \\ r = 0 \text{ or } N(r) < N(a). \end{cases}$$

Since $a, c \in I$ we note that $r = c - aq \in I$. If $r \neq 0$ then r is a nonzero element of I with size strictly smaller than a, which a contradiction. It follows that $r = 0$ and hence $c = aq \in aR$. Since this holds for any $c \in I$ we have shown that $I \subseteq aR$ as desired.

(ii): Next we show that $aR = bR$ if and only if $a \sim b$. If one of a or b is zero then so is the other. So let us assume that a and b are both nonzero. For one direction, suppose that $a \sim b$ so that $a = bu$ and $b = au^{-1}$ for some unit $u \in R$. For all $r \in R$ it follows that $ar = b(ur) \in bR$ and $br = a(u^{-1}r) \in aR$. Hence we have $aR \subseteq bR$ and $bR \subseteq aR$. For the other direction, suppose that $aR = bR$. Since $a \in aR$ this implies that $a \in bR$ and hence $a = bu$ for some $u \in R$. Similarly, since $b \in bR = aR$ we have $b = av$ for some $v \in R$. Then since R is a domain and $b \neq 0$ we find that u and v are units:

$$\begin{aligned} b &= av \\ b &= buv \\ b(1 - uv) &= 0 \\ 1 - uv &= 0. \end{aligned}$$

Hence $a \sim b$.

(iii): Let $p \in R$ be prime and consider the quotient ring R/pR. We will use the Extended Euclidean Algorithm from Chapter 3 to prove that R/pR is a field. This is the same proof that we used to show that $\mathbb{Z}/p\mathbb{Z}$ is a field. So consider any nonzero element of the quotient ring: $a + pR \neq 0 + pR$. By definition this means that $p \nmid a$. Since p is prime this implies that $\gcd(a, p) = 1$, hence from the Extended Euclidean Algorithm we can find $b, c \in R$ such that $ab + pc = 1$. Finally, we conclude that

$$(a + pR)(b + pR) = ab + pR = (1 - pc) + pR = 1 + pR,$$

so that $a + pR$ has a multiplicative inverse.

Conversely, suppose that $p \in R$ is not prime. That is, suppose that we have $p = ab$ for some $a, b \in R$ both non-associate to p. In particular, this implies that $p \nmid a$ since $p|a$ and $a|p$ would

[106] If R itself is a field then we can also allow $p = 0$, since $R/0R \cong R$.

imply $p \sim a$. Similarly, we have $p \nmid b$. But then we have two nonzero cosets $a + pR \neq 0 + pR$ and $b + pR \neq 0 + pR$ whose product in R/pR zero:

$$(a + pR)(b + pR) = ab + pR = p + pR = 0 + pR.$$

Hence R/pR is not a domain. □

Note that a field is trivially a PID since the zero and unit ideals are principal:

$$\{0\} = 0R \quad \text{and} \quad R = 1R.$$

More interesting examples come from our favorite Euclidean domains $\mathbb{Z}$ and $\mathbb{F}[x]$:

- Every ideal of $\mathbb{Z}$ has the form $n\mathbb{Z}$ for some $n \in \mathbb{Z}$. Recall that $\mathbb{Z}^\times = \{\pm 1\}$, hence we have

  $$m\mathbb{Z} = n\mathbb{Z} \quad \Longleftrightarrow m = \pm n.$$

  It follows that each ideal can be expressed uniquely in the form $n\mathbb{Z}$ for some $n \geqslant 0$.

- Every ideal of $\mathbb{F}[x]$ has the form $f(x)\mathbb{F}[x]$ for some $f(x) \in \mathbb{F}[x]$. Recall that the units of $\mathbb{F}[x]$ are the non-zero constants $\mathbb{F}^\times$. Thus we have

  $$f(x)\mathbb{F}[x] = g(x)\mathbb{F}[x] \quad \Longleftrightarrow \quad f(x) = \lambda g(x) \text{ for some } \lambda \in \mathbb{F}^\times.$$

  It follows that every non-zero ideal of $\mathbb{F}[x]$ can be expressed uniquely in the form $m(x)\mathbb{F}[x]$ for some **monic** polynomial $m(x) \in \mathbb{F}[x]$ (i.e., with leading coefficient 1).

Just for context, let me briefly mention the two simplest examples[107] of non-PIDs:

$$\mathbb{Z}[x] \quad \text{and} \quad \mathbb{F}[x, y].$$

Indeed, one can check that the following sets are non-principal ideals of $\mathbb{Z}[x]$ and $\mathbb{F}[x, y]$:

$$\begin{aligned} 2\mathbb{Z}[x] + x\mathbb{Z}[x] &= \{2f(x) + xg(x) : f(x), g(x) \in \mathbb{Z}[x]\}, \\ x\mathbb{F}[x, y] + y\mathbb{F}[x, y] &= \{xf(x, y) + yg(x, y) : f(x, y), g(x, y) \in \mathbb{F}[x, y]\}. \end{aligned}$$

It is much harder to classify the ideals of these rings, so we won't even try.

As with groups, the modern study of rings is expressed in terms of homomorphism and isomorphism. These concepts are packaged together in the First Isomorphism Theorem for Rings. Most of this follows from the First Isomorphism Theorem for (Abelian) Groups. We just need to include the multiplicative structure.

[107] The original example of a ring that is not a PID is the ring $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5}\}$. The lack of unique prime factorization in rings such as these frustrated early attempts to prove Fermat's Last Theorem.

Definition of Ring Homomorphism

Consider rings $(R, +, \cdot, 0, 1)$ and $(R', +, \cdot, 0', 1')$ and a function $\varphi : R \to R'$. We say that φ is a *ring homomorphism* when

(i) $\varphi(a + b) = \varphi(a) + \varphi(b)$,

(ii) $\varphi(ab) = \varphi(a)\varphi(b)$,

(iii) $\varphi(1) = 1'$.

The first axiom says that $\varphi : (R, +, 0) \to (R', +, 0')$ is a group homomorphism, which implies that $\varphi(0) = 0'$ as follows:

$$\begin{aligned} 0 + 0 &= 0 \\ \varphi(0 + 0) &= \varphi(0) \\ \varphi(0) + \varphi(0) &= \varphi(0) \\ \varphi(0) + \varphi(0) - \varphi(0) &= \varphi(0) - \varphi(0) \\ \varphi(0) &= 0'. \end{aligned}$$

However, the second axiom $\varphi(ab) = \varphi(a)\varphi(b)$ does not imply that $\varphi(1) = 1'$, because we are not necessarily allowed to divide in a ring. Indeed, if we try to use the same proof idea then we get stuck:

$$\begin{aligned} 1 \cdot 1 &= 1 \\ \varphi(1 \cdot 1) &= \varphi(1) \\ \varphi(1)\varphi(1) &= \varphi(1). \end{aligned}$$

Now we cannot conclude that $\varphi(1) = 1'$ because we are not allowed to "divide both sides by $\varphi(1)$". Hence we must include $\varphi(1) = 1'$ as an axiom.

We should think of a ring $(R, +, \cdot, 0, 1)$ as an abelian group $(R, +, 0)$ with some extra decorations. Thus the kernel of a ring homomorphism $\varphi : R \to R'$ is defined as the set of $a \in R$ such that $\varphi(a) = 0'$. The First Isomorphism Theorem confirms that this is the correct definition.

The First Isomorphism Theorem for Rings

Consider a ring homomorphism $\varphi : (R, +, \cdot, 0, 1) \to (R', +, \cdot, 0', 1')$. We define the image and kernel as follows:

$$\operatorname{im} \varphi = \{a' \in R' : \exists a \in R, \varphi(a) = a'\},$$

$$\ker\varphi = \{a \in R : \varphi(a) = 0'\}.$$

I claim that $\operatorname{im}\varphi \subseteq R'$ is a subring and $\ker\varphi \subseteq R$ is an ideal.[108] Furthermore, I claim that the following map is a well-defined ring isomorphism:

$$\begin{array}{rccc} \tilde{\varphi}: & R/\ker\varphi & \to & \operatorname{im}\varphi \\ & a+\ker\varphi & \mapsto & \varphi(a). \end{array}$$

Proof. A *subring* $S \subseteq R$ is a subset that is closed under the ring operations $+, \cdot$ and contains the special elements $0, 1$. Consider any two elements of the image: $a' = \varphi(a)$ and $b' = \varphi(b)$. Since $a' + b' = \varphi(a) + \varphi(b) = \varphi(a+b)$ and $a'b' = \varphi(a)\varphi(b) = \varphi(ab)$, we see that $a' + b'$ and $a'b'$ are also in the image. And since $\varphi(0) = 0'$ and $\varphi(1) = 1'$, we see that $0'$ and $1'$ are in the image. Hence $\operatorname{im}\varphi \subseteq R'$ is a subring.

Next we show that $\ker\varphi \subseteq R$ is an ideal. Since $\ker\varphi$ was defined in terms of the additive structure $(R, +, 0)$ we already know from the previous chapter that $\ker\varphi$ is an additive subgroup.[109] Thus we only need to check the second axioms for ideals. Suppose that $a \in R$ and $b \in \ker\varphi$, so that $\varphi(b) = 0'$. Then we have $\varphi(ab) = \varphi(a)\varphi(b) = \varphi(a)\cdot 0' = 0'$, so that $ab \in \ker\varphi$.

Finally, we check that $\tilde{\varphi}(a+\ker\varphi) = \varphi(a)$ is a well-defined ring isomorphism. We already know from the previous chapter that this is a well-defined isomorphism of additive groups. Hence we only need to check that $\tilde{\varphi}$ is a ring homomorphism. Indeed, it preserves multiplication because

$$\begin{aligned} \tilde{\varphi}((a+\ker\varphi)(b+\ker\varphi)) &= \tilde{\varphi}(ab+\ker\varphi) \\ &= \varphi(ab) \\ &= \varphi(a)\varphi(b) \\ &= \tilde{\varphi}(a+\ker\varphi)\tilde{\varphi}(b+\ker\varphi). \end{aligned}$$

And it preserves the unit element because

$$\tilde{\varphi}(1+\ker\varphi) = \varphi(1) = 1'.$$

□

8.2 The Minimal Polynomial Theorem

In Chapter 4 we developed the theory of "modular arithmetic" in the ring $\mathbb{Z}$. Now we pursue the analogous theory in the ring of polynomials $\mathbb{F}[x]$ over a field $\mathbb{F}$. Even though the two theories are analogous, they are still very different. The ideal theory of $\mathbb{F}[x]$ is encoded via "evaluation homomorphisms".

