Introduction to Groups, Rings and Fields

HT and TT 2011

H. A. Priestley

0. Familiar algebraic systems: review and a look ahead.

GRF is an ALGEBRA course, and specifically a course about algebraic structures. This introductory section revisits ideas met in the early part of Analysis I and in Linear Algebra I, to set the scene and provide motivation.

0.1 Familiar number systems

Consider the traditional number systems

$\mathbb{N} = \{0, 1, 2, \dots\}$	the natural numbers
$\mathbb{Z} = \{\, m - n \mid m, n \in \mathbb{N} \,\}$	the integers
$\mathbb{Q} = \{\, m/n \mid m, n \in \mathbb{Z},\ n \neq 0 \,\}$	the rational numbers
$\mathbb{R}$	the real numbers
$\mathbb{C}$	the complex numbers

for which we have

$$\mathbb{N} \subset \mathbb{Z} \subset Q \subset \mathbb{R} \subset \mathbb{C}.$$

These come equipped with the familiar arithmetic operations of sum and product.

The real numbers: Analysis I built on the real numbers. Right at the start of that course you were given a set of assumptions about $\mathbb{R}$, falling under three headings: (1) Algebraic properties (laws of arithmetic), (2) order properties, (3) Completeness Axiom; summarised as saying the real numbers form a *complete ordered field*.

(1) **The algebraic properties of $\mathbb{R}$** You were told in Analysis I:

Addition: for each pair of real numbers a and b there exists a unique real number $a + b$ such that

- $+$ is a commutative and associative operation;
- there exists in $\mathbb{R}$ a zero, 0, for addition: $a + 0 = 0 + a = a$ for all $a \in \mathbb{R}$;
- for each $a \in \mathbb{R}$ there exists an additive inverse $-a \in \mathbb{R}$ such that $a+(-a) = (-a)+a = 0$.

Multiplication: for each pair of real numbers a and b there exists a unique real number $a \cdot b$ such that

- $\cdot$ is a commutative and associative operation;
- there exists in $\mathbb{R}$ an identity, 1, for multiplication: $a \cdot 1 = 1 \cdot a = a$ for all $a \in \mathbb{R}$;
- for each $a \in \mathbb{R}$ with $a \neq 0$ there exists an additive inverse $a^{-1} \in \mathbb{R}$ such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Addition and multiplication together: for all $a, b, c \in \mathbb{R}$, we have the distributive law $a \cdot (b + c) = a \cdot b + a \cdot c$.

Avoiding collapse: we assume $0 \neq 1$. 0 for addition, 1 for multiplication

On the basis of these arithmetic laws *and no further assumptions* you were able to prove various other rules, such as the property $\forall a, b \in \mathbb{R}$, we have $a \cdot b = 0$ implies $a = 0$ or $b = 0$ (that is, there are no divisors of 0).

Q: why does "completeness axiom" matter? what's the consequence?

https://math.stackexchange.com/questions/2072709/what-is-the-role-of-associative-and-commutative-properties-in-mathematics-and-wh

(2) **Order properties**: $\mathbb{R}$ comes equipped with an order relation $<$ whereby each real number is classified uniquely as positive (> 0), negative (< 0), or zero. The properties ((P1)–(P3) in Analysis I handout) tell us how order interacts with $+$ and $\cdot$ and so provide rules for manipulating inequalities. (In addition they imply the trichotomy law: for all $a, b \in \mathbb{R}$, we have exactly one of $a > b$, $a < b$ or $a = b$. This allows us to think of $\mathbb{R}$ as a 'number line'.) Note that the modulus function draws on the order structure.

least upper bound if bounded above.

(3) **Completeness Axiom**: Concerns the order relation. Central to the development of real analysis.

The complex numbers, $\mathbb{C}$: In summary, $\mathbb{C}$ has arithmetic properties just the same as those for $\mathbb{R}$. There is no total order on $\mathbb{C}$ compatible with the arithmetic operations. Important good feature: polynomials (with real or complex coefficients) always have a full complement of roots in $\mathbb{C}$.

The rational numbers, $\mathbb{Q}$: Same arithmetic and order properties as for $\mathbb{R}$. Completeness Axiom fails. → e.g. $\{x \in \mathbb{Q} \mid x^2 < 2\}$, no least upper bound in $\mathbb{R}$, not in $\mathbb{Q}$.

The integers, $\mathbb{Z}$: Arithmetic behaves as for $\mathbb{Q}$ and $\mathbb{R}$ with the critical exception that not every non-zero integer has an inverse for multiplication: for example, there is no $n \in \mathbb{Z}$ such that $2 \cdot n = 1$.

The natural numbers, $\mathbb{N}$ are what number theory is all about. But $\mathbb{N}$'s arithmetic is defective: we can't in general perform either subtraction or division, so we shall usually work in $\mathbb{Z}$ when talking about such concepts as factorisation. $\mathbb{N}$, ordered by $0 < 1 < 2 \ldots$, also has an important ancillary role in the study of the rings of integers and polynomials (see Sections 3,4,5).

Restricting operations to subsets: We have $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. The sum and product on each of $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ are those they inherit from $\mathbb{R}$. For a non-empty subset S of $\mathbb{R}$, we say that S is *closed under* $+$ if $a, b \in S$ implies $a + b \in S$, and likewise for $\cdot$. In this terminology $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ are closed under $+$ and $\cdot$.

The operations $+$ and $\cdot$ on $\mathbb{R}$ are subject to a list of **axioms** (rules), as recalled in (1) above. Observe that these axioms are of two kinds:

($\forall$) those which have only universal quantifiers $\forall$;

($\exists$) those which contain an existential quantifier $\exists$ and so assert the existence of something.

Examples of axioms of type ($\forall$) for $\mathbb{R}$ are commutativity and associativity of both $+$ and $\cdot$, and the distributive law. For example, commutativity of $+$ says

$$(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})\, a + b = b + a.$$

An axiom of type ($\exists$) for $\mathbb{R}$ is that asserting that we have a zero element for addition:

$$(\exists 0 \in \mathbb{R})\big(\forall a \in \mathbb{R}\big)\, a + 0 = 0 + a = a.$$

Let S be any non-empty subset of $\mathbb{R}$ closed under $+$ and $\cdot$. Then any axiom of type ($\forall$) which holds in $(\mathbb{R}; +, \cdot)$ also holds in $(S; +, \cdot)$—it is *inherited*. By contrast, an axiom of type ($\exists$) may or may not hold on S: it depends whether or not the element or elements whose existence in $\mathbb{R}$ it guarantees actually belong to S.

Aside: do the number systems exist? [Informal comments in lecture.]

Q: in reality, encoding of $\mathbb{R}$ is limited by precision of computers, thus operation is actually over $\mathbb{Q}$ rather than $\mathbb{R}$., as a field?

0.2 An informal overview of algebraic structures. [Remarks in lecture.]

Just as geometric vectors provide motivation for the study of abstract vector spaces, so the number systems give prototypes for mathematical structures worthy of investigation.

$(\mathbb{R}; +, \cdot)$ and $(\mathbb{Q}; +, \cdot)$ serve as examples of **fields**,

$(\mathbb{Z}; +, \cdot)$ is an example of a **ring** which is not a field.

