MTH 561: Abstract Algebra
Course Notes

Fall 2021
Drew Armstrong

Contents

1 Complex Numbers **2**
1.1 Cardano's Formula . . . 2
1.2 Complex Numbers as a Ring . . . 5
1.3 Complex Numbers as a Vector Space . . . 7
1.4 Complex Numbers as a Field . . . 11
1.5 Complex Numbers as Linear Functions . . . 12
1.6 Euler's Formula and Roots of Unity . . . 21

2 Introduction to Polynomials **29**
2.1 Rings of Polynomials . . . 29
2.2 Descartes' Theorem . . . 31
2.3 Polynomials: Functions or Formal Expressions? . . . 35
2.4 Concept of a Splitting Field . . . 37

3 Unique Prime Factorization **39**
3.1 Definition of Euclidean Domains . . . 39
3.2 The Euclidean Algorithm . . . 45
3.3 The Vector Euclidean Algorithm . . . 47
3.4 Unique Prime Factorization . . . 53
3.5 Irreducible Polynomials . . . 56

4 Some Number Theory **58**
4.1 Modular Arithmetic . . . 58
4.2 Some Finite Fields . . . 61
4.3 The Euler-Fermat Theorem . . . 64
4.4 The Chinese Remainder Theorem . . . 70

5 The Fundamental Theorem of Algebra **77**
5.1 Leibniz' Mistake . . . 77
5.2 Partial Fractions . . . 80
5.3 Equivalent Statements of the FTA . . . 80
5.4 Intermediate Value Theorem . . . 83
5.5 Descartes and Euler on Quartic Equations . . . 83
5.6 Waring's Method . . . 83
5.7 Laplace's Proof of the FTA . . . 83

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:

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

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)

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]

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

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

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.

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)

[3] 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.

(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"

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

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

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 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.

[4] The proof is extremely boring.

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)

(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:

$$(a+bi)+(c+di) = (a+c)+(b+d)i$$
$$a(b+ci) = (ab)+(ac)i.$$

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

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

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:

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

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

$$\alpha\alpha^* = |\alpha|^2$$
$$\alpha(\alpha^*/|\alpha|^2) = 1,$$

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

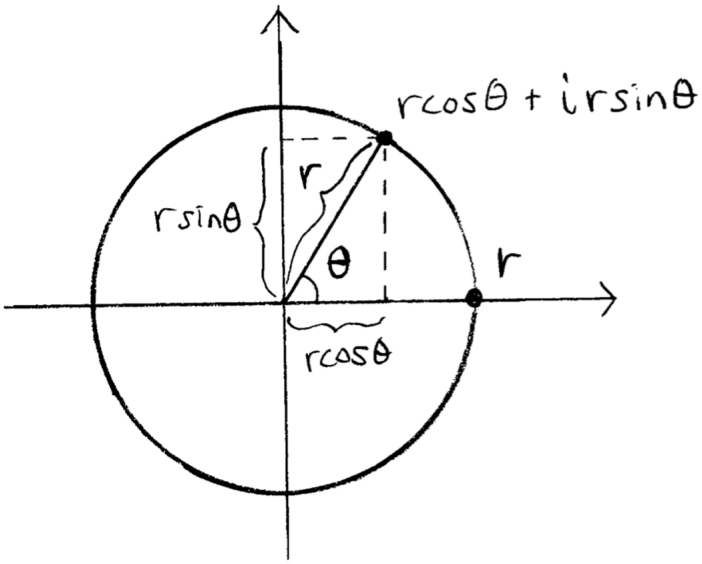
Based on the isomorphism $\mathbb{C} \cong \mathbb{R}^2$ we can view the complex number $a + bi$ as the point (a, b) in the Cartesian plane. But we can also express points of $\mathbb{R}^2$ in polar coordinates. That is, for any pair of real numbers (a, b), not both zero, there exist a unique pair of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

$$\alpha = r(\cos\theta + i\sin\theta),$$
$$\beta = s(\cos\lambda + i\sin\lambda).$$

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

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

$$= M(\mathbf{u}),$$

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:

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

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.