[108] The image is almost never an ideal and the kernel is almost never a subring.

[109] It's easy enough to check it again. Given $a, b \in \ker\varphi$ we have $\varphi(a-b) = \varphi(a) - \varphi(b) = 0' - 0' = 0'$ and hence $a - b \in \ker\varphi$.

Evaluation Homomorphisms

Let $\mathbb{E} \subseteq \mathbb{F}$ be a field extension. Then for any element $\alpha \in \mathbb{E}$ we have a ring homomorphism $\mathbb{F}[x] \to \mathbb{E}$ defined by evaluating polynomials at $x = \alpha$:

$$\begin{array}{rccc} \varphi_{\alpha/\mathbb{F}}: & \mathbb{F}[x] & \to & \mathbb{E} \\ & f(x) & \mapsto & f(\alpha). \end{array}$$

We can view this as the unique ring homomorphism $\mathbb{F}[x] \to \mathbb{E}$ that fixes elements of $\mathbb{F}$ and sends x to α. We include $\mathbb{F}$ in the notation "$\varphi_{\alpha/\mathbb{F}}$" to indicate that this homomorphism fixes elements of $\mathbb{F}$. The symbol "/" here is not mathematical; it is an abbreviation for the English word "over". The symbol "$\alpha/\mathbb{F}$" indicates that we are thinking of α as an element of a field extension of $\mathbb{F}$.

The unique ring homomorphism $\varphi_{\alpha/\mathbb{F}} : \mathbb{F}[x] \to \mathbb{E}$ is analogous to the unique group homomorphism $\varphi_a : (\mathbb{Z}, +, 0) \to G$ that sends an integer $k \in \mathbb{Z}$ to the power $a^k \in G$. Just as the image $\langle a \rangle = \operatorname{im} \varphi_a \subseteq G$ is the *subgroup generated by a* (i.e., the smallest subgroup of G that contains a), the image of $\varphi_{\alpha/\mathbb{F}}$ is the "subring of $\mathbb{E}$ generated by α over $\mathbb{F}$".

Adjoining an Element to a Field

Consider an element of a field extension $\alpha \in \mathbb{E} \subseteq \mathbb{F}$ with corresponding evaluation homomorphism $\varphi_{\alpha/\mathbb{F}} : \mathbb{F}[x] \to \mathbb{E}$. We denote the image by

$$\mathbb{F}[\alpha] := \operatorname{im} \varphi_{\alpha/\mathbb{F}} = \{f(\alpha) : f(x) \in \mathbb{F}[x]\}.$$

Being the image of a ring homomorphism, $\mathbb{F}[\alpha]$ is necessarily a subring of $\mathbb{E}$. I claim that it is the **smallest subring of $\mathbb{E}$ that contains α and $\mathbb{F}$**. Based on this idea, we refer to the ring[110] $\mathbb{F}[\alpha]$ as "$\mathbb{F}$ *adjoin* α".

Proof. Let $R \subseteq \mathbb{E}$ be any subring containing $\mathbb{F}$ and α, and consider an arbitrary polynomial

$$f(x) = a_0 + a_1 + \cdots + a_n x^n \in \mathbb{F}[x].$$

Since $a_0, \ldots, a_n, \alpha \in R$ and since R is closed under addition and multiplication, we see that

$$f(\alpha) = a_0 + a_1\alpha + \cdots + a_n\alpha^n \in R.$$

[110] If α is a root of some polynomial over $\mathbb{F}$ then we will prove below in the Minimal Polynomial Theorem that $\mathbb{F}[\alpha]$ is actually a field. This is surprising.

Since this holds for any $f(x) \in \mathbb{F}[x]$ we conclude that $\mathbb{F}[\alpha] \subseteq R$ as desired. □

Now we discuss the kernel of $\varphi_{\alpha/\mathbb{F}}$. There are two essentially different cases.

Algebraic vs Transcendental Elements over a Field

Consider an element $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ with evaluation homomorphism $\varphi_{\alpha/\mathbb{F}} : \mathbb{F}[x] \to \mathbb{E}$. Since $\mathbb{F}[x]$ is a PID we know that the kernel of $\varphi_{\alpha/\mathbb{F}}$ is a principal ideal. There are two essentially different cases:

- If $\ker \varphi_{\alpha/\mathbb{F}} = \{0\}$ then we say that α *is transcendental over* $\mathbb{F}$. For example, Lindemann proved in 1882 that $\pi = 3.14\cdots$ is transcendental over $\mathbb{Q}$. It is generally quite difficult to prove that a given complex number is transcendental over $\mathbb{Q}$.[111]
- If $\ker \varphi_{\alpha/\mathbb{F}} \neq \{0\}$ then since $\mathbb{F}[x]$ is a PID there exists a unique monic polynomial $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ such that

  $$\ker \varphi_{\alpha/\mathbb{F}} = m_{\alpha/\mathbb{F}}(x)\mathbb{F}[x] = \{m_{\alpha/\mathbb{F}}(x)g(x) : g(x) \in \mathbb{F}[x]\}.$$

  Equivalently, for all $f(x) \in \mathbb{F}[x]$ we have

  $$f(\alpha) = 0 \quad \Longleftrightarrow \quad m_{\alpha/\mathbb{F}}(x) \,\big|\, f(x) \text{ in the ring } \mathbb{F}[x].$$

  In this case we say that α *is algebraic over* $\mathbb{F}$ and we call $m_{\alpha/\mathbb{F}}(x)$ the *minimal polynomial for* α *over* $\mathbb{F}$.

The concept of a minimal polynomial is a direct generalization of Descartes' Factor Theorem. Indeed, for any element $\alpha \in \mathbb{F}$ and for any polynomial $f(x) \in \mathbb{F}[x]$, Descartes says that

$$f(\alpha) = 0 \quad \Longleftrightarrow \quad (x - \alpha) \,\big|\, f(x) \text{ in the ring } \mathbb{F}[x].$$

In other words, if $\alpha \in \mathbb{F}$ then the minimal polynomial of α over $\mathbb{F}$ is $m_{\alpha/\mathbb{F}}(x) = x - \alpha$.

We also saw a slightly more general example last semester. For any real polynomial $f(x) \in \mathbb{R}[x]$ and for a fixed square root $i = \sqrt{-1}$, we showed that

$$f(i) = 0 \quad \Longleftrightarrow \quad (x^2 + 1) \,\big|\, f(x) \text{ in the ring } \mathbb{R}[x],$$

[111] One can show that the algebraic numbers over $\mathbb{Q}$ are countable, but the complex numbers are uncountable. The most famous transcendental number that cannot be proved to be so is the *Euler-Mascheroni constant*:

$$\gamma = \lim_{n\to\infty}\left(-\log n + \sum_{k=1}^{n} \frac{1}{k}\right) \approx 0.577.$$

In fact, no one even knows how to prove that γ is irrational.

so that $m_{i/\mathbb{R}}(x) = x^2 + 1$ is the minimal polynomial of i over $\mathbb{R}$. On the other hand, the minimal polynomial of i over $\mathbb{C}$ is $m_{i/\mathbb{C}}(x) = x - i$. This is why we include the field in the notation for minimal polynomials.

Since the evaluation homomorphism $\varphi_{i/\mathbb{R}} : \mathbb{R}[x] \to \mathbb{C}$ is surjective, it follows from the First Isomorphism Theorem that

$$\frac{\mathbb{R}[x]}{(x^2+1)\mathbb{R}[x]} = \frac{\mathbb{R}[x]}{\ker \varphi_{i/\mathbb{R}}} \cong \operatorname{im} \varphi_{i/\mathbb{R}} = \mathbb{R}[i] = \mathbb{C}.$$

This is Cauchy's construction of the complex numbers, which we discussed in section 6.8. Recall that we did a lot of work in Chapter 1 to construct the complex numbers and to prove their basic properties. The following theorem is a generalization of this construction. We will apply it in the next two sections when we construct finite fields.

The Minimal Polynomial Theorem

Consider an element $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ with evaluation homomorphism $\varphi_{\alpha/\mathbb{F}} : \mathbb{F}[x] \to \mathbb{E}$. Let α be algebraic over $\mathbb{F}$ with minimal polynomial $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$. Then we have the following:

(1) The minimal polynomial $m_{\alpha/\mathbb{F}}(x)$ is **irreducible** over $\mathbb{F}$. Furthermore, if $f(\alpha) = 0$ for some irreducible monic polynomial $f(x) \in \mathbb{F}[x]$ then $f(x) = m_{\alpha/\mathbb{F}}(x)$.

(2) The subring $\mathbb{F}[\alpha] \subseteq \mathbb{E}$ is actually a **field**.

(3) If $d = \deg(m_{\alpha/\mathbb{F}})$ then every element $\beta \in \mathbb{F}[\alpha]$ has a **unique** expression of the form

$$\beta = b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_{d-1}\alpha^{d-1},$$

for some elements $b_0, b_1, \ldots, b_{d-1}$.

Proof. (1): Let $m_{\alpha/\mathbb{F}}(x) = g(x)h(x)$ for some $g(x), h(x) \in \mathbb{F}[x]$. Substituting α gives

$$0 = m_{\alpha/\mathbb{F}}(\alpha) = g(\alpha)h(\alpha),$$

which implies that $g(\alpha) = 0$ or $h(\alpha) = 0$ since we are working in a domain. Without loss, suppose that $g(\alpha) = 0$. By definition of $m_{\alpha/\mathbb{F}}(x)$ this means that $m_{\alpha/\mathbb{F}}(x)|g(x)$. On the other hand we have $g(x)|m_{\alpha/\mathbb{F}}(x)$ by assumption. Since we are working in a domain this implies that $m_{\alpha/\mathbb{F}}(x) \sim g(x)$. Hence $m_{\alpha/\mathbb{F}}(x)$ is irreducible over $\mathbb{F}$.

Now let $f(x) \in \mathbb{F}[x]$ be monic and irreducible over $\mathbb{F}$, with $f(\alpha) = 0$. By definition of the minimal polynomial we have $m_{\alpha/\mathbb{F}}(x)|f(x)$. Then since $f(x)$ is irreducible we have $m_{\alpha/\mathbb{F}}(x) = \lambda f(x)$ for some $\lambda \in \mathbb{F}$. Finally, since $m_{\alpha/\mathbb{F}}(x)$ and $f(x)$ are both monic we have $\lambda = 1$.