We may ask which other familiar structures come equipped with addition and multiplication operations sharing some or all of the properties we have encountered in the number systems. Here are some examples we might consider:

(1) $n \times n$ real matrices, $\mathrm{M}_n(\mathbb{R})$, or complex matrices, $\mathrm{M}_n(\mathbb{C})$, with the usual matrix addition and multiplication. Note that, except when $n = 1$, multiplication is not commutative. and that the no zero-divisor property fails: $AB = 0$ does not imply $A = 0$ or $B = 0$ in general.

(2) **Real-valued functions**. You know how to add and multiply pairs of real-valued functions on $\mathbb{R}$, and in Analysis II, you discovered that sums and products of continuous/differentiable/twice-differentiable/... functions are continuous/differentiable/twice-differentiable/.... All these sets of functions have good arithmetic properties (which you take for granted when using such things as the Algebra of Limits).

(3) **Polynomials with real coefficients**, $\mathbb{R}[x]$: You can think of these as real-valued functions; you do addition and multiplication of polynomials this way. Polynomials (except non-zero constants) do not have inverses for multiplication, but otherwise they behave rather well. In fact they share important features with the integers: the property of having no zero divisors and the process of 'long division'. We explore these ideas in Sections 4 and 5.

Two operations or one? The structures most familiar to you have two operations. But you have also met structures with a single operation, for example $\mathrm{Sym}(n)$, the permutations of an n-element set, with the operation of composition.

In the early part of the course we shall focus on structures with two (linked) operations. Most of our motivating examples are of this sort, and we shall not stray far from everyday mathematics. But we don't want to have long, unstructured, lists of axioms. So it will be expedient to modularise. Therefore before we categorise and study structures with two 'arithmetical' operations we should collect together some basic material on sets equipped with just one operation. Structures with one basic operation include groups.

1. Binary operations, and a first look at groups

1.1 Binary operations. Let S be a non-empty set. A map

(bop) $$\star\colon S \times S \to S, \qquad (a,b) \mapsto a \star b$$

is called a *binary operation* on S. So $\star$ takes 2 inputs a, b from S and produces a single output $a \star b \in S$. In this situation we may say that 'S is closed under $\star$'.

Aside: A *unary operation* on a non-empty set S is a map from S to S. Examples are

$$a \mapsto (-a) \text{ on } \mathbb{Z}, \qquad a \mapsto 2a \text{ on } \mathbb{C}.$$

Let $\star$ be a binary operation on a set S. We say

- $\star$ is *commutative* if, for all $a, b \in S$,
$$a \star b = b \star a.$$
- $\star$ is *associative* if, for all $a, b, c \in S$,
$$a \star (b \star c) = (a \star b) \star c$$
(note that (bop) ensures that each side of this equation makes sense). If $\star$ is associative we can *unambiguously* write $a \star b \star c$ to denote either of the iterated products. Very convenient.

Provable fact: Let $\star$ be an associative binary operation on a set S and let $x_1, \ldots, x_n \in S$. Then $x_1 \star x_2 \star \ldots \star x_n$ can be unambiguously defined.

In summary: we shall want to consider binary operations which are associative, but we do not restrict to those which are commutative.

intuitively why?

1.2 Examples of binary operations.

(1) Addition, $+$, is a commutative and associative binary operation on each of the following:
$$\mathbb{N}, \quad \mathbb{Z}, \quad \mathbb{Q}, \quad \mathbb{R}, \quad \mathbb{C}, \quad \mathrm{M}_{m,n}(\mathbb{R}) \ (m, n \geqslant 1), \quad \text{real polynomials}.$$
But $+$ is NOT a binary operation on the set $S = \{0, 1\}$: we have $1 \in S$ but $1 + 1 = 2 \notin S$.

(2) Multiplication, $\cdot$, is an associative and commutative binary operation on each of the following:
$$\mathbb{N}, \quad \mathbb{Z}, \quad \mathbb{Q}, \quad \mathbb{R}, \quad \mathbb{C}, \quad \text{real polynomials}.$$
Matrix multiplication is an associative binary operation on $\mathrm{M}_n(\mathbb{R})$; for $n \geqslant 2$ this is NOT commutative.
Here, and in (1) too, there is no reason to restrict to *real* matrices and polynomials. We could equally well have considered matrices with entries drawn from $\mathbb{C}$ or polynomials with complex coefficients.

(3) Assume that S is a non-empty subset of $\mathbb{R}$ closed under multiplication and such that $S \neq \{0\}$. Then $\cdot$ is a binary operation on $S \setminus \{0\}$. e.g. $S = \{2^k \mid k \in \mathbb{Z}\}$

(4) On a finite set S a binary operation may be specified by a table: for example:

$+$	0	1
0	0	1
1	1	0

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

You may recognise these tables as specifying binary arithmetic, that is, addition and multiplication mod (or modulo) 2.

Exercise example: Formulate addition and multiplication tables for 'arithmetic modulo 3' on the set $\{0,1,2\}$ and for 'arithmetic modulo 4' on $\{0,1,2,3\}$. [We'll look systematically at arithmetic modulo n later on.]

(5) **Exercise example:** By constructing appropriate tables give examples of (i) a binary operation on $\{0,1\}$ which is not commutative and (ii) a binary operation on $\{0,1\}$ which is not associative.

(6) **Scalar product on** $\mathbb{R}^3$ is given by

$$(a_1,a_2,a_3)\cdot(b_1,b_2,b_3)=a_1b_1+a_2b_2+a_3b_3.$$

Here the input is two vectors in $\mathbb{R}^3$ but the output is a real number, NOT another element of $\mathbb{R}^3$. Thus scalar product is NOT a binary operation on $\mathbb{R}^3$.

(7) **Vector product on** $\mathbb{R}^3$ is given by

$$(a_1,a_2,a_3)\wedge(b_1,b_2,b_3)=(a_2b_2-a_3b_3,a_3b_3-a_1b_1,a_1b_1-a_2b_2).$$

It defines a map from $\mathbb{R}^3\times\mathbb{R}^3$ to $\mathbb{R}^3$. With $\cdot$ standing for scalar product,

$$(\mathbf{a}\wedge\mathbf{b})\wedge c=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a},$$

whereas

$$\mathbf{a}\wedge(\mathbf{b}\wedge c)=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}.$$

From this we see easily that associativity fails when, for example,

$$\mathbf{a}=(1,0,0),\quad \mathbf{b}=\mathbf{c}=(1,1,1).$$

Furthermore, and this is very unlike 'ordinary' algebra,

$$\mathbf{a}\wedge(\mathbf{b}\wedge\mathbf{c})+\mathbf{b}\wedge(\mathbf{c}\wedge\mathbf{a})+\mathbf{c}\wedge(\mathbf{a}\wedge\mathbf{b})=0.$$

1.3 An important example: composition of maps. Let X be a set. Consider the set S of maps $f\colon X\to X$. For $f,g\in S$ define the *composition* $g\circ f$ by

$$\forall x\in X\quad (g\circ f)(x)=g(f(x)).$$

Then $\circ$ is a binary operation on S. We claim that $\circ$ is associative. Let $f,g,h\in S$. We require to show that, for each $x\in X$, we have $(h\circ(g\circ f))(x)=((h\circ g)\circ f)(x)$. But

$$(h\circ(g\circ f))(x)=h((g\circ f)(x))=h(g(f(x)))=(h\circ g)(f(x))=((h\circ g)\circ f)(x),$$

as required.