- **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 $S_r(\mathbf{u}) = r\mathbf{u}$

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

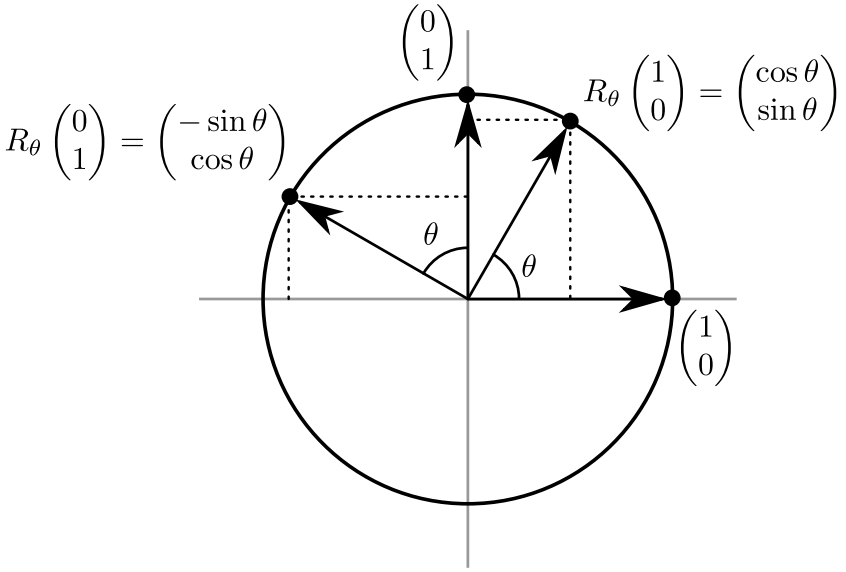
is linear. The corresponding matrix is

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

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 $\mathbf{e}_2 = (0, 1)$:



  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]

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

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

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

[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.

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

[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.

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.

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

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

$n \geqslant 1$ 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

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

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:

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

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

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.

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

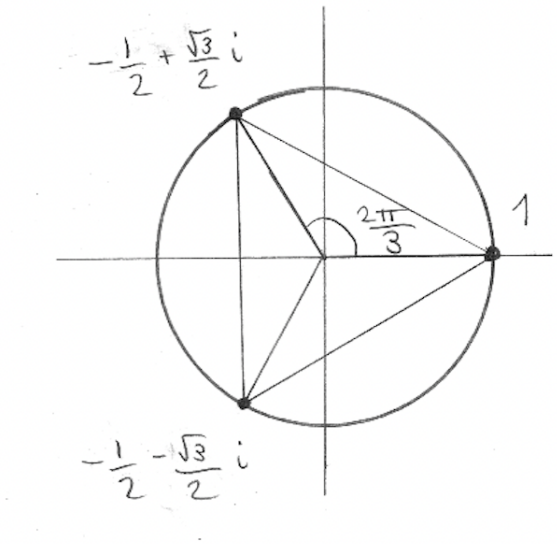
  $$\alpha'\omega^0 = \alpha' \quad \text{and} \quad \alpha'\omega^1 = -\alpha'.$$

  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:



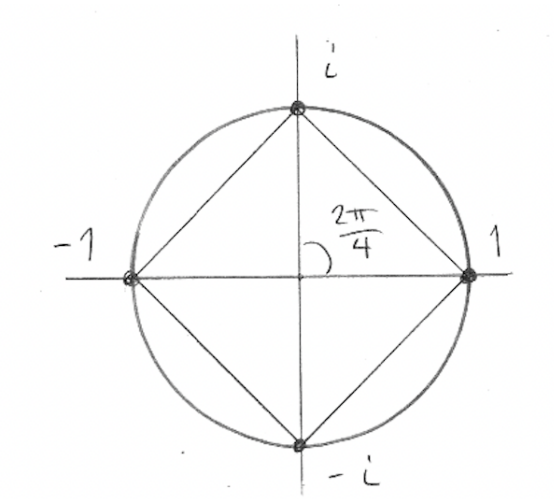
[9] We also need to know that an equation of degree n can have **no more than n roots**. You will prove this on the homework and we will discuss it more in the next section.

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

- $n = 4$: 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:



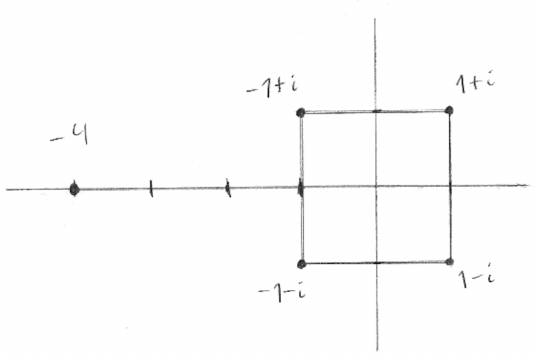
More generally, let's compute the 4th roots of $\alpha = -4$. First we express $\alpha = 4e^{\pi i}$ in polar form, so the principal 4th root is