(2): The The First Isomorphism Theorem for Rings tells us that

$$\mathbb{F}[\alpha] = \operatorname{im}\varphi_{\alpha/\mathbb{F}} \cong \frac{\mathbb{F}[x]}{\ker\varphi_{\alpha/\mathbb{F}}} = \frac{\mathbb{F}[x]}{m_{\alpha/\mathbb{F}}(x)\mathbb{F}[x]}.$$

Then since $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ is a prime element of a Euclidean domain, it follows from the previous section that this quotient ring is a field.

(3): Let $d = \deg(m_{\alpha/\mathbb{F}})$ and consider an arbitrary element $\beta \in \mathbb{F}[\alpha]$. By definition, we can write $\beta = f(\alpha)$ for some polynomial $f(x) \in \mathbb{F}[x]$. Divide this $f(x)$ by the minimal polynomial $m_{\alpha/\mathbb{F}}(x)$ to obtain polynomials $q(x), r(x) \in \mathbb{F}[x]$ satisfying

$$\begin{cases} f(x) = m_{\alpha/\mathbb{F}}(x)q(x) + r(x), \\ r(x) = 0 \text{ or } \deg(r) < \deg(m_{\alpha/\mathbb{F}}). \end{cases}$$

Since $r(x) = 0$ or $\deg(r) < \deg(m_{\alpha/\mathbb{F}}) = d$, we can write

$$r(x) = b_0 + b_1 x + \cdots b_{d-1}x^{d-1},$$

for some elements $b_0, \ldots, b_{d-1} \in \mathbb{F}$. Then substitute $x = \alpha$ to obtain

$$\begin{aligned} \beta &= f(\alpha) \\ &= m_{\alpha/\mathbb{F}}(\alpha)q(\alpha) + r(\alpha) \\ &= 0 \cdot q(\alpha) + r(\alpha) \\ &= r(\alpha) \\ &= b_0 + b_1\alpha + \cdots + b_{d-1}\alpha^{d-1}. \end{aligned}$$

To prove uniqueness of this expression, suppose that we have

$$b_0 + b_1\alpha + \cdots + b_{d-1}\alpha^{d-1} = c_0 + c_1\alpha + \cdots + c_{d-1}\alpha^{d-1}$$

for some $b_0, \ldots, b_{d-1}, c_0, \ldots, c_{d-1} \in \mathbb{F}$. We wish to show that $b_i = c_i$ for all i. To do this, we define polynomials $r(x) = b_0 + b_1x + b_{d-1}x^{d-1}$ and $s(x) = c_0 + c_1x + \cdots + c_{d-1}x^{d-1}$. We will be done if we can show that $r(x) - s(x)$ is the zero polynomial, since then the coefficients of $r(x)$ and $s(x)$ will be equal.

By assumption we have $r(\alpha) = s(\alpha)$ and hence $r(\alpha) - s(\alpha) = 0$. In other words, we have $r(x) - s(x) \in \ker\varphi_{\alpha/\mathbb{F}}$, which implies that $r(x) - s(x)$ is divisible by $m_{\alpha/\mathbb{F}}(x)$. If $r(x) - s(x) \neq 0$ then this gives a contradiction:

$$d = \deg(m_{\alpha/\mathbb{F}}) \leqslant \deg(r - s) \leqslant \max\{\deg(r), \deg(s)\} < d.$$

Hence $r(x) - s(x) = 0$ as desired. □

To end this section we will discuss a few examples. Each example will require us to prove that a certain polynomial is irreducible over $\mathbb{Q}$, and each time we will use the following trick.

The Rational Root Test

For any polynomial $f(x) \in \mathbb{Z}[x]$[112] there is a finite algorithm to determine all of the rational or roots of $f(x)$, or to prove that no such roots exist. Suppose that

$$f(x) = c_0 + c_1 x + \cdots + c_n x^n \in \mathbb{Z}[x],$$

with $c_n \neq 0$. If $f(a/b) = 0$ for some integers $a, b \in \mathbb{Z}$ with $\gcd(a, b) = 1$ then we must have $a|c_0$ and $b|c_n$. This gives a finite list of possible roots $a/b \in \mathbb{Q}$.

Proof. Suppose that $f(\alpha) = 0$ for some $\alpha \in \mathbb{Q}$ and write $\alpha = a/b$ with $a, b \in \mathbb{Z}$ and $\gcd(a, b) = 1$. Substitute a/b into the expression for $f(x)$ and multiply both sides by n:

$$\begin{aligned} f(a/b) &= 0 \\ c_0 + c_1(a/b) + \cdots + c_n(a/b)^n &= 0 \\ c_0 b^n + c_1 a b^{n-1} + \cdots + c_n a^n &= 0. \end{aligned}$$

We find that b divides $c_n a^n$ because

$$\begin{aligned} c_0 b^n + c_1 a b^{n-1} + \cdots + c_{n-1} a^{n-1} b &= -c_n a^n \\ b(c_0 b^{n-1} + c_1 a b^{n-2} + \cdots + c_{n-1} a^{n-1}) &= -c_n a^n. \end{aligned}$$

Then since $b|c_n a^n$ and $\gcd(a, b) = 1$ we must have $b|c_n$.[113] A similar proof shows that $a|c_0$. □

Furthermore, we recall the following result from Chapter 3.

Irreducible Polynomials of Small Degree

Let $f(x) \in \mathbb{F}[x]$ have degree 2 or 3. Then

$$f(x) \text{ is irreducible over } \mathbb{F} \iff f(x) \text{ has no root in } \mathbb{F}.$$

Examples.

[112] We can also allow $f(x) \in \mathbb{Q}[x]$ since the roots of $f(x)$ are the same as the roots of $m \cdot f(x) \in \mathbb{Z}[x]$ where $m \in \mathbb{Z}$ is least common multiple of the denominators of the coefficients of $f(x)$.

[113] Recall: If $b|ca$ and $\gcd(a, b) = 1$ then we can write $1 = ax + by$ for some $x, y \in \mathbb{Z}$ and then multiply both sides by c to get $c = cax + cby = b(\text{something})$, hence $b|c$.

• **Square Roots of Integers.** Consider any integer $d \in \mathbb{Z}$ and a fixed square root $\sqrt{d} \in \mathbb{C}$. Suppose that $\sqrt{d} \notin \mathbb{Z}$, so that $\sqrt{d} \notin \mathbb{Q}$. In other words, assume that the polynomial $x^2 - d \in \mathbb{Q}[x]$ has no roots in $\mathbb{Q}$. From the above result this implies that $x^2 - d$ is irreducible over $\mathbb{Q}$, hence it is the minimal polynomial for $\sqrt{d}$ over $\mathbb{Q}$:

$$m_{\sqrt{d}/\mathbb{Q}}(x) = x^2 - d \in \mathbb{Q}[x].$$

Thus from the Minimal Polynomial Theorem we obtain a field by adjoining $\sqrt{d}$ to $\mathbb{Q}$:

$$\mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} : \text{for unique } a, b \in \mathbb{Q}\}.$$

In this case, division is achieved by "rationalizing the denominator":

$$\frac{1}{a+b\sqrt{d}} = \frac{1}{a+b\sqrt{d}} \cdot \frac{a-b\sqrt{d}}{a-b\sqrt{d}} = \left(\frac{a}{a^2-b^2d}\right) + \left(\frac{-b}{a^2-b^2d}\right)\sqrt{d}.$$

• **Cube Roots of 1.** Let $\omega \in \mathbb{C}$ be any primitive 3rd root of unity, for example $\omega = e^{2\pi i/3}$. Recall[114] that ω is a root of the cyclotomic polynomial

$$\Phi_3(x) = x^2 + x + 1.$$

I claim that $\Phi_3(x)$ is the minimal polynomial of ω over $\mathbb{Q}$. Since the degree is 2 we only need to show that $\Phi_3(x)$ has no root in $\mathbb{Q}$ and for this we use the Rational Root Test. Suppose that $\Phi_3(a/b) = 0$ for some $a, b \in \mathbb{Z}$ with $\gcd(a, b) = 1$. Then we must have $a|1$ and $b|1$, hence $a/b = \pm 1$. But we see that $+1$ and -1 are not roots of $x^2 + x + 1$. We conclude that $\Phi_3(x)$ is irreducible[115], and hence

$$m_{\omega/\mathbb{Q}}(x) = \Phi_3(x) = x^2 + x + 1.$$

Thus from the MPT we obtain the following field by adjoining ω to $\mathbb{Q}$:

$$\mathbb{Q}[\omega] = \{a + b\omega : \text{for unique } a, b \in \mathbb{Q}\}.$$

This time it is not so clear how to perform division, since we don't know how to define "conjugation". Instead we pursue a brute force approach. Suppose that elements $a + b\omega$ and $c + d\omega$ satisfy

$$(a + b\omega)(c + d\omega) = 1 + 0\omega.$$

We assume that $a, b \in \mathbb{Q}$ are known and we try to solve for $c, d \in \mathbb{Q}$. Expand the left hand side and use the identity $\omega^2 + \omega + 1 = 0$ to obtain

$$\begin{aligned}(a + b\omega)(c + d\omega) &= ac + (ad + bc)\omega + bd\omega^2 \\ &= ac + (ad + bc)\omega + bd(-1 - \omega) \\ &= (ac - bd) + (ad + bc - bd)\omega.\end{aligned}$$

[114] Since $0 = \omega^3 - 1 = (\omega - 1)(\omega^2 + \omega + 1)$ and $\omega - 1 \neq 0$ we must have $\omega^2 + \omega + 1 = 0$.

[115] More generally, it is true that any cyclotomic polynomial $\Phi_n(x)$ is irreducible over $\mathbb{Q}$, and hence is the minimal polynomial of any primitive nth root of unity over $\mathbb{Q}$.

Since the coefficients are unique, comparing with the right hand side gives the following system of two linear equations in the two unknowns $c, d \in \mathbb{Q}$:

$$\begin{cases} ac & - & bd & = & 1, \\ bc & + & (a-b)d & = & 0. \end{cases}$$

This can be solved by inverting the coefficient matrix:

$$\begin{aligned} \begin{pmatrix} a & -b \\ b & a-b \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} c \\ d \end{pmatrix} &= \begin{pmatrix} a & -b \\ b & a-b \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} c \\ d \end{pmatrix} &= \frac{1}{a(a-b)+b^2} \begin{pmatrix} a-b & b \\ -b & a \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} c \\ d \end{pmatrix} &= \frac{1}{a^2+b^2-ab} \begin{pmatrix} a-b \\ -b \end{pmatrix}. \end{aligned}$$

We conclude that

$$\frac{1}{a+b\omega} = \left(\frac{a-b}{a^2+b^2-ab} \right) + \left(\frac{-b}{a^2+b^2-ab} \right) \omega.$$

In retrospect, we see that the map $a + b\omega \mapsto (a-b) - b\omega$ plays the role of "conjugation" in the field $\mathbb{Q}[\omega]$. In the next example it will not be so easy to find a "conjugation" map.