In particular composition is an associative binary operation on the set $\mathrm{Sym}(n)$ of permutations of $\{1,\dots,n\}$. For $n\geqslant 3$, $\circ$ on $\mathrm{Sym}(n)$ is not commutative. [**Exercise**: give an example.]

e.g. $n=7$, $f(x)=x+1 \pmod n$, $g(x)=2x \pmod n$, $(f\circ g)(5)=4\neq(g\circ f)(5)=5$

1.4 The definition of a group. Let G be a non-empty set and let $\star$ be a binary operation on G:

(bop) $$\star\colon G\times G\to G,\qquad (a,b)\mapsto a\star b.$$

Then $(G;\star)$ is a *group* if the following axioms are satisfied:

(G1) **associativity**: $a\star(b\star c)=(a\star b)\star c$ for all $a,b,c\in G$

G2) **identity element**: there exists $e\in G$ such that $a\star e=e\star a=a$ for all $a\in G$.

(G3) **inverses**: for any $a \in G$ there exists $a^{-1} \in G$ such that $a \star a^{-1} = a^{-1} \star a = e$.

If in addition the following holds

(G4) **commutativity**: $a \star b = b \star a$ for all $a, b \in G$

then $(G; \star)$ is called an *abelian group*, or simply a *commutative group*.

If the set G is finite, we define the *order* of G to be the number of elements in G, and denote it $|G|$. Otherwise we say that G has infinite order.

Remarks: Note that (bop) is an essential part of the definition. and that (G2) must precede (G3) because (G3) refers back to the element e.

Fact: if $(G; \star)$ is a group then the identity e is unique and the inverse of any a in G is uniquely determined by a. [Proof an exercise later.]

First examples of groups

(1) $(\mathbb{Z}; +)$, $(\mathbb{Q}; +)$, $(\mathbb{R}; +)$, $(\mathbb{C}; +)$ are abelian groups.

(2) Let V be a real vector space. Then, forgetting the scalar multiplication, $(V; +)$ is an abelian group. (See LAI notes, Section 2.)

(3) $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{C} \setminus \{0\}, \cdot)$ are abelian groups.

(4) The invertible $n \times n$ complex matrices form a group under matrix multiplication. This is an important group, denoted $\mathrm{GL}(n, \mathbb{C})$ or $\mathrm{GL}_n(\mathbb{C})$. For $n > 1$ this group is non-abelian.

$(\mathbb{R}; \cdot)$ is not group, because 0 doesn't have inverse.

$(\mathbb{Z} \setminus \{0\}; \cdot)$ is not a group, because most $a \in \mathbb{Z}$, $\frac{1}{a} \notin \mathbb{Z}$.

$\det(D) \neq 0$

Why matrix $(n \geq 2)$ is not commutative?

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}_{2\times 2} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}_{2\times 2} = \begin{bmatrix} a_{11}\cdot b_{11} + a_{12}\cdot b_{21} & a_{11}\cdot b_{12} + a_{12}\cdot b_{22} \end{bmatrix}$$

$\neq$ $\neq$

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{11}\cdot a_{11} + b_{12}\cdot a_{21} & b_{11}\cdot a_{12} + b_{12}\cdot a_{22} \end{bmatrix}$$

Our next example is of sufficient importance to deserve a theorem all to itself.

1.5 Theorem (bijections on a set X)**.** *Fix a non-empty set X and let*

$$\mathrm{Sym}(X) = \{\, f : X \to X \mid f \text{ is a bijection} \,\},$$

"1 to 1"

and let $\circ$ denote composition of maps.

(1) *$\circ$ is a binary operation on* $\mathrm{Sym}(X)$.

(2) $(\mathrm{Sym}(X); \circ)$ *is a group.* → but usually not an abelian group.

Proof. (1) To verify (bop) we need the fact that the composition of two bijections is a bijection. You met this in Introduction to Pure Mathematics.

For (2) we must verify (G1)–(G3).

(G1) Composition of maps is always associative (see 1.3), so in particular composition of bijective maps is associative. So (G1) holds.

(G2) It's up to us to exhibit an identity element, confirm that it belongs to $\mathrm{Sym}(X)$ and serves as e in (G2). Note that the identity map id is indeed a bijection and satisfies $\mathrm{id} \circ f = f \circ \mathrm{id} = f$.

(G3) Fix $f \in \mathrm{Sym}(X)$. Then, as a bijection, f has an inverse map f^{-1} which is also a bijection and satisfies $f \circ f^{-1} = \mathrm{id} = f^{-1} \circ f$. □

As a special case we have that $\mathrm{Sym}(n)$, the permutations of an n-element set, is a group under the usual composition of permutations.

2. Rings, fields and integral domains

Rings ≈ abelian group + 1 more op.

2.1 The definition of a ring. A structure $(R, +, \cdot)$ is a *ring* if R is a non-empty set and $+$ and $\cdot$ are binary operations:

??

$$+ : R \times R \to R, \qquad (a, b) \mapsto a + b$$
$$\cdot : R \times R \to R, \qquad (a, b) \mapsto a \cdot b$$

such that

Addition: $(R, +)$ is an abelian group, that is,

(A1) **associativity**: for all $a, b, c \in R$ we have $a + (b + c) = (a + b) + c$

(A2) **zero element**: there exists $0 \in R$ such that for all $a \in R$ we have $a + 0 = 0 + a = a$

(A3) **inverses**: for any $a \in R$ there exists $-a \in R$ such that $a + (-a) = (-a) + a = 0$

(A4) **commutativity**: for all $a, b \in R$ we have $a + b = b + a$

Multiplication:

NOTE: no "inverse" requirement on multiplication op.

(M1) **associativity**: for all $a, b, c \in R$ we have $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

Addition and multiplication together (distributive)

(D) for all $a, b, c \in R$,

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot b + b \cdot c.$$

interesting example of significance of commutativity

We sometimes say 'R is a ring', taken it as given that the ring operations are denoted $+$ and $\cdot$. As in ordinary arithmetic we shall frequently suppress $\cdot$ and write ab instead of $a \cdot b$.

We do NOT demand that multiplication in a ring be commutative. As a consequence we must postulate distributivity as 2 laws, since neither follows from the other in general.

Notation: subtraction and division We write $a - b$ as shorthand for $a + (-b)$ and a/b as shorthand for $a \cdot (1/b)$ when $1/b$ exists.

2.2 Special types of rings: definitions. Assume $(R; +, \cdot)$ is a ring. We say R is a *commutative ring* if its multiplication $\cdot$ is commutative, that is,

(M4) **Commutativity**: $a \cdot b = b \cdot a$ for all $a, b \in R$.

We say R is a *ring with* 1 (or *ring with identity*) if there exists an identity for multiplication, that is,

(M2) **identity element**: there exists $1 \in R$ such that for all $a \in R$ we have $a \cdot 1 = 1 \cdot a = a$.

no "inverse" axiom requirement on multiplication

2.3 Examples of rings.

Number systems

(1) All of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are commutative rings with identity (with the number 1 as the identity).