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

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 $x^4 + 4$:

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

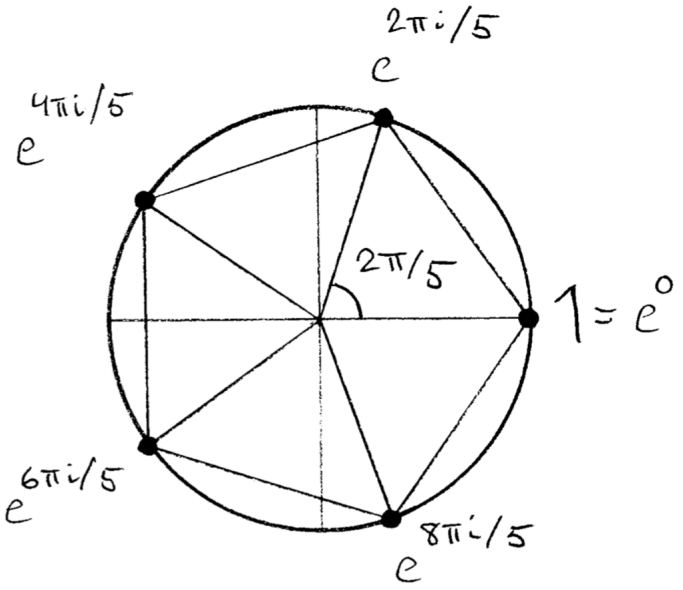
In 1702, Gottfried Leibniz claimed that the polynomial x^4+4 cannot be factored over the real numbers. However, we can show that he was wrong by grouping the four complex roots into "conjugate pairs":

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

- $n = 5$: 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.

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

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

[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.

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:

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

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]

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

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

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

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]

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

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 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]

$$\begin{array}{r|l} & x^2 - 3x - \frac{1}{2} \\ \hline 2x^2 + 1 & 2x^4 - 6x^3 + x - 1 \\ & \underline{-2x^4 - x^2} \\ & -6x^3 - x^2 + x \\ & \underline{6x^3 + 3x} \\ & -x^2 + 4x - 1 \\ & \underline{x^2 + \frac{1}{2}} \\ & 4x - \frac{1}{2} \end{array}$$

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

[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 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

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

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:

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

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:

$$f(x) = x + 0,$$
$$g(x) = x^3 + 0x^2 + 0x + 0.$$

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 $\mathbb{E}$. 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 $\mathbb{R}$.

[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.

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

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

Review for the First Exam

- hello

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

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

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

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

$a, b \in \mathbb{Z}$ 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:

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

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]

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

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

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

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

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

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

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

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

$$\begin{aligned} 1 &\leqslant |k| \\ |a| &\leqslant |a||k| \\ |a| &\leqslant |ak| \\ |a| &\leqslant |b|. \end{aligned}$$

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

[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.

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 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.

[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.

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*

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

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 $\mathrm{Div}(a, b)$ 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:

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

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 $\mathrm{Div}(a, b)$. 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:

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

(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:

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

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:

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

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 Vector 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.

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:

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

Hence from the lemma in the previous section we have:

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

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

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

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

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*.

The Vector 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 $r\mathbf{x} + r'\mathbf{x}' = (rx + r'x', ry +$

[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*.

$r'y', rz + r'z')$ 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:

$$\begin{aligned}
a &= bq_1 + r_1, \\
b &= r_1q_2 + r_2, \\
r_1 &= r_2q_3 + r_3, \\
&\vdots \\
r_{i-2} &= r_{i-1}q_i + r_i, \\
&\vdots \\
r_{n-2} &= r_{n-1}q_n + r_n, \\
r_{n-1} &= r_nq_{n+1} + 0,
\end{aligned}$$

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

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

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

[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.

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 Vector 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.

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 Vector 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] 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

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

[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.

even though the GCD is not generally unique. If a, b are coprime then it follows from the Vector 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

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

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.

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 Vector 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 $k \in R$:

$$\begin{aligned} ax + py &= 1 \\ abx + pby &= b \\ pkx + pby &= b \\ p(kx + by) &= b. \end{aligned}$$

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:

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

$$p_1 p_2 \cdots p_k = uu' p_1 q_2 \cdots q_\ell$$
$$p_2 \cdots p_k = uu' q_2 \cdots q_\ell.$$

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

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

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

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?