• **Cube Roots of** 2. Let $\omega \in \mathbb{C}$ be any fixed cube root of 2, so that $\omega^3 = 2$. One can see using the Rational Root Test that any root $a/b \in \mathbb{Q}$ (in lowest terms) of the polynomial $x^3 - 2$ must satisfy $a|2$ and $b|1$, hence $a/b = \pm 1$ or ± 2. But none of these is a root of $x^3 - 2$. Thus $x^3 - 2$ is irreducible over $\mathbb{Q}$ and must be the minimal polynomial of α over $\mathbb{Q}$:

$$m_{\alpha/\mathbb{Q}}(x) = x^3 - 2.$$

It follows that we obtain a field by adjoining α to $\mathbb{Q}$:

$$\mathbb{Q}[\alpha] = \{a + b\alpha + c\alpha^2 : \text{for unique } a, b, c \in \mathbb{Q}\}.$$

To practice computations in this field, let's compute the inverse of $1 + \alpha + \alpha^2$. From the Minimal Polynomial Theorem we know that there exist unique $a, b, c \in \mathbb{Q}$ satisfying

$$(1 + \alpha + \alpha^2)(a + b\alpha + c\alpha^2) = 1 + 0\alpha + 0\alpha^2.$$

In order to solve for a, b, c, we expand the left hand side and use the fact that $\alpha^3 = 2$:

$$\begin{aligned} (1 + \alpha + \alpha^2)(a + b\alpha + c\alpha^2) = a &+ b\alpha + c\alpha^2 \\ &+ a\alpha + b\alpha^2 + c\alpha^3 \\ &+ a\alpha^2 + b\alpha^3 + c\alpha^4 \end{aligned}$$

$$\begin{aligned} &= a + b\alpha + c\alpha^2 \\ &\qquad + a\alpha + b\alpha^2 + 2c \\ &\qquad + a\alpha^2 + 2b + 2c\alpha \\ &= (a + 2b + 2c) + (a + b + 2c)\alpha + (a + b + c)\alpha^2. \end{aligned}$$

Since the coefficients are unique, comparing coefficients with the right hand side gives a system of three linear equations in the three unknowns $a, b, c \in \mathbb{Q}$:

$$\begin{cases} a + 2b + 2c = 1, \\ a + b + 2c = 0, \\ a + b + c = 0. \end{cases}$$

After a bit of work, one sees that $(a, b, c) = (-1, 1, 0)$, so that

$$\frac{1}{1 + \alpha + \alpha^2} = -1 + 1\alpha + 0\alpha^2.$$

More generally, the equation $(r + s\alpha + t\alpha^2)(a + b\alpha + c\alpha^2) = 1 + 0\alpha + 0\alpha^2$ leads the following system of linear equations in the unknowns a, b, c:

$$\begin{cases} ra + 2tb + 2sc = 1, \\ sa + rb + 2tc = 0, \\ ta + sb + rc = 0. \end{cases}$$

My computer says that the solution is

$$(a, b, c) = \left(\frac{r^2 - 2st}{\Delta}, \frac{2t^2 - rs}{\Delta}, \frac{s^2 - rt}{\Delta} \right),$$

where $\Delta = r^3 + 2s^3 + 4t^3 - 6rst$ is the determinant of the coefficient matrix. Clearly it is not worthwhile to do these calculations by hand.

Remark on "rationalizing the denominator": This time each element of the field $\mathbb{Q}[\alpha]$ will have **two conjugates**, obtained by replacing α with one of the other two roots of $x^3 - 2$; namely, $\omega\alpha$ or $\omega^2\alpha$, where ω is a primitive 3rd root of unity. Denote these "conjugation maps",[116] together with the identity map, by

$$\begin{aligned} \sigma_1(r + s\alpha + t\alpha^2) &:= r + s(\omega\alpha) + t(\omega\alpha)^2, \\ \sigma_2(r + s\alpha + t\alpha^2) &:= r + s(\omega^2\alpha) + t(\omega^2\alpha)^2. \end{aligned}$$

For a given element $\beta = r + s\alpha + t\alpha^2 \in \mathbb{F}[\alpha]$, one can check that

$$\begin{aligned} \sigma_1(\beta)\sigma_2(\beta) &= (r^2 - 2st) + (2t^2 - rs)\alpha + (s^2 - rt)\alpha^2, \\ \beta\sigma_1(\beta)\sigma_2(\beta) &= (r^3 + 2s^3 + 4t^3 - 6rst) + 0\alpha + 0\alpha^2. \end{aligned}$$

[116] For a given $\beta \in \mathbb{F}[\alpha]$ the complex numbers $\sigma_1(\beta)$ and $\sigma_2(\beta)$ do not live in $\mathbb{F}[\alpha]$, but this doesn't matter.

So the rationalization of the denominator of $1/\beta$ is achieved by multiplying the numerator and the denominator by **both of the conjugates** $\sigma_1(\beta)$ and $\sigma_2(\beta)$:

$$\frac{1}{\beta} = \frac{1}{\beta} \cdot \frac{\sigma_1(\beta)\sigma_2(\beta)}{\sigma_1(\beta)\sigma_2(\beta)} = \frac{1}{r^3 + 2s^3 + 4t^3 - 6rst} \left((r^2 - 2st) + (2t^2 - rs)\alpha + (s^2 - rt)\alpha^2\right).$$

Further investigations of this kind lead to the subject of *Galois theory*, which we will study in the next chapter.[117]

8.3 The Classification of Finite Fields

Given a field extension $\mathbb{E} \supseteq \mathbb{F}$ and an element $\alpha \in \mathbb{E}$ that is algebraic over $\mathbb{F}$, the Minimal Polynomial Theorem tells us that the subring $\mathbb{F}[\alpha] = \{f(\alpha) : f(x) \in \mathbb{F}[x]\} \subseteq \mathbb{E}$ generated by α is actually a field. To be precise, we have

$$\mathbb{F}[\alpha] \cong \frac{\mathbb{F}[x]}{m_{\alpha/\mathbb{F}}(x)\mathbb{F}[x]},$$

where $m_{\alpha/\mathbb{F}}(x)$ is the unique monic irreducible polynomial in $\mathbb{F}[x]$ having α as a root. This quotient ring is a field since R/pR is a field for any prime element $p \in R$ of a Euclidean domain.

In the next two sections, we will turn this construction around. That is, instead of starting with an element of a field extension $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ and ending with an irreducible polynomial $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$, we will start with an irreducible polynomial $m(x) \in \mathbb{F}[x]$ and end up with a field extension $\mathbb{E} \supseteq \mathbb{F}$ containing some element $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ such that $m(\alpha) = 0$.

Why would we do this? There are two reasons:

- We know that any non-constant polynomial $f(x) \in \mathbb{Q}[x]$ has a root (in fact, all of its roots) in the field $\mathbb{C}$ of complex numbers. This is the content of the Fundamental Theorem of Algebra. However, our proof of the FTA was based on the assumption that the roots already exist in some field extension $\mathbb{E} \supseteq \mathbb{C}$, and proceeded to show that the roots must actually be in $\mathbb{C}$. We did not yet verify this assumption, which is called Kronecker's Theorem.
- The same idea should work for polynomials over the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. That is, given some non-constant polynomial $f(x) \in \mathbb{F}_p[x]$, Kronecker's Theorem will tell us that there exists some field $\mathbb{E} \supseteq \mathbb{F}_p$ where $f(x)$ has all of its roots. The construction of $\mathbb{E}$ is analogous to the construction of the complex numbers, since it involves the adjunction of some "imaginary elements". But it does not directly involve the complex numbers because $\mathbb{C}$ does not contain a subfield isomorphic to $\mathbb{F}_p$.

Our main application of these ideas will be to construct all possible finite fields.

Example: A Field of Size Four. Consider the field of two elements, $\mathbb{F}_2 = \{0, 1\}$. I claim that the polynomial $x^2 + x + 1 \in \mathbb{F}_2[x]$ is irreducible over $\mathbb{F}_2$. Since this polynomial has degree

[117] No we won't, because we don't have time.

2, we only have to show that it has no roots in $\mathbb{F}_2$. And this is easy because there are only two elements to check:

x	0	1
x^2+x+1	1	1

Then since x^2+x+1 is a prime element of the Euclidean domain $\mathbb{F}_2[x]$, it follows that the quotient ring is a field. Let's call it

$$\mathbb{E} = \frac{\mathbb{F}_2[x]}{(x^2+x+1)\mathbb{F}_2[x]}.$$

I claim that this field has four elements. To prove this we first consider the quotient homomorphism, sending a polynomial $f(x) \in \mathbb{F}_2[x]$ to the coset of $(x^2+x+1)\mathbb{F}_2[x]$ that it generates:

$$\begin{array}{rccc} \varphi: & \mathbb{F}_2[x] & \to & \mathbb{E} \\ & f(x) & \mapsto & f(x)+(x^2+x+1)\mathbb{F}_2[x]. \end{array}$$

Recall that we can view $\mathbb{F}_2 \subseteq \mathbb{F}_2[x]$ as the subring of constant polynomials. Similarly, we can view $\mathbb{F}_2 \subseteq \mathbb{E}$ as the subring of cosets generated by constant polynomials. To be precise, the homomorphism φ restricted to $\mathbb{F}_2$ is injective:

$$\begin{array}{rccc} \varphi: & \mathbb{F}_2 & \to & \mathbb{E} \\ & a & \mapsto & a+(x^2+x+1)\mathbb{F}_2[x]. \end{array}$$

Indeed, if $\varphi(a) = \varphi(b)$ then the constant polynomial $a-b \in \mathbb{F}_2[x]$ is in the coset $(x^2+x+1)\mathbb{F}_2[x]$, which implies that $a-b$ is divisible by x^2+x+1. For reasons of degree this is only possible if $a-b=0$, and hence $a=b$.