(2) $\mathbb{N}$ is NOT a ring for the usual addition and multiplication. These are binary operations and we do have a zero element, namely 0, so axiom (A2) holds. However (A3) (existence of additive inverses) fails: there is no $n \in \mathbb{N}$ for which $1 + n = 0$, for example. .

Q: what's the consequence? (tangible restriction?)

(3) Consider the set of even integers, denoted $2\mathbb{Z}$, with the usual addition and multiplication. This is a commutative ring without an identity. To verify that (M2) fails it is not sufficient just to say that the integer 1 does not belong to $2\mathbb{Z}$. Instead we argue as follows. Suppose for contradiction that there were an element $e \in 2Z$ such that $n \cdot e = n$ for all $n \in 2\mathbb{Z}$. In particular $2e = 2$, from which we deduce that e would have to be 1. Since $1 \notin 2\mathbb{Z}$ we have a contradiction.

Matrix rings Under the usual matrix addition and multiplication $\mathrm{M}_n(\mathbb{R})$ and $\mathrm{M}_n(\mathbb{C})$, are rings with 1, but are not commutative (unless $n = 1$). If we restrict to invertible matrices we no longer have a ring, because there is then no zero for addition.

Polynomials Polynomials, with real coefficients, form a commutative ring with identity under the usual addition and multiplication; we denote this by $\mathbb{R}[x]$.

Modular arithmetic Binary arithmetic on $\{0, 1\}$ (see 1.2(4)) gives us a 2-element commutative ring with identity. More generally we get a commutative ring with identity if we consider addition and multiplication mod n on $\{0, 1, \ldots, n-1\}$. Details in Section 6.

polynomials is an individual algebraic structure, because of ordering

$\mathbb{R}[x] = r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n$

different from Matrix

$M_n(\mathbb{R}) = \begin{bmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{bmatrix}$

We next record some basic facts about rings. To simplify the statements we have restricted attention to commutative rings.

2.4 Calculational rules for rings. *Assume that $(R; +, \cdot)$ is a commutative ring. Let $a, b, c \in R$.*

(i) *If $a + b = a + c$ then $b = c$.*

(ii) *If $a + a = a$ then $a = 0$.*

(iii) $-(-a) = a$.

(iv) $0a = 0$.

(v) $-(ab) = (-a)b = a(-b)$.

Assume in addition that R has an identity 1 Then

(vi) $(-1)a = -a$.

(vii) *If $a \in R$ has a multiplicative identity a^{-1} then $ab = 0$ implies $b = 0$.*

Proof. Very similar to the do-it-once-in-a-lifetime, axiom-chasing proofs you have already seen, or had to do, in the context of $\mathbb{R}$, or (in the case of (i)–(iv)), seen for addition in a vector space. □

2.5 Subrings and the Subring Test. [Compare with the Subspace Test from LAI]

Let $(R; +, \cdot)$ be a ring and let S be a non-empty subset of R. Then $(S; +, \cdot)$ is a *subring* of R if it is a ring with respect to the operations it inherits from R.

The Subring Test Let $(R; +, \cdot)$ be a ring and let $S \subseteq R$. Then $(S; +, \cdot)$ is a subring of R if (and only if) S is non-empty and the following hold:

(SR1) $a + b \in S$ for any $a, b \in S$;

(SR2) $a - b \in S$ for any $a, b \in S$;

(SR3) $ab \in S$ for any $a, b \in S$.

Here (SR1) and (SR3) are just the (bop) conditions we require for S. (A1), (M1) and (D) are inherited from R. The only ($\exists$) axioms are (A2) and (A3). Since $S \neq \varnothing$ we can pick some $c \in S$. Then, by (SR2), $0 = c - c \in S$. By (SR2) again, $-a = 0 - a \in S$.

Exercise: prove that if S is a subring of R then, with an obvious notation, necessarily $0_S = 0_R$ and $(-a)_S = (-a)_R$.

Examples

(1) $\mathbb{Z}$ and $\mathbb{Q}$ are subrings of $\mathbb{R}$;

(2) $\mathbb{R}$, regarded as numbers of the form $a + 0\mathrm{i}$ for $a \in \mathbb{R}$, is a subring of $\mathbb{C}$. interesting

(3) In the polynomial ring $\mathbb{R}[x]$, the polynomials of even degree form a subring but the polynomials of odd degree do NOT form a subring because $x \cdot x = x^2$ is not of odd degree.

(4) $n\mathbb{Z} := \{nk \mid k \in \mathbb{Z}\}$ is a subring of $\mathbb{Z}$ for any $n \in \mathbb{N}$.

More examples on Problem sheet 1.

2.6 New rings from old: products and functions.

Products of rings Let R_1 and R_2 be rings. Define binary operations on $R_1 \times R_2$ coordinatewise: for $r_1, r_1' \in R_1$ and $r_2, r_2' \in R_2$, let

$$(r_1, r_2) + (r_1', r_2') := (r_1 + r_1, r_2 + r_2'),$$
$$(r_1, r_2) \cdot (r_1', r_2') := (r_1 \cdot r_1', r_2 \cdot r_2').$$

good heuristics on how to build new structures

Consider the ring axioms in two groups:

- **Type ($\forall$) axioms** (no $\exists$ quantifier): associativity of addition (A1) and multiplication (M1); commutativity of addition (A4); distributivity (D). All of these hold because they hold in each coordinate and the operations are defined coordinatewise. Also if multiplication is commutative in R_1 and R_2 then so is multiplication in $R_1 \times r_2$
- **Type ($\exists$) axioms** (containing a $\exists$ quantifier): existence of zero (A2) and additive inverses (A3). Here we need to exhibit the required elements: $(0, 0)$ serves as zero and $(-r_1, -r_2)$ as the additive inverse of (r_1, r_2). Also. if R_1 and R_2 are rings with identity, so is $R_1 \times R_2$, with identity $(1, 1)$.

Example: $\mathbb{R}^2$ and more generally $\mathbb{R}^n$ for $n \geqslant 2$ is a commutative ring with 1 under the coordinatewise operations derived from $\mathbb{R}$.

Rings of functions: [Lecture example]

Let R be a ring and X a non-empty set. Denote by $R^X := \{ f \colon X \to R$, with sum, $f + g$, and product, $\cdot$, defined pointwise:

$$\forall x \in X \quad (f + g)(x) = f(x) + g(x),$$
$$\forall x \in X \quad (f \cdot g)(x) = f(x) \cdot g(x).$$

Then R^X is a ring, which is commutative (has a 1) if R is commutative (has a 1). Many instances of rings can be viewed as subrings of rings of the form R^X.

In the remainder of this section we consider some very special, but very important, classes of rings, of which our most familiar rings provide examples.

2.7 Fields and integral domains A commutative ring with identity, $(R, +, \cdot)$ satisfies all of the axioms (A1)–(A4), (that is, $(R, +)$ is an abelian group); (M1), (M2), (M4); (D) (distributivity). What's 'missing' here is

(M3) **inverses**: for all $a \in R$ with $a \neq 0$ there exists $1/a \in R$ (alternatively written a^{-1}) such that $a \cdot 1/a = 1/a \cdot a = 1$.