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*:

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

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:

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

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 $a \equiv a \mod n$.
- **Symmetric.** Let $a \equiv b \mod n$ 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 $b \equiv a \mod n$.
- **Transitive.** Let $a \equiv b \mod n$ and $b \equiv c \mod n$ 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 $a \equiv c \mod n$.

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,

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

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 $\{0, 1, \ldots, n-1\}$. 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 $\{0, 1, \ldots, n-1\}$. 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

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

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 $\{0, 1, \ldots, n-1\}$ 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 $\mathbb{Z}/n\mathbb{Z} = \{0, 1, \ldots, n-1\}$ 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]

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

$$b \equiv c.$$

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 $\mathbb{Z}/7\mathbb{Z}$:

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

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 Vector 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

[29] All congruences are mod p.

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 Vector 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 "psedorandom" 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.

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

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

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

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

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

$$a^k \equiv a^\ell$$

$$a^k \cdot a^{-\ell} \equiv a^\ell \cdot a^{-\ell}$$
$$a^{k-\ell} \equiv 1.$$

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

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

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

k	$3^k \bmod 11$
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 $\mathrm{ord}_{11}(3) = 5$, and the sequence repeats after every 5 steps.

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

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

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

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

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:

(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 $R^\times = R \backslash \{0\}$.

[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*.

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.

- 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

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

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

$$ax + ny = 1$$

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

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

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 $d \in \mathbb{Z}$ 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

$$a^{p-1} \equiv 1 \bmod p \text{ for all } a \in \mathbb{Z} \text{ such that } p \nmid a.$$

[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.

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.

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:

$$a^{p-1} \equiv 1 \bmod p \quad \text{for all integers } a \in \mathbb{Z} \text{ satisfying } p \nmid a.$$

[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.

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 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 $a \bmod mn$ to the pair of congruence classes $(a \bmod m, a \bmod n)$. Here is an example with $m = 2$ and $n = 3$:

$a \bmod 6$	$(a \bmod 2, a \bmod 3)$
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.

[39] The theorem was named by Leonard Dickson in 1929 and this notation has become standard.

In practical terms, this example tells us that each system of congruences $c \equiv a \bmod 2$ and $c \equiv b \bmod 3$ has a unique solution mod 6. For example, the final row of the table tells us that

$$\left\{ \begin{array}{rcl} c & \equiv & 1 \bmod 2 \\ c & \equiv & 2 \bmod 3 \end{array} \right\} \iff c \equiv 5 \bmod 6.$$

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 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:

$$\begin{array}{rccc} \varphi: & \mathbb{Z}/mn\mathbb{Z} & \to & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ & a \bmod mn & \mapsto & (a \bmod m, a \bmod n). \end{array}$$

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:

$$\begin{array}{rccc} \varphi^{-1}: & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} & \to & \mathbb{Z}/mn\mathbb{Z} \\ & (a \bmod m, b \bmod n) & \mapsto & ? \bmod mn. \end{array}$$

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 Vector 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:

$$\left\{ \begin{array}{rcl} c & \equiv & a \bmod m \\ c & \equiv & b \bmod n \end{array} \right\} \iff c \equiv any + bmx \bmod mn.$$

To prove this we only need to check that $\varphi(any + bmx) = (a, b)$. In other words, we need

[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.

to check that

$$any + bmx \equiv a \bmod m,$$
$$any + bmx \equiv b \bmod n.$$

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]

$$\left\{ \begin{array}{rcl} c & \equiv & a \bmod 2 \\ c & \equiv & b \bmod 3 \end{array} \right\} \iff c \equiv 3a - 2b \bmod 6.$$

We can use the same method to solve multiple simultaneous congruences by induction. Recall Sun Zu's system of congruences:

$$\left\{ \begin{array}{rcl} c & \equiv & 2 \bmod 3, \\ c & \equiv & 3 \bmod 5, \\ c & \equiv & 2 \bmod 7. \end{array} \right.$$

First we take $m = 3$ and $n = 5$ and observe that $3(2) + 5(-1) = 1$, so that

$$\left\{ \begin{array}{rcl} c & \equiv & 2 \bmod 3 \\ c & \equiv & 3 \bmod 5 \end{array} \right\} \iff c \equiv 2 \cdot 5(-1) + 3 \cdot 3(2) \equiv 8 \bmod 15.$$