Another way to say this is that the kernel of $\varphi : \mathbb{F}_2 \to \mathbb{E}$ is the zero ideal. Then from the First Isomorphism Theorem we have $\varphi(\mathbb{F}_2) = \operatorname{im}\varphi \cong \mathbb{F}_2/\{0\} = \mathbb{F}_2$. We will identify $\mathbb{F}_2$ with the subring $\varphi(\mathbb{F}_2) \subseteq \mathbb{E}$ by writing "a" instead of $a+(x^2+x+1)\mathbb{F}_2[x]$.[118]

So we have constructed a field extension $\mathbb{E} \supseteq \mathbb{F}_2$. In fact, I claim that $\mathbb{E} = \mathbb{F}_2[\alpha]$ for some special element $\alpha \in \mathbb{E}$. Indeed, let α be the coset generated by x:

$$\alpha := x+(x^2+x+1)\mathbb{F}_2[x].$$

Note that any element of $\mathbb{E}$ looks like $f(x)+(x^2+x+1)\mathbb{F}_2[x]$ for some polynomial $f(x) \in \mathbb{F}_2[x]$. Let's say $f(x) = \sum_k a_k x^k$. Then from the ring operations in $\mathbb{E}$ we have[119]

$$\begin{aligned} f(\alpha) &= \sum_k \big(a_k + (x^2+x+1)\mathbb{F}_2[x]\big)\big(x+(x^2+x+1)\mathbb{F}_2[x]\big)^k \\ &= (\sum_k a_k x^k) + (x^2+x+1)\mathbb{F}_2[x] \end{aligned}$$

[118] This is common practice in algebra. We do the same thing when we identify the integer $a \in \mathbb{Z}$ with the fraction $a/1 \in \mathbb{Q}$, or the real number $a \in \mathbb{R}$ with the complex number $a+0i \in \mathbb{C}$.

[119] This is hard to process at first, but I assure you that there is nothing interesting going on here.

$$= f(x) + (x^2 + x + 1)\mathbb{F}_2[x].$$

Furthermore, we observe that this element α is a root of $x^2 + x + 1$ because

$$\alpha^2 + \alpha + 1 = x^2 + x + 1 + (x^2 + x + 1)\mathbb{F}_2[x] = 0 + (x^2 + x + 1)\mathbb{F}_2[x].$$

But we already know that $x^2 + x + 1$ is irreducible over $\mathbb{F}_2$ and hence $x^2 + x + 1$ is the minimal polynomial of α over $\mathbb{F}_2$:

$$m_{\alpha/\mathbb{F}_2}(x) = x^2 + x + 1.$$

Finally, it follows from the Minimal Polynomial Theorem that

$$\mathbb{E} = \mathbb{F}_2[\alpha] = \{a + b\alpha : \text{for unique } a, b \in \mathbb{F}_2\}.$$

Since $\mathbb{F}_2$ consists of only two elements $\{0, 1\}$ we conclude that $\mathbb{E}$ has four elements:

$$\mathbb{E} = \{0 + 0\alpha, 1 + 0\alpha, 0 + 1\alpha, 1 + 1\alpha\} = \{0, 1, \alpha, 1 + \alpha\}.$$

The addition and multiplication tables for $\mathbb{E}$ are obtained by reducing the coefficients mod 2 and using the fact that $\alpha^2 = -1 - \alpha = 1 + \alpha$:

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

$\times$	0	1	α	$1+\alpha$
0	0	0	0	0
1	0	1	α	$1+\alpha$
α	0	α	$1+\alpha$	1
$1+\alpha$	0	$1+\alpha$	1	α

The steps leading up to these tables were abstract, but once we have them it is easy to teach a computer how to work with this field. In fact, such fields are used extensively in cryptography and error correcting codes.

Example: A Field of Size Eight. If we can find an irreducible of polynomial of degree 3 in the ring $\mathbb{F}_2[x]$ then the same reasoning as in the previous example will yield a field of size $2^3 = 8$. In fact, the ring $\mathbb{F}_2[x]$ contains exactly two irreducible polynomials of degree 3:

x	0	1
$x^3 + x^2 + 1$	1	1
$x^3 + x + 1$	1	1

Choosing the first of these gives the following field of size eight:

$$\mathbb{E} = \{a + b\alpha + c\alpha^2 : \text{for unique } a, b, c \in \mathbb{F}_2, \text{ where } \alpha^3 + \alpha^2 + 1 = 0\}.$$

Here is the multiplication table:

$\times$	0	1	α	$1+\alpha$	α^2	$1+\alpha^2$	$\alpha+\alpha^2$	$1+\alpha+\alpha^2$
0	0	0	0	0	0	0	0	0
1	0	1	α	$1+\alpha$	α^2	$1+\alpha^2$	$\alpha+\alpha^2$	$1+\alpha+\alpha^2$
α	0	α	α^2	$\alpha+\alpha^2$	$1+\alpha^2$	$1+\alpha+\alpha^2$	1	$1+\alpha$
$1+\alpha$	0	$1+\alpha$	$\alpha+\alpha^2$	$1+\alpha^2$	1	α	$1+\alpha+\alpha^2$	α^2
α^2	0	α^2	$1+\alpha^2$	1	$1+\alpha+\alpha^2$	$1+\alpha$	α	$\alpha+\alpha^2$
$1+\alpha^2$	0	$1+\alpha^2$	$1+\alpha+\alpha^2$	α	$1+\alpha$	$\alpha+\alpha^2$	α^2	1
$\alpha+\alpha^2$	0	$\alpha+\alpha^2$	1	$1+\alpha+\alpha^2$	α	α^2	$1+\alpha$	$1+\alpha^2$
$1+\alpha+\alpha^2$	0	$1+\alpha+\alpha^2$	$1+\alpha$	α^2	$\alpha+\alpha^2$	1	$1+\alpha^2$	α

You might think that the other polynomial x^3+x+1 gives a different field of size eight, but we will prove that any two finite fields of the same size must be isomorphic. More generally, we have the following important theorem. It is remarkable that finite fields are completely understood. This situation is much different, for example, from the theory of finite groups, which can be arbitrarily complicated.

The Classification of Finite Fields

(1) Any finite field has size p^k where p is prime and $k \geqslant 1$.

(2) There exists a field of size p^k for any prime p and integer $k \geqslant 0$.

(3) Any two finite fields of the same size are isomorphic.

It is common to write $q = p^k$ and to denote the unique field of size q by

$$\mathbb{F}_q \quad \text{or} \quad \mathrm{GF}(q).$$

The notation GF stands for "Galois Field", since the study of finite fields beyond $\mathbb{Z}/p\mathbb{Z}$ was initiated by Galois. However, he did not express his results in this language since the concept of a "field" was not invented until the 1870s. Dedekind used the word *Körper* (body) and Kronecker used the word *Bereich* (realm). The English translation "field" was given by E.H. Moore (1896), who first proved this theorem in its modern form.

In the remainder of this section we will prove (1) and then in the next section we will use the theory of splitting fields to prove (2) and (3). In order to prove (1) we need the following concept.

The Characteristic of a Field

For any ring R there exists a unique ring homomorphism $\varphi : \mathbb{Z} \to R$ from the ring of integers. The kernel of φ, being an ideal of $\mathbb{Z}$ must have the form $n\mathbb{Z}$ for some unique integer $n \geqslant 0$. We call this integer the *characteristic* of the ring:

$$\text{char}(R) := n.$$

If $\mathbb{E}$ is a field then we must have $\text{char}(\mathbb{E}) = 0$ or $\text{char}(\mathbb{E}) = p$ with p prime. If $\mathbb{E}$ is a **finite field** then we must have $\text{char}(\mathbb{E}) \neq 0$. Hence there exists[120] a prime p such that $\mathbb{E}$ contains a subring isomorphic to $\mathbb{F}_p$:

$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} = \mathbb{Z}/\ker\varphi \cong \text{im}\,\varphi \subseteq \mathbb{E}.$$

Proof. Any ring homomorphism $\varphi : \mathbb{Z} \to R$ must, in particular, be a group homomorphism $\varphi : (\mathbb{Z}, +, 0) \to (R, +, 0)$ sending 1 to 1. We know from the previous chapter that a unique such homomorphism exists and is given by the following definition:[121]

$$\varphi(k) = k \cdot 1 := \begin{cases} \overbrace{1 + 1 + \cdots + 1}^{k \text{ times}} & \text{if } k \geqslant 1, \\ 0 & \text{if } k = 0, \\ \underbrace{-1 - 1 - \cdots - 1}_{-k \text{ times}} & \text{if } k \leqslant -1. \end{cases}$$

One can check that this function also preserves multiplication, hence is a ring homomorphism.

Since the kernel of φ is an ideal of $\mathbb{Z}$ we have $\ker\varphi = n\mathbb{Z}$ for some unique integer $n \geqslant 0$. Then from the First Isomorphism Theorem we see that R contains a subring isomorphic to $\mathbb{Z}/n\mathbb{Z}$:

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\ker\varphi \cong \text{im}\,\varphi \subseteq R.$$

If the ring R is finite, then we must have $n \geqslant 1$ because $\mathbb{Z}/0\mathbb{Z} \cong \mathbb{Z}$ is infinite.

Now suppose that $\mathbb{E} = R$ is a field with $\text{char}(\mathbb{E}) = n$, and let $S \subseteq \mathbb{E}$ be any subring. Suppose that we have $a, b \in S$ with $ab = 0$ and $a \neq 0$. Then since a^{-1} exists in $\mathbb{E}$ we must have

$$\begin{aligned} ab &= 0 \\ aa^{-1}b &= a^{-1}0 \\ b &= 0. \end{aligned}$$

[120] We will see below that this prime is unique.

[121] In the previous chapter we used multiplicative language. Recall: For any group element $a \in G$ there exists a unique group homomorphism $\varphi : (\mathbb{Z}, +, 0) \to G$ sending k to "a^k". When the group structure on G is additive, we will write "$k \cdot a$" instead.

Hence S is a domain. In particular, $\operatorname{im}\varphi \subseteq \mathbb{E}$ must be a domain. Then since $\mathbb{Z}/n\mathbb{Z} \cong \operatorname{im}\varphi$, the theorem on Quotients of Euclidean Domains implies that $n = 0$ or $n = p$ is prime. And if $\mathbb{E}$ is a **finite field** then the case $n = 0$ is impossible. □

From this and a bit of linear algebra we obtain part (1) of the Classification of Finite Fields.