Definition of a field: A structure $(R, +, \cdot)$, where $+$ and $\cdot$ are binary operations on R is a *field* if (A1)–(A4), (M1)–(M4), and (D) hold, and $0 \neq 1$. This can be expressed in a more modular way as follows $(R, +, \cdot)$ is a field if

(A) $(R, +)$ is an abelian group;
(M) $(R \setminus \{0\}, \cdot)$ is an abelian group;
(D) the distributive laws hold.

polynomials are not field, but could be for modulus multiplication.

Exercise: convince yourself that the two formulations of the definition of a field are the same.

Examples of fields: The following are fields: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. The following commutative rings with identity fail to be fields: $\mathbb{Z}$, $\mathbb{K}[x]$; here K denotes $\mathbb{R}$, or $\mathbb{C}$, or any other field, and $K[x]$ the polynomials with coefficients in K.

Definitions: In a commutative ring we call an element $a \neq 0$ a *zero divisor* if there exists $b \neq 0$ such that $a \cdot b = 0$. A commutative ring with identity in which $0 \neq 1$ is an *integral domain* (ID) if it has no zero divisors.

Observation (cancellation property): in an integral domain $ab = ac$ and $a \neq 0$ implies $b = c$.

Examples of integral domains

(1) We claim that any field is an integral domain. To prove this, assume that $(R, +, \cdot)$ is a field and let $a, b \in R$ be such that $a \cdot b = 0$. If $a \neq 0$ then a^{-1} exists, and we have

$$0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b) = (a^{-1} \cdot a) \cdot b = 1 \cdot b = b.$$

and likewise with the roles of a and b reversed.

$(\mathbb{Z} \setminus \{0\}, \cdot)$ is not abelian group.

(2) $\mathbb{Z}$ and $K[x]$ are integral domains which fail to be fields.

(3) $\mathbb{R}^2$, with coordinatewise addition and multiplication (see 2.3) is a commutative ring with identity which fails to be an integral domain (and so is not a field):

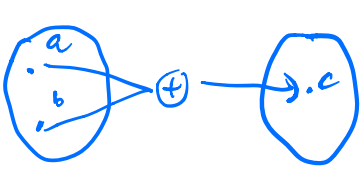
$$(0, 1) \cdot (1, 0) = (0, 0).$$

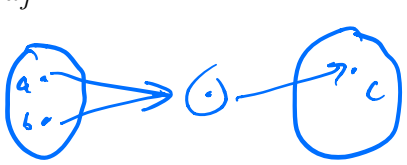
important example.

2.8 Theorem. *A finite integral domain is a field.*

Proof. Let R be a finite integral domain and list its distinct elements as $a_1, \ldots, a_n$. Let $b \in R$, $b \neq 0$. We have to show there exists $c \in R$ such that $bc = 1$. To do this we invoke the cancellation property of an ID: $ba_i = ba_j$ implies $a_i = a_j$. Therefore $\{ba_1, \ldots, ba_n\}$ contains n elements and is all of R. In particular, there exists j such that $ba_j = 1$. □

Finite ID





2.9 Units. Let R be a commutative ring with 1. Then $0 \neq u \in R$ is a *unit* if there exists $v \in R$ such that $uv = vu = 1$; we write as usual $v = u^{-1}$. Thus the units are those elements (necessarily non-zero) which have multiplicative inverses.

$\mathbb{Z}_n^*$ consists of units.

Examples

(1) In a field, every non-zero element is a unit.

(2) In $\mathbb{Z}$, the units are ± 1.

(3) In the ring of real or complex polynomials, or more generally the ring $K[x]$ of polynomials over a field K, the units are the non-zero constant polynomials.

↳ not for higher order

Summing up what we have about $\mathbb{Z}$ so far.:–

2.10 Theorem (the integers as a ring)**.**

(i) *$\mathbb{Z}$ is a commutative ring with* 1.

(ii) *$n \in \mathbb{Z}$ has a multiplicative inverse in $\mathbb{Z}$ (that is, n is a unit) if and only if $n = \pm 1$.*

(iii) *$\mathbb{Z}$ fails to be a field, but is an integral domain:*

(ID) *$ab = 0$ implies $a = 0$ or $b = 0$.*

3. Interlude: properties of the natural numbers

This short section draws attention to some facts about $\mathbb{N}$ which underpin our results on integers and polynomials in the next two sections.

3.1 Basic facts about the natural numbers (reminders). Recall that we have adopted the convention that

$$\mathbb{N} = \{0, 1, 2, \dots\}$$

(i.e. $\mathbb{N}$ includes 0). Thus $\mathbb{N}$ consists of the **non-negative** integers. When we refer to *positive* integers we mean those $n \in \mathbb{N}$ for which $n > 0$, or equivalently $n \neq 0$.

ASSUMED FACTS

- **The well-ordering property of** $\mathbb{N}$: every non-empty set of natural numbers contains a least element.
- **Unboundedness**: $\mathbb{N}$ is not bounded above.

Consequences of the well-ordering property

- There is no natural number n such that $0 < n < 1$.

  Proof: If there were, there would be a least such, m say. But then (exercise) $0 < m \cdot m < m < 1$, contradiction.

- **The Going Down Lemma** (strictly decreasing sequences in $\mathbb{N}$) *If (s_n) is a sequence of natural numbers such that*

  $$s_0 > s_1 > \ldots > s_n \ldots \geqslant 0$$

  then there exists $N \in \mathbb{N}$ such that $s_N = 0$.

  Proof Assume for contradiction that $s_n > 0$ for all $n \in \mathbb{N}$. Then

  $$S := \{\, s_n \mid n \in \mathbb{N} \,\}$$

  is a non-empty set of natural numbers, and so, by the well-ordering property of $\mathbb{N}$, the set S has a least member, s_m say. But then $s_{m+1} \in S$ and $s_{m+1} < s_m$, and we have the required contradiction. $\square$

4. Integers

We now look at what we can say about division in $\mathbb{Z}$. [Lecture example: how does school 'long division' work?]

4.1 Theorem (the division algorithm for $\mathbb{N}$). *Let $a, b \in \mathbb{N}$ with $b \neq 0$. Then there exist unique natural numbers q and r such that $a = qb + r$ and $0 \leqslant r < b$.*

Proof. Assume $a < b$. Then $a = 0b + a$ so $q = 0$ and $r = a$ do the job.

Now assume $a \geqslant b > 0$. Consider

$$S = \{\, m \in \mathbb{N} \mid mb - a > 0 \,\}.$$

Then $S \neq \varnothing$ because $\mathbb{N}$ (as a subset of $\mathbb{R}$) is not bounded above. By the Well-ordering Property there exists a least element, k say, in S. Note $0 \notin S$ and so $k \geqslant 1$. Let $q = k - 1$ and $r = a - qb$. Then $q \in \mathbb{N} \setminus S$ (by leastness of k) so $r \geqslant 0$. Also $q + 1 = k \in S$ so that $(q + 1)b > a$. Hence $b > r$. □

Corollary (the division algorithm for $\mathbb{Z}$). *Let $a, b \in \mathbb{Z}$ with $b \neq 0$. Then there exist unique integers q and r such that $a = qb + r$ and $0 \leqslant r < |b|$.*

Proof. [omitted from lecture]

(i) $a \geqslant 0$, $b > 0$. Apply the theorem.