Hence we have

$$\left\{ \begin{array}{rcl} c & \equiv & 2 \bmod 3 \\ c & \equiv & 3 \bmod 5 \\ c & \equiv & 2 \bmod 7 \end{array} \right\} \iff \left\{ \begin{array}{rcl} c & \equiv & 8 \bmod 15 \\ c & \equiv & 2 \bmod 7 \end{array} \right\}.$$

Then we take $m = 15$ and $n = 7$ and observe that $15(1) + 7(-2) = 1$, so that

$$\left\{ \begin{array}{rcl} c & \equiv & 8 \bmod 15 \\ c & \equiv & 2 \bmod 7 \end{array} \right\} \iff c \equiv 8 \cdot 7(-2) + 2 \cdot 15(1) \equiv 23 \bmod 105.$$

[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.

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.

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 $m, n \geqslant 1$:

$$\begin{array}{rccc} \varphi: & \mathbb{Z}/mn\mathbb{Z} & \to & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ & a \bmod mn & \mapsto & (a \bmod m, a \bmod n). \end{array}$$

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

$$\gcd(a, p^m) = 1 \quad \iff \quad p \nmid a.$$

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

$$\begin{aligned}
\phi(n) &= \phi(p_1^{n_1})\phi(p_2^{n_2}) \cdots \phi(p_k^{n_k}) \\
&= p_1^{n_1}\left(1 - \frac{1}{p_1}\right) p_2^{n_2}\left(1 - \frac{1}{p_2}\right) \cdots p_k^{n_k}\left(1 - \frac{1}{p_k}\right) \\
&= p_1^{n_1} p_2^{n_2} p_k^{n_k}\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\left(1 - \frac{1}{p_k}\right)
\end{aligned}$$

$$= n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right)$$
$$= n \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

where the product is taken over the prime divisors of n.

5 The Fundamental Theorem of Algebra

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:

$$\int x^n dx = \begin{cases} x^{n+1}/(n+1) & \text{if } n \neq -1 \\ \log|x| & \text{if } n = -1 \end{cases} \quad \text{and} \quad \int \frac{1}{x^2+1} dx = \arctan(x).$$

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:

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

[43] We will discuss this method in detail below.

[44] "Squaring the circle" refers to arctan and "squaring the hyperbola" refers to log.

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]

$$\begin{aligned}\frac{1}{x^4+4} &= \frac{1}{(x^2-2x+2)(x^2+2x+2)} \\ &= \frac{A+Bx}{x^2-2x+2} + \frac{C+Dx}{x^2+2x+2}.\end{aligned}$$

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 Vector 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 $x^2 - 2x + 2 \bmod 4x$:

$$\begin{array}{r|l} & \frac{1}{4}x - \frac{1}{2} \\ \hline 4x & x^2 - 2x + 2 \\ & \underline{-x^2} \\ & \quad -2x \\ & \quad \underline{2x} \\ & \qquad\quad 2 \end{array}$$

[45] This follows from a general theorem on partial fractions which we will prove below.

[46] We say that V is an $\mathbb{F}[x]$-module.

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:

$$\frac{1}{(x^2+2x+2)(x^2-2x+2)} = \frac{(2-x)/8 \cdot (x^2+2x+2)}{(x^2+2x+2)(x^2-2x+2)} + \frac{(2+x)/8 \cdot (x^2-2x+2)}{(x^2+2x+2)(x^2-2x+2)}$$
$$\frac{1}{x^4+4} = \frac{(2-x)/8}{x^2-2x+2} + \frac{(2+x)/8}{x^2+2x+2}$$

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 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.

Theorem on Partial Fractions

5.3 Equivalent Statements of the FTA

The original statement of the Fundamental Theorem of Algebra says 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. In this section we will still not prove that the FTA is true. Instead we will increase our understanding of what the FTA actually says by examining several equivalent 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 of Equivalent Forms of the $\mathbb{C}$FTA.

(1$\mathbb{C}$)⇒(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[47] 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, hence $f(x)$ splits over $\mathbb{C}$. ▫

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.

[47] We could also phrase this as a formal proof by induction.

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 of Equivalent Forms of the $\mathbb{R}$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 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.

(2$\mathbb{R}$)⇒(3$\mathbb{R}$): 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.

(3$\mathbb{R}$)⇒(1$\mathbb{R}$): 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 $\mathbb{R}$FTA and $\mathbb{C}$FTA

blah

5.4 Intermediate Value Theorem

5.5 Descartes and Euler on Quartic Equations

5.6 Waring's Method

5.7 Laplace's Proof of the FTA