Proof of (1). Let $\mathbb{E}$ be a finite field. From the previous theorem we have $\mathbb{F}_p \subseteq \mathbb{E}$ for some prime p. Thus we can view $\mathbb{E}$ as a vector space over $\mathbb{F}_p$, defining scalar multiplication $\mathbb{F}_p \times \mathbb{E} \to \mathbb{E}$ via the field multiplication. Since $\mathbb{E}$ is a finite set, this must be a finite-dimensional vector space. In other words, there exists a finite basis $\alpha_1, \ldots, \alpha_k \in \mathbb{E}$ such that every element $\beta \in \mathbb{E}$ can be expressed **uniquely** in the form

$$\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_k\alpha_k,$$

for some $b_1, \ldots, b_k \in \mathbb{F}_p$.[122] This gives a bijection between elements of $\mathbb{E}$ and k-tuples of elements of $\mathbb{F}_p$, which implies that

$$\#\mathbb{E} = \#\mathbb{F}_p^k = (\#\mathbb{F}_p)^k = p^k.$$

□

Remark: It follows from this proof that if $\mathbb{F}_{p_1} \subseteq \mathbb{E}$ and $\mathbb{F}_{p_2} \subseteq \mathbb{E}$ for primes p_1, p_2 then we must have $p_1 = p_2$. Indeed, this would imply that

$$p_1^{k_1} = \#\mathbb{E} = p_2^{k_2}$$

for some integers $k_1, k_2 \geqslant 1$, which can only happen if $p_1 = p_2$.

8.4 Existence and Uniqueness of Splitting Fields

In this section we complete the Classification of Finite Fields by proving (2) that finite fields of sizes p^k exist, and (3) that any two finite fields of the same size are isomorphic. We will prove this in four steps:

- Any polynomial $f(x) \in \mathbb{F}[x]$ over any field $\mathbb{F}$ has a splitting field.
- Any splitting field of $x^{p^k} - x \in \mathbb{F}_p[x]$ has size p^k.
- Any field of size p^k is a splitting field for the polynomial $x^{p^k} - x \in \mathbb{F}_p[x]$.
- Any two splitting fields for the same polynomial are isomorphic.

The first step is called Kronecker's Theorem. In addition to proving the existence of finite fields, this result also completes our proof of the Fundamental Theorem of Algebra.

[122] Start with the whole field $S = \mathbb{E}$. If any element of S is expressible as an $\mathbb{F}_p$-linear combination of the other elements of S, throw it away. Continue until no element of S is expressible as an $\mathbb{F}_p$-linear combination of the others. The result will be the desired basis.

Kronecker's Theorem

Consider a non-constant polynomial $f(x) \in \mathbb{F}[x]$ of degree n over a field $\mathbb{F}$. We say that $\mathbb{E} \supseteq \mathbb{F}$ is a *splitting field for $f(x)$ over* $\mathbb{F}$ when the following two properties hold:

- There exist $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$ such that $f(x) = (x-\alpha_1)\cdots(x-\alpha_n)$.[123] In other words, *$f(x)$ splits over* $\mathbb{E}$.
- For any field $\mathbb{E} \supseteq \mathbb{E}' \supseteq \mathbb{F}$ such that $f(x)$ splits over $\mathbb{E}'$, we must have $\mathbb{E}' = \mathbb{E}$.

Kronecker's Theorem says that

splitting fields always exist.

The slogan of the proof is to "pretend hard enough and things will work out".

Proof. Consider any irreducible factor $m_1(x)|f(x)$ in $\mathbb{F}[x]$ and consider the field

$$\mathbb{F}_1 := \frac{\mathbb{F}[x]}{m_1(x)\mathbb{F}[x]}.$$

As in the previous section, we will view $\mathbb{F} \subseteq \mathbb{F}_1$ as the subring of (cosets generated by) constant polynomials. Now let $\alpha_1 \in \mathbb{F}_1$ denote the coset generated by x:

$$\alpha_1 := x + m_1(x)\mathbb{F}[x].$$

As in our discussion of fields of size four, it is a "trivial fact" that $m_1(\alpha_1) = 0$ in the field $\mathbb{F}_1$, the only difficultly being that this notation hides all the details of the construction. Furthermore, since $m_1(x)$ is irreducible over $\mathbb{F}$ the MPT implies that $\mathbb{F}_1 = \mathbb{F}[\alpha_1]$.

Since $m_1(\alpha_1) = 0$ and $m_1(x)|f(x)$ we have $f(\alpha_1) = 0$, hence by Descartes' Theorem we have

$$f(x) = (x - \alpha_1)f_1(x) \text{ for some } f_1(x) \in \mathbb{F}_1[x].$$

If $f_1(x)$ is constant then we are done. Otherwise, let $m_2(x)|f_1(x)$ be an irreducible factor in the ring $\mathbb{F}_1[x]$ and consider the field

$$\mathbb{F}_2 := \frac{\mathbb{F}_1[x]}{m_2(x)\mathbb{F}_1[x]}.$$

Think of $\mathbb{F}_1 \subseteq \mathbb{F}_2$ as the subring of cosets generated by constant polynomials[124] and let $\alpha_2 \in \mathbb{F}_2$ be the coset generated by x:

$$\alpha_2 := x + m_2(x)\mathbb{F}_1[x].$$

[123] It does no harm to assume that $f(x)$ is monic.

[124] To be completely precise, these are cosets of cosets. Now you see why it is necessary to abuse the notation.

As before, we have $m_2(\alpha_2) = 0$, which since $m_2(x)$ is irreducible over $\mathbb{F}_1$ implies that $\mathbb{F}_2 = \mathbb{F}_1[\alpha_2] = \mathbb{F}[\alpha_1, \alpha_2]$.[125] Also, since $m_2(x)|f_1(x)$ we have $f_1(\alpha_2) = 0$ and hence

$$f_1(x) = m_2(x)f_2(x) \text{ for some } f_2(x) \in \mathbb{F}_2[x].$$

If $f_2(x)$ is constant then we are done. Otherwise, we repeat the process to create a chain of field extensions $\mathbb{F} \subseteq \mathbb{F}_1 \subseteq \cdots \subseteq \mathbb{F}_n$ and elements $\alpha_i \in \mathbb{F}_i$ such that

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).$$

In other words, $f(x)$ splits over $\mathbb{F}_n$. By construction we also have $\mathbb{F}_n = \mathbb{F}[\alpha_1, \ldots, \alpha_n]$ which means that $\mathbb{F}_n$ itself is the smallest subring of $\mathbb{F}_n$ containing $\mathbb{F}$ and the elements $\alpha_1, \ldots, \alpha_n$.

Finally, suppose that $f(x)$ splits over a field $\mathbb{E}$ where $\mathbb{F}_n \supseteq \mathbb{E} \supseteq \mathbb{F}$. In this case we will show that $\mathbb{E} = \mathbb{F}_n$ and hence that $\mathbb{F}_n$ is a splitting field for $f(x)$ over $\mathbb{F}$. By assumption there exist $\beta_1, \ldots, \beta_n \in \mathbb{E}$ such that

$$f(x) = (x - \beta_1)(x - \beta_2) \cdots (x - \beta_n).$$

But then in the ring $\mathbb{F}_n[x]$ we have

$$(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) = (x - \beta_1)(x - \beta_2) \cdots (x - \beta_n).$$

Substituting $x = \alpha_1$ gives

$$0 = (\alpha_1 - \alpha_1)(\alpha_1 - \alpha_2) \cdots (\alpha_1 - \alpha_n) = (\alpha_1 - \beta_1)(\alpha_1 - \beta_2) \cdots (\alpha_1 - \beta_n),$$

which implies that $\alpha_1 - \beta_j = 0$ (and hence $\alpha_1 = \beta_j$) for some j. Similarly, by substituting $x = \alpha_i$ we find that every α_i is equal to some β_j and hence is an element of $\mathbb{E}$. We have shown that the ring $\mathbb{E}$ contains $\mathbb{F}$ and the elements $\alpha_1, \ldots, \alpha_n$, and it follows that $\mathbb{E} = \mathbb{F}_n$ as desired. □

What do you think of this proof? There is a reason that it didn't get written down until the 1880s. We will use the existence of splitting fields to prove the existence of finite fields of every possible size. But for this we need two more lemmas.

Repeated Roots

For any field $\mathbb{F}$ we define the *formal derivative* $D : \mathbb{F}[x] \to \mathbb{F}[x]$ by[126]

$$D\left(\sum a_k x^k\right) = \sum k \cdot a_k x^{k-1}.$$

This satisfies all of the usual algebraic properties of derivatives, such as the product rule.

Now consider any element of a field extension, $\alpha \in \mathbb{E} \supseteq \mathbb{F}$. We say that α is a *repeated root* of $f(x)$ when[127]

$$f(x) = (x - \alpha)^2 g(x) \text{ for some } g(x) \in \mathbb{E}[x].$$

[125] We use the notation $\mathbb{F}[\alpha_1, \alpha_2]$ to denote the field $\mathbb{F}[\alpha_1][\alpha_2]$.

I claim that

$$\alpha \text{ is a repeated root of } f(x) \quad\Longleftrightarrow\quad f(\alpha)=0 \text{ and } Df(\alpha)=0.$$

Proof. Consider an element of a field extension $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ and a polynomial $f(x) \in \mathbb{F}[x]$. First we suppose that α is a repeated root of $f(x)$. That is, we suppose that $f(x) = (x-\alpha)^2 g(x)$ for some $g(x) \in \mathbb{E}[x]$. Then from the product rule we have

$$Df(x) = 2(x-\alpha)g(x) + (x-\alpha)^2 Dg(x),$$

and hence

$$Df(\alpha) = 2(\alpha-\alpha)g(\alpha) + (\alpha-\alpha)Dg(\alpha) = 0.$$

Conversely, consider any polynomial $f(x) \in \mathbb{F}[x]$ such that $f(\alpha) = 0$ and $Df(\alpha) = 0$. Since $f(\alpha) = 0$, Descartes' Factor Theorem in the ring $\mathbb{E}[x]$ tells us that

$$f(x) = (x-\alpha)g(x) \text{ for some } g(x) \in \mathbb{E}[x].$$

Now we compute the derivative of $f(x)$ using the product rule:

$$Df(x) = g(x) + (x-\alpha)Dg(x).$$

Since $Df(\alpha) = 0$ we must have

$$0 = Df(\alpha) = g(\alpha) + (\alpha-\alpha)Dg(\alpha) = g(\alpha).$$

But then Descartes' Factor Theorem in the ring $\mathbb{E}[x]$ says that

$$g(x) = (x-\alpha)h(x) \text{ for some } h(x) \in \mathbb{E}[x],$$

and hence α is a repeated root of $f(x)$:

$$\begin{aligned} f(x) &= (x-\alpha)g(x) \\ &= (x-\alpha)(x-\alpha)h(x) \\ &= (x-\alpha)^2 h(x). \end{aligned}$$

□

[126] Given $k \in \mathbb{Z}$ and $a_k \in \mathbb{F}$, the element $k \cdot a_k \in \mathbb{F}$ is defined repeated addition or subtraction. More precisely, we define $k \cdot a_k = \varphi(k)$ where $\varphi : (\mathbb{Z}, +, 0) \to (\mathbb{F}, +, 0)$ is the unique group homomorphism sending 1 to a_k.