(ii) $a \geqslant 0$, $b < 0$. Then $-b > 0$ and by the theorem there exist $q', r' \in \mathbb{N}$ with $a = q'(-b) + r'$ and $0 \leqslant r' < -b = |b|$. Let $q = -q'$ and $r = r'$.

(iii) $a < 0$, $b > 0$. By the theorem, there exist $q'', r'' \in \mathbb{Z}$ such that $-a = q''b + r''$ and $0 \leqslant r'' < b$. If $r'' = 0$, take $q = -q''$ and $r = 0$. If $r'' > 0$, take $r = b - r''$ and note that $0 < r < b$. Also let $q = (-q'' - 1)$ and note that $a = -q''b - r'' = -q''b + (r - b) = qb + r$, as we require.

(iv) $a < 0$, $b < 0$. Apply the result from case (iii) with b replaced by $-b$, in the same way as in case (ii). □

Uniqueness of q and r in the Division Algorithm

Assume there is another pair of integers $\tilde{q}$ and $\tilde{r}$ such that $a = \tilde{q}b + \tilde{r}$ with $0 \leqslant \tilde{r} < |b|$. Then, from $a = qb + r = \tilde{q}b + \tilde{r}$, we deduce that $(q - \tilde{q})b = \tilde{r} - r$. So $\tilde{r} = r$ would force $q = \tilde{q}$ (note that property (ID) is used here).

So suppose for a contradiction that $\tilde{r} \neq r$. Suppose $b > 0$ (the case $b < 0$ works similarly). Without loss of generality, $\tilde{r} > r$ and hence $q > \tilde{q}$. But then

$$b > \tilde{r} \geqslant \tilde{r} - r = (q - \tilde{q})b > 0.$$

But this would give $0 < (q - \tilde{q}) < 1$, with $(q - \tilde{q}) \in \mathbb{N}$ and this is impossible. □

We next give an application of the Division Algorithm which will be important later on, specifically in our study of groups.

4.2 Theorem (subrings of $\mathbb{Z}$)**.**

(1) *Let S be a non-empty subset of $\mathbb{Z}$ closed under addition and subtraction. Then there exists $n \in \mathbb{N}$ such that*

$$S = n\mathbb{Z}; = \{nk \mid k \in \mathbb{Z}\}.$$

(2) *The subrings of $\mathbb{Z}$ are exactly the sets of the form $n\mathbb{Z}$ for $n \in \mathbb{Z}$.*

Proof. (1) Since $S \neq \varnothing$ there exists some $m \in S$. Then $0 = m - m \in S$.

If $S = \{0\}$, then $S = 0\mathbb{Z}$.

Now assume $S \neq \{0\}$. Then there exists some $m \neq 0$ in S. Also $-m = 0 - m \in S$. Therefore S contains a positive integer. By the Well-order Property of $\mathbb{N}$ there exists a least positive member of S; call it n. We claim $S = n\mathbb{Z}$.

We can prove by induction that $n \in S$ implies $n\mathbb{Z} \subseteq S$ (see Problem sheet 2, question 2). Now let $a \in S$. By the Division Algorithm there exist $q, r \in \mathbb{Z}$ such that $a = qn + r$, with $0 \leqslant r < n$. But then $r = a - qn \in S$. By leastness of n we must have $r = 0$. Therefore $a \in n\mathbb{Z}$.

(2) We noted earlier that every set $n\mathbb{Z}$ is a subring of $\mathbb{Z}$. (1) supplies the converse. □

4.3 Divisors. Given integers a and b, with $b \neq 0$, we say that b is a *factor* (or a *divisor*) of a if there exists an integer q such that $a = qb$. If b is a divisor of a we write $b|a$.

Note: Every $b \in \mathbb{Z}$ is such that $b|0$. But $0|b$ never holds (by definition).

The next lemma records elementary but useful properties of divisors.

Divisors Lemma for $\mathbb{Z}$. *Let a, b and c be integers.*

(i) *Assume $a \neq 0$. Then $a|1$ implies $a = \pm 1$.*

(ii) *Assume $a, b \neq 0$. Then $a|b$ and $b|a$ implies $a = \pm b$.*

(iii) *Assume $a \neq 0$. Then $a|b$ and $a|c$ implies $a|(mb + nc)$ for any integers m and n.*

Proof. (i) $a|1$ is just a way of saying that a is a unit, so (i) follows from Theorem 2.10(ii). We leave (ii) and (iii) as exercises. $\square$

4.4 The highest common factor of two integers: definitions and uniqueness. Let a and b be two non-zero integers. Assume that c is a positive integer such that

(HCF1) $c|a$ and $c|b$;

(HCF2) for all integers d such that $d|a$ and $d|b$, then $d|c$.

Then c is called the *highest common factor* (or the *greatest common divisor*) of a and b, and denoted $\mathrm{hcf}(a,b)$ (or $\gcd(a,b)$). It is convenient, by convention, to extend the definition by taking, for any integer a,

$$\mathrm{hcf}(a,0) = \mathrm{hcf}(0,a) = a.$$

Notes: Observe that if c satisfies (HCF1) and (HCF2) above then so does $(-1)c = -c$. By requiring the hcf of a and b to be positive we ensure it is unique, provided it exists (Exercise: check uniqueness, using the Divisors Lemma from 4.3). So we really can talk about *the* hcf rather than *an* hcf of given integers, provided there is an hcf at all.

Observe that for any $a, b \in \mathbb{Z} \setminus \{0\}$,

$$\mathrm{hcf}(a,b) = \mathrm{hcf}(a,|b|) = \mathrm{hcf}(|a|,b) = \mathrm{hcf}(|a|,|b|).$$

We can therefore restrict attention to positive integers.

But do hcf's exist? How do we find them if they do? The following lemma is a key ingredient in proving the existence of hcf's.

4.5 Invariance Lemma (for hcf's of positive integers) *Let a and b be positive integers. Write*

$$a = qb + r \quad \textit{with } q, r \in \mathbb{N} \textit{ and } 0 \leqslant r < b.$$

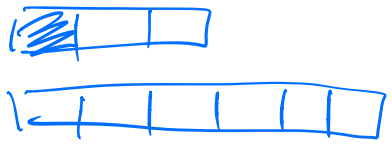
Assume $\mathrm{hcf}(b,r)$ *exists. Then* $\mathrm{hcf}(a,b)$ *exists and equals* $\mathrm{hcf}(b,r)$.

Proof. Let $c := \mathrm{hcf}(b,r)$. We shall show that c satisfies conditions (HCF1) and (HCF2) in the definition of $\mathrm{hcf}(a,b)$. First note that $c|b$ and $c|r$. Hence $c|qb + 1r = a$ by Divisors Lemma, part (iii). Now assume that $d|a$ and $d|b$. From property (iii) in the Divisors Lemma we get $d|r = a - qb$. Hence $d|c$ by condition (HCF2) applied to the pair b and r. □

In the process of moving from the pair (a,b) to the pair (b,r)

- the hcf is left unchanged (it is called the **'invariant'**);
- the second component (called the **'variant'**) changes from one positive integer to another *whose magnitude is strictly smaller.*

[:**Lecture example** : computing $\mathrm{hcf}(72,15)$.]



4.6 Theorem (Euclid's algorithm, from Euclid's *Elements*, Book 7, Propositions 1 and 2)**.**

(1) *Let a and b be positive integers, then* $\mathrm{hcf}(a,b)$ *exists;*

(2) (*the* **hcf formula**) *there exist integers m and n such that*

$$\mathrm{hcf}(a,b) = ma + nb.$$

Proof. Without loss of generality we may assume that $0 < b \leqslant a$. In what follows, all quantities appearing are integers and at each step the division algorithm is invoked. We can write

$$\begin{aligned}
a &= q_0 b + r_0 && \text{where } 0 < r_0 < b \\
b &= q_1 r_0 + r_1 && \text{where } 0 < r_1 < r_0 \\
r_0 &= q_2 r_1 + r_2 && \text{where } 0 < r_2 < r_1 \\
&\dots \quad \dots && \dots \\
r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1} && \text{where } 0 < r_{k-1} < r_{k-2} \\
r_{k-2} &= q_k r_{k-1} + r_k && \text{where } 0 = r_k
\end{aligned}$$

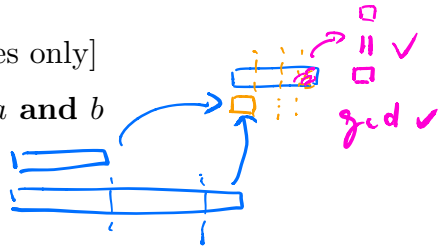
where we stop as soon as we reach a zero remainder. This must happen in a finite number of steps, by the Going Down Lemma (see Section 3). The Invariance Lemma 4.5 tells us that

$$\mathrm{hcf}(a,b) = \mathrm{hcf}(b,r_0) = \ldots = \mathrm{hcf}(r_{k-2}, r_{k-1}) = r_{k-1}.$$

To obtain the hcf formula we retrace our steps: write r_{k-1} in terms of r_{k-2}, then substitute the resulting formula for r_{k-1} into the equation for r_{k-3} and so on until a formula in a and b of the required form is obtained. More formally: argue by induction on k. □

Exercise example. (corrected) Find the highest common factor of $a = 38793$ and $b = 89531$ and express it as $ma + nb$. [Ans: $\mathrm{hcf}(a,b) = 1 = 13469 \cdot 38793 + (-5836) \cdot 89531$ (note m and n are not unique—recall Problem sheet 2, Question 4).

4.7 Implementing Euclid's Algorithm in practice. [web notes only]



Euclid's Algorithm to compute the hcf of positive integers a and b

Input a `and` b.

Step 1. `Let` $x := \min\{a,b\}$ `and` $y := \max\{a,b\}$.

Step 2. `Find integers` q `and` $0 \leqslant r < x$ `such that` $y = qx + r$.

Step 3. `IF` $r = 0$, `THEN output` y `ELSE put` $x := b$ `and` $y := r$ `and GOTO Step 2.`

Extending Euclid's algorithm to incorporate obtaining the hcf formula. By refining the strategy in Theorem 4.6 by adding some extra book-keeping we can obtain a simple 'one-pass' procedure for calculating both $\mathrm{hcf}(a,b)$ and possible values for m, n in $\mathrm{hcf}(a,b) = ma + nb$. Notation: When, from the division algorithm, we have $a = qb + r$ with $0 \leqslant r < b$, let us write $q := q(a,b)$.

Input: positive integers a and b.

Generate natural numbers r_i, q_i, m_i and n_i as follows:

Step 1: `Set`

$$\begin{aligned}
r_0 &:= \max\{a,b\}, \quad m_0 := 1, \quad n_0 := 0, \\
r_1 &:= \min\{a,b\}, \quad n_1 := 0, \quad n_1 := 1, \\
i &:= 1.
\end{aligned}$$

Step 2: `WHILE` $r_i \neq 0$ `REPEAT`

$q_{i+1} := q(r_{i-1}, r_i);$

$r_{i+1} := r_{i-1} - q_{i+1}r_i, \quad m_{i+1} := m_{i-1} - q_{i+1}m_i, \quad n_{i+1} := n_{i-1} - q_{i+1}n_i;$

`add 1 to` i.

Step 3: `Set`

$$k := i - 1;$$
$$c := r_{k-1},$$
$$m := m_{k-1},$$
$$n := n_{k-1}.$$

[A programming detail: note that in Step 2, at the end of each pass, i is incremented to $i+1$; it is only at the start of the *next* pass through the while loop that the termination condition $r_{i+1} = 0$ is picked up.]

The following hold

(i) the process reaches Step 3, and so has terminated, after finitely many steps;

(ii) $c = \text{hcf}(a, b)$;

(iii) $c = ma + nb$.

(i) and (ii) have already been proved in Theorem 4.6. For (iii), note that what we need is $r_{k-1} = m_{k-1}a + n_{k-1}b$ and this we prove by induction. By construction, if $r_j = m_j a + n_j b$ holds at the start of the loop when $j = i$ then it also holds, for $j = i+1$, after the loop has been executed once. Also $r_j = m_j a + n_j b$ holds for $j = 0, 1$ by definition.

Euclidian Algorithm :

fast way to compute $\gcd(a,n) = ?$

Extended Euclidian Algorithm :

given $\gcd(a,n) = r$, calculate s & t, s.t. $s \cdot n + t \cdot a = r$

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

↳ why useful? EEA used to calculate inverse modulo n :

$a^{-1} \equiv ? \mod n$

$a \cdot a^{-1} \equiv 1 \mod n$

$\therefore \gcd(a,n) = 1$

$\rightarrow s \cdot n + t \cdot a = 1$

$\downarrow \mod n$

$(s \cdot n) \bmod n + (t \cdot a) \bmod n \equiv 1 \mod n$

$t \cdot a \equiv 1 \cdot \mod n$

$\Rightarrow t = a^{-1}$

Tabulating the sums: Form a table, one row at a time, as follows. The first row has entries -, $r_0 = \max\{a, b\}$, $m_0 = 1$, $n_0 = 0$ and the second has entries -, $r_1 = \min\{a, b\}$, $m_1 = 0$ and $n_1 = 1$. Then, so long as $r_i \neq 0$, the $(i+1)$th row is filled in from left to right by first calculating q_{i+1} by the division algorithm and then r_{i+1}, m_{i+1} and n_{i+1} are calculated as per Step 2. ;

Example. To find the hcf of 134 and 28.

Q	R	M	N
-	134	1	0
-	28	0	1
4	22	1	-4
1	6	-1	5
3	4	4	-19
1	2	-5	24
2	0		

We have

$$2 = (-5) \cdot 134 + 24 \cdot 28.$$

Exercise. Rework the exercise from 4.6 tabulating the calculations.

4.8 Coprime integers: definition and a useful fact. Two non-zero integers a and b are said to be *coprime* if $\text{hcf}(a, b) = 1$. In this case, by the hcf formula from 4.6, there exist integers m and n such that $1 = ma + nb$. We claim the converse is true too. Assume that $1 = ma + nb$ for some integers m and n. Let d be a positive integer such that $d|a$ and $d|b$. Then $d|ma + nb = 1$ by property (iii) in the Divisors Lemma. By property (ii) in the same lemma, $d = 1$. Hence $\text{hcf}(a, b) = 1$, as claimed.