[127] We do not exclude the possibility that $g(\alpha) = 0$, in which case $f(x)$ is disible by $(x-\alpha)^3$ or some higher power.

Freshman's Dream

Let R be a ring of prime characteristic $p > 0$. Then for any elements $a, b \in R$ we have

$$(a \pm b)^p = a^p \pm b^p,$$

and by induction we have

$$(a \pm b)^{p^k} = a^{p^k} \pm b^{p^k} \text{ for any integer } k \geqslant 0.$$

We will apply this in the special case when $R = \mathbb{E} \supseteq \mathbb{F}_p$ is a field extension of $\mathbb{F}_p$.

Proof. For any ring R and for any integer $e \geqslant 0$ we have the Binomial Theorem:

$$(a+b)^e = a^e + \binom{e}{1} \cdot a^{e-1}b + \binom{e}{2} \cdot a^{e-2}b^2 + \cdots + \binom{e}{e-1} \cdot ab^{e-1} + b^e.$$

Suppose that $\text{char}(R) = n$. That is, suppose that the unique ring homomorphism from the integers, $\varphi : \mathbb{Z} \to R$, has kernel $n\mathbb{Z}$. By definition, this means that for nonzero $a \in R$ and $k \in \mathbb{Z}$ we have $k \cdot a = 0$ if and only if k is divisible by n.

If $p \in \mathbb{Z}$ is prime and $\text{char}(R) = p$, our first goal is to show that

$$(a+b)^p = a^p + b^p.$$

From the above remarks, it is enough to show that $\binom{p}{r}$ is divisible by p for all $1 \leqslant r \leqslant p-1$. Recall the formula for binomial coefficients:

$$\binom{p}{r} = \frac{p!}{r!(p-r)!} = \frac{p(p-1)(p-2)\cdots 2 \cdot 1}{r(r-1)\cdots 2 \cdot 1 \cdot (p-r)(p-r-1)\cdots 2 \cdot 1}.$$

We know that this is an integer, hence every prime factor in the denominator must be canceled by a prime factor in the numerator. Thus we need to show that the prime p occurs with higher multiplicity in the numerator than it does in the denominator. In fact, I claim that p occurs with multiplicity 1 in the numerator and with multiplicity 0 in the denominator. Indeed, we clearly have $p|p!$. But p does not divide the product $(p-1)(p-2)\cdots 2 \cdot 1$ because if it did then by Euclid's Lemma p would divide one of the factors, but each factor is smaller than p. Similarly, if $1 \leqslant r \leqslant p-1$ then p does not divide the product $r!(p-r)!r(r-1)\cdots 2 \cdot 1 \cdot (p-r)(p-r-1)\cdots 2 \cdot 1$ since each factor in this product is smaller than p.

Now I claim that we also have

$$(a-b)^p = a^p - b^p.$$

If $p = 2$ then this follows from the previous because $a = -a$ in a ring of characteristic 2. So suppose that $p > 2$. Since p is prime this implies that p is odd, so that

$$(a-b)^p = a^p - \binom{p}{1} \cdot a^{p-1}b + \binom{p}{2} \cdot a^{p-2}b^2 - \cdots + (-1)^{p-1}\binom{p}{p-1} \cdot ab^{p-1} + (-1)^p \cdot b^p$$

$$
\begin{aligned}
&= a^p - 0 + 0 - \cdots + 0 - b^p \\
&= a^p - b^p.
\end{aligned}
$$

Finally, we observe by induction that for any integer $k \geqslant 1$ we have

$$
\begin{aligned}
(a \pm b)^{p^k} &= \left((a \pm b)^{p^{k-1}}\right)^p \\
&= \left(a^{p^{k-1}} \pm b^{p^{k-1}}\right)^p \\
&= \left(a^{p^{k-1}}\right)^p \pm \left(b^{p^{k-1}}\right)^p \\
&= a^{p^k} \pm b^{p^k}.
\end{aligned}
$$

□

Proof of (2). For any prime power p^k we want to show that a field of size p^k exists. The key trick, used by Gauss and Galois, is to consider the following polynomial:

$$g(x) = x^{p^k} - x \in \mathbb{F}_p[x].$$

Let $\mathbb{E} \supseteq \mathbb{F}_p$ be a splitting field for $g(x)$ over $\mathbb{F}_p$, which exists by Kronecker's Theorem. Then I claim that $\#\mathbb{E} = p^k$. To prove this, we consider the set of roots of $g(x)$ in the field $\mathbb{E}$:

$$S := \{\alpha \in \mathbb{E} : g(\alpha) = 0\}.$$

The proof will follow from two facts:

(i) $\#S = p^k$,

(ii) $S = \mathbb{E}$.

To prove (i) we first observe that $\#S \leqslant p^k$ since a polynomial of degree p^k can have at most p^k roots in any field. Now consider the formal derivative of $g(x)$ in the ring $\mathbb{F}_p[x]$:

$$Dg(x) = p^k x^{p^k - 1} - 1 = 0x^{p^k} - 1 = -1.$$

Since this polynomial is constant and nonzero it cannot have a root in any field extension. It follows from the lemma that $g(x)$ cannot have a repeated root in any field extension. Since $g(x)$ splits[128] in $\mathbb{E}$ this implies that $g(x)$ has p^k distinct roots in $\mathbb{E}$, hence $\#S = p^k$.

To prove (ii) we will show that (surprisingly!) the subset $S \subseteq \mathbb{E}$ is actually a **subfield**. Then since $\mathbb{E}$ is a splitting field for $g(x)$ and since $g(x)$ splits over S it will follow that $S = \mathbb{E}$. There are two things to check:

- Let $\alpha, \beta \in S$ so that $\alpha^{p^k} = \alpha$ and $\beta^{p^k} = \beta$. Then

  $$(\alpha\beta^{-1})^{p^k} = \alpha^{p^k}\left(\beta^{p^k}\right)^{-1} = \alpha\beta^{-1},$$

  so that $\alpha\beta^{-1} \in S$.

[128] Indeed, $\mathbb{E}$ is a splitting field for $g(x)$.

- Let $\alpha,\beta \in S$ so that $\alpha^{p^k} = \alpha$ and $\beta^{p^k} = \beta$. Then from the Freshman's Dream we have

  $$(\alpha-\beta)^{p^k} = \alpha^{p^k} - \beta^{p^k} = \alpha - \beta,$$

  so that $\alpha - \beta \in S$.

□

This abstract existence proof is not particularly useful. For practical purposes it is better to start with an irreducible polynomial in $\mathbb{F}_p[x]$ of degree k. The existence of finite fields of every size is equivalent to the statement that there exist irreducible polynomials in $\mathbb{F}_p[x]$ of every degree. Apparently there exist fast algorithms for finding such polynomials.[129]

To complete the Classification of Finite Fields, it only remains to show that any two finite fields of the same size are isomorphic. This is the hardest part, and the proof was apparently not written down until 1896.[130] We will use the following theorem, which has applications beyond finite fields.

Uniqueness of Splitting Fields

Consider a non-constant polynomial $f(x) \in \mathbb{F}[x]$ over a field $\mathbb{F}$. Suppose that $\mathbb{E} \supseteq \mathbb{F}$ and $\mathbb{E}' \supseteq \mathbb{F}$ are splitting fields for $f(x)$. Then there exists a ring isomorphism $\mathbb{E} \cong \mathbb{E}'$.

The proof will use induction on the degree of the polynomial. In order to facilitate the induction step we will actually prove a more general statement.

Uniqueness of Splitting Fields (General Statement)

Any isomorphism of fields $\varphi : \mathbb{F} \to \mathbb{F}'$ induces an isomorphism of polynomial rings $\varphi : \mathbb{F}[x] \to \mathbb{F}'[x]$ by acting on coefficients:

$$\begin{array}{rccc} \varphi : & \mathbb{F}[x] & \to & \mathbb{F}'[x] \\ & \sum_k a_k x^k & \mapsto & \sum_k \varphi(a_k) x^k. \end{array}$$

We use the same symbol φ for both isomorphisms to save notation. We will also write $f^\varphi(x) = \varphi(f(x))$ to save notation.[131]

Now consider a non-constant polynomial $f(x) \in \mathbb{F}[x]$ and its image polynomial $f^\varphi(x) \in \mathbb{F}'[x]$. If $\mathbb{E} \supseteq \mathbb{F}$ is a splitting field for $f(x)$ and if $\mathbb{E}' \supseteq \mathbb{F}'$ is a splitting field for $f^\varphi(x)$ then I claim that there exists an isomorphism $\tilde{\varphi} : \mathbb{E} \to \mathbb{E}'$ with the property that $\tilde{\varphi}(a) = \varphi(a)$

[129] `https://arxiv.org/abs/0905.1642`

[130] E.H. Moore, *A Doubly-Infinite System of Simple Groups*, 1896.

for all $a \in \mathbb{F}$. Here is a diagram:

$$\begin{tikzcd}
\mathbb{E} \arrow[r, "\tilde{\varphi}"] & \mathbb{E}' \\
\mathbb{F} \arrow[u] \arrow[r, "\varphi"'] & \mathbb{F}' \arrow[u]
\end{tikzcd}$$

The vertical arrows here are just the "inclusion homomorphisms", sending elements of $\mathbb{F}$ and $\mathbb{F}'$ to themselves. The idea of this diagram is that the two composite functions from $\mathbb{F}$ to $\mathbb{E}'$ yield the same function.[132]

The induction step is a bit complicated, so we isolate this as a separate lemma.

The Lifting Lemma

We are given the following data:

- An isomorphism of fields $\varphi : \mathbb{F} \to \mathbb{F}'$.
- An irreducible polynomial $m(x) \in \mathbb{F}[x]$.
- Some field extensions $\mathbb{E} \supseteq \mathbb{F}$ and $\mathbb{E}' \supseteq \mathbb{F}'$.
- Elements $\alpha \in \mathbb{E}$ and $\beta \in \mathbb{E}'$ such that $m(\alpha) = 0$ and $m^{\varphi}(\beta) = 0$.

Since $m(x) \in \mathbb{F}[x]$ is irreducible, the image polynomial $m^{\varphi}(x) \in \mathbb{F}'[x]$ is also irreducible. It follows from the Minimal Polynomial Theorem that $m(x)$ and $m^{\varphi}(x)$ are the minimal polynomials of $\alpha/\mathbb{F}$ and $\beta/\mathbb{F}'$, respectively. Furthermore, the subrings $\mathbb{F}[\alpha] \subseteq \mathbb{E}$ and $\mathbb{F}'[\beta] \subseteq \mathbb{E}'$ are actually fields. Finally, we obtain an isomorphism of fields $\hat{\varphi} : \mathbb{F}[\alpha] \to \mathbb{F}'[\beta]$ by composing the following three isomorphisms:

$$\mathbb{F}[\alpha] \cong \frac{\mathbb{F}[x]}{m(x)\mathbb{F}[x]} \cong \frac{\mathbb{F}'[x]}{m^{\varphi}(x)\mathbb{F}'[x]} \cong \mathbb{F}'[\beta].$$

The middle isomorphism is induced by φ. From the definitions, we observe that $\hat{\varphi}(\alpha) = \beta$ and that $\hat{\varphi}(a) = \varphi(a)$ for all $a \in \mathbb{F}$, where we think of $\mathbb{F} \subseteq \mathbb{F}[\alpha]$ and $\mathbb{F}' \subseteq \mathbb{F}'[\beta]$ as subfields.

We can summarize all of this information with a commutative diagram:

[131] Every author seems to invent their own notation for this operation, i.e., the operation of acting on coefficients by a ring homomorphism.

[132] We say that the diagram "commutes". Such "commutative diagrams" originated in the 1940s, which is quite late. But they have recently taken over most branches of algebra.

$$\begin{tikzcd}
\mathbb{E} & \mathbb{E}' \\
\mathbb{F}[\alpha] \arrow[u] \arrow[r, "\hat{\varphi}"] & \mathbb{F}'[\beta] \arrow[u] \\
\mathbb{F} \arrow[u] \arrow[r, "\varphi"'] & \mathbb{F}' \arrow[u]
\end{tikzcd}$$

We call this the Lifting Lemma because it lifts the isomorphism $\varphi : \mathbb{F} \to \mathbb{F}'$ to an isomorphism of field extensions $\hat{\varphi} : \mathbb{F}[\alpha] \to \mathbb{F}'[\beta]$. To prove the uniqueness of splitting fields we will adjoin the roots of a polynomial one at a time, each time lifting the original isomorphism, until we obtain an isomorphism between the splitting fields.

Proof of Uniqueness of Splitting Fields. Consider an isomorphism of fields $\varphi : \mathbb{F} \to \mathbb{F}'$ and a non-constant polynomial $f(x) \in \mathbb{F}[x]$. Let $\mathbb{E} \supseteq \mathbb{F}$ be a splitting field of $f(x)$ and let $\mathbb{E}' \supseteq \mathbb{F}$ be a splitting field of the image polynomial $f^{\varphi}(x) \in \mathbb{F}'[x]$. Our goal is to construct an isomorphism $\tilde{\varphi} : \mathbb{E} \to \mathbb{E}'$.

To begin, we consider any irreducible factor $m(x)|f(x)$ in the ring $\mathbb{F}[x]$. The image polynomial $m^{\varphi}(x)$ will be an irreducible factor of $f^{\varphi}(x)$ in the ring $\mathbb{F}'[x]$. Choose any element $\alpha_1 \in \mathbb{E}$ such that $m(\alpha_1) = 0$, which must exist because $\mathbb{E}$ is a splitting field,[133] and define $\beta_1 := \varphi(\alpha_1)$ so that $m^{\varphi}(\beta_1) = 0$. From the conditions of the Lifting Lemma, we obtain an isomorphism of fields $\varphi_1 : \mathbb{F}[\alpha_1] \to \mathbb{F}'[\beta_1]$, as in the following diagram:

$$\begin{tikzcd}
\mathbb{E} & \mathbb{E}' \\
\mathbb{F}[\alpha_1] \arrow[u] \arrow[r, "\varphi_1"] & \mathbb{F}'[\beta_1] \arrow[u] \\
\mathbb{F} \arrow[u] \arrow[r, "\varphi"'] & \mathbb{F}' \arrow[u]
\end{tikzcd}$$

Since $m(\alpha_1) = 0$ and $m(x)|f(x)$ we must have $f(\alpha_1) = 0$. Hence from Descartes' Factor Theorem we obtain

$$f(x) = (x - \alpha_1)f_1(x) \text{ for some } f_1(x) \in \mathbb{F}[\alpha_1][x].$$

If $f_1(x)$ is constant then we stop. Otherwise, we consider any irreducible factor $m_1(x)|f_1(x)$ in the ring $\mathbb{F}[\alpha_1][x]$. Since $f(x)$ splits in $\mathbb{E}$ and since $m_1(x)|f_1(x)|f(x)$, we can find some $\alpha_2 \in \mathbb{E}$

[133] Since $f(x)$ splits in $\mathbb{E}[x]$ and $m(x)|f(x)$, the uniqueness of prime factorization in $\mathbb{E}[x]$ implies that $m(x)$ also splits in $\mathbb{E}[x]$.

such that $m_1(\alpha_2) = 0$. Define $\beta_2 := \varphi_1(\alpha_1)$ so that $m_1^{\varphi_1}(\beta_1) = 0$. Then we can apply the Lifting Lemma again to obtain an isomorphism of fields $\varphi_2 : \mathbb{F}[\alpha_1, \alpha_2] \to \mathbb{F}'[\beta_1, \beta_2]$ making the following diagram commute:[134]

$$\begin{tikzcd}
\mathbb{E} & \mathbb{E}' \\
\mathbb{F}[\alpha_1, \alpha_2] \arrow[u] \arrow[r, "\varphi_2"] & \mathbb{F}'[\beta_1, \beta_2] \arrow[u] \\
\mathbb{F}[\alpha_1] \arrow[u] \arrow[r, "\varphi_1"] & \mathbb{F}'[\beta_1] \arrow[u] \\
\mathbb{F} \arrow[u] \arrow[r, "\varphi"'] & \mathbb{F}' \arrow[u]
\end{tikzcd}$$

We repeat this process until we obtain complete factorizations[135]

$$\begin{aligned} f(x) &= (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n), \\ f^{\varphi}(x) &= (x - \beta_1)(x - \beta_2) \cdots (x - \beta_n), \end{aligned}$$

and an isomorphism of fields $\varphi_n : \mathbb{F}[\alpha_1, \ldots, \alpha_n] \to \mathbb{F}'[\beta_1, \ldots, \beta_n]$. Since $f(x)$ splits over $\mathbb{F}[\alpha_1, \ldots, \alpha_n]$ and since $\mathbb{E}$ is a splitting field (i.e. is minimal with respect to splitting f) we must have $\mathbb{E} = \mathbb{F}[\alpha_1, \ldots, \alpha_n]$. Similarly, we must have $\mathbb{E}' = \mathbb{F}'[\beta_1, \ldots, \beta_n]$. Thus we have produced the desired isomorphism $\mathbb{E} \cong \mathbb{E}'$. □

Remark: Our original goal was to prove that any two splitting fields $\mathbb{E} \supseteq \mathbb{F}$ and $\mathbb{E}' \supseteq \mathbb{F}$ for the same polynomial $f(x) \in \mathbb{F}[x]$ are isomorphic. We obtain this from the above result by starting with the identity isomorphism $\varphi : \mathbb{F} \to \mathbb{F}$. Isomorphisms between splitting fields are certainly not unique. For example, if $\mathbb{E} \supseteq \mathbb{F}$ is a splitting field for $f(x) \in \mathbb{F}[x]$ then there might exist many *automorphisms* of $\mathbb{E}$ (i.e., self-isomorphisms $\varphi : \mathbb{E} \to \mathbb{E}$) fixing the elements of $\mathbb{F}$. The collection of such automorphisms is called the *Galois group of $f(x)$ over $\mathbb{F}$*. This is a fascinating topic, which we have no time to pursue.

Finally, we can complete our Classification of Finite Fields. The last step is to prove that any two finite fields of the same size are isomorphic. We will do this by showing that any two finite fields of the same size are splitting fields for the same polynomial.

Proof of (3). Let $\mathbb{E}$ and $\mathbb{E}'$ be fields of size p^k, with p prime. Then we must have $\text{char}(\mathbb{E}) = \text{char}(\mathbb{E}') = p$. Indeed, suppose that $\text{char}(\mathbb{E}) = p'$ for some prime p'. Then from the proof at the end of the last section we must have

$$p^k = \#\mathbb{E} = p_1^{k_1}$$

[134] We use the notation $\mathbb{F}[\alpha_1, \alpha_2]$ to denote the field $\mathbb{F}[\alpha_1][\alpha_2]$.

[135] It does no harm to assume that $f(x)$ is monic.

for some integer k_1, which can only happen if $p_1 = p$. Thus each of $\mathbb{E}$ and $\mathbb{E}'$ contains a field isomorphic to $\mathbb{F}_p$. I claim that each of $\mathbb{E}$ and $\mathbb{E}'$ is a splitting field for the polynomial $g(x) = x^{p^k} - x \in \mathbb{F}_p[x]$, from which it will follow that $\mathbb{E} \cong \mathbb{E}'$.

To prove this, I claim that **every element of $\mathbb{E}$ is a root of** $g(x)$. Indeed, since $\mathbb{E}$ is a field of size p^k, the group of units $(\mathbb{E}^\times, \cdot, 1)$ has size $p^k - 1$. Hence for any nonzero element $\alpha \in \mathbb{E}$ the generalized Euler-Fermat theorem tells us that

$$\alpha^{p^k - 1} = 1,$$

and multiplying both sides by α gives

$$\begin{aligned} \alpha^{p^k} &= \alpha \\ \alpha^{p^k} - \alpha &= 0 \\ g(\alpha) &= 0. \end{aligned}$$

This last equation also holds for $\alpha = 0$, hence it holds for every element of $\mathbb{E}$. In other words, we can write

$$g(x) = \prod_{\alpha \in \mathbb{E}} (x - \alpha).$$

It follows from this that $\mathbb{E}$ is a splitting field for $g(x)$, since any subfield of $\mathbb{E}$ must necessarily omit some of root of $g(x)$. A parallel argument shows that $\mathbb{E}'$ is a splitting field for $g(x)$. □

THE END

9 Introduction to Galois Theory

Nope, no time.