4.9 Definition: prime numbers. Let $n \in \mathbb{N}$ with $n \neq 0, 1$. Then n is called *prime* or *a prime* if n has exactly two positive divisors (itself and 1). According to our definition, 1 is NOT a prime!

4.10 Proposition (division by primes)**.** *Assume that a and b are integers and that p is a prime such that $p|ab$. Then $p|a$ or $p|b$.*

Proof. If $p|a$ there is nothing to prove. On the other hand, if p does not divide a, then $\text{hcf}(a, p) = 1$ and we can find $m, n \in \mathbb{Z}$ such that $1 = ma + np$. Now $b = mab + bmn$. By the Divisors Lemma (iii) we get $p|b$. □

4.11 The Fundamental Theorem of Arithmetic. *Every positive integer $a > 1$ can be written as a product of prime factors, and this factorisation is unique up to the order in which the factors are written. Equivalently, for every positive integer $a > 1$ there exist a unique positive integer k, unique primes $p_1, \ldots, p_k$ and unique positive integers $\beta_1, \ldots, \beta_k$ such that*

$$a = p_1^{\beta_1} \cdots p_k^{\beta_k}.$$

We can now see the reason for disqualifying 1 from being a prime:

$$1 = 1 \cdot 1 = \cdot 1 \cdot 1 \cdot 1 \cdot \ldots$$

so we don't get a *unique* factorisation.

The proof of the Fundamental Theorem of Arithmetic is not examinable in Mods. You can find a proof in *A guide to Abstract Algebra*, C. Whitehead (Macmillan Mathematical Guides). The result will be revisited in Part A Algebra next year, and treated in a more general context.

Remark: There is an obvious extension of the theorem to the factorisation of an integer $\notin \{0, \pm 1\}$, in terms of products of elements of the form $\pm p$, where p is prime. Now we only have uniqueness 'up to order of factors and $\pm$s'.

4.12 A classic theorem. *There are infinitely many primes.* [web notes only]

Proof. We assume for a contradiction that there are only finitely many primes $p_1, p_2, \ldots, p_n$. Let $a = p_1 p_2 \cdots p_n + 1$. Note that since a is bigger than every p_i it cannot itself be a prime. As a is not a prime, it must have a divisor other than a and 1. In particular it must have a prime divisor. Hence there exists $j \in \{1, \ldots, n\}$ such that $p_j | a$. We can also see that $p_j | p_1 p_2 \cdots p_n$. Therefore p_j divides $a - p_1 p_2 \cdots p_n = 1$. We have a contradiction, since 1 is the only positive divisor of 1 and 1 is not prime. $\square$

5. Polynomials

Our aim in this section is to reveal that there are close similarities between the ring $\mathbb{Z}$ of integers and the ring $K[x]$ of polynomials with coefficients in a field K.

5.1 Polynomials. [with a light touch on the formalities]

Examples The following are (real) polynomials:

$$x^2,$$
$$2 + \pi x + 7x^2,$$
$$\sqrt{2} + 3x - 5x^3 = \sqrt{2} + 3x + 0.x^2 - 5x^3,$$
$$42$$

but

$$1 + x + x^2 + \ldots + x^r + \ldots = \sum_{r=0}^{\infty} x^r$$

is NOT a polynomial.

Definition: A (*real*) *polynomial* in the variable x is an expression

$$f := a_0 + a_1 x + \cdots + a_n x^n$$

bounded exponent

where n is a non-negative integer and $a_i \in \mathbb{R}$ for $i = 0, \ldots, n$.

We shall also allow ourselves alternatively to write polynomials 'top-down':

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

Note that we are dealing with a *finite* number of terms $a_i x^i$ (compare with a power series $\sum_{n=0}^{\infty} a_r x^r$ which need not be a finite sum).

The set of all polynomials in the variable x and coefficients in $\mathbb{R}$ is denoted $\mathbb{R}[x]$. We shall, of course, want to think of polynomials as (defining) functions on $\mathbb{R}$.

In the expression $f = a_0 + a_1 x + \cdots + a_n x^n$ some or all of the coefficients a_i could be zero. If they are all zero we have the *zero polynomial*, denoted 0. A non-zero polynomial f can ~~in~~ be written as $f = a_0 + a_1 x + \cdots + a_n x^n$, where $a_n \neq 0$, and we call n the *degree* of f and denote it by $\deg f$ (∂f in some books). Note that the degree of the zero polynomial is not defined. A polynomial of degree zero is usually called a *constant*.

5.2 Rules for manipulating polynomials. [Justifications non-examinable] Let

$$f = a_0 + a_1 x + \cdots + a_n x^n \text{ and } g = b_0 + b_1 x + \cdots + b_m x^m$$

be two polynomials.

- **Equality**: the polynomials f and g are the same if and only if either both are 0 or $\deg f = \deg g$ ($= k$, say) and $a_i = b_i$ for all $i \leqslant k$.

  As usual we write $f(a)$ for 'f evaluated at $a \in \mathbb{R}$', that is,

  $$f(a) := a_0 + a_1 a + \cdots + a_n a^n.$$

  FACT: f and g are equal as polynomials if and only if $f(a) = g(a)$ for all $a \in \mathbb{R}$. [Outline proof: consider in turn the coefficients of $x^0, x^1, x^2, \ldots$.]

- **Addition**: let $k = \max\{n, m\}$ and define $a_i = 0$ for any i such that $n < i \leqslant k$ and $b_i = 0$ for any i such that $m < i \leqslant k$. Then the sum $f + g$ is defined to be the polynomial $c_0 + c_1x + \cdots + c_kx^k$ where $c_i = a_i + b_i$ for $i = 0, \ldots, k = \max\{n, m\}$.
- **Additive inverse**: for each polynomial $f = a_0 + a_1x + \ldots a_nx^n$ there exists a unique polynomial $(-f) := (-a_0) + (-a_1)x + \cdots + (-a_n)x^n$ with the property that $f + (-f) = 0$.
- **Multiplication**: the product $f \cdot g$ is defined to be the polynomial $c_0 + c_1x + \cdots + c_kx^k$ where $c_j = \sum_{i=0}^{j} a_i b_{j-i}$ for all $j = 0, \ldots, k = n + m$.

FACTS about degrees For non-zero $f, g \in \mathbb{R}[x]$,

- if $f + g \neq 0$ then $\deg(f + g) \leqslant \max\{\deg f, \deg g\}$;
- $\deg(f \cdot g) = \deg f + \deg g$ (in particular $f \cdot g \neq 0$).

5.3 Theorem ($\mathbb{R}[x]$ as a ring)**.**

(i) *$\mathbb{R}[x]$ is a commutative ring with 1, with multiplicative identity the constant polynomial 1.*

(ii) *$f \in \mathbb{R}[x]$ has a multiplicative inverse (that is, f is a unit) if and only if f is a non-zero constant.*

(iii) *$\mathbb{R}[x]$ is not a field but is an integral domain, that is,* (i) *holds and*

(ID) $$f \cdot g = 0 \text{ implies } f = 0 \text{ or } g = 0.$$

Proof. (i) follows from the rules for addition and multiplication, together with the ring properties that hold in $\mathbb{R}$.

For (ii) and (iii) we appeal to the fact that $\deg(f \cdot g) = \deg f + \deg g$ for non-zero polynomials f and g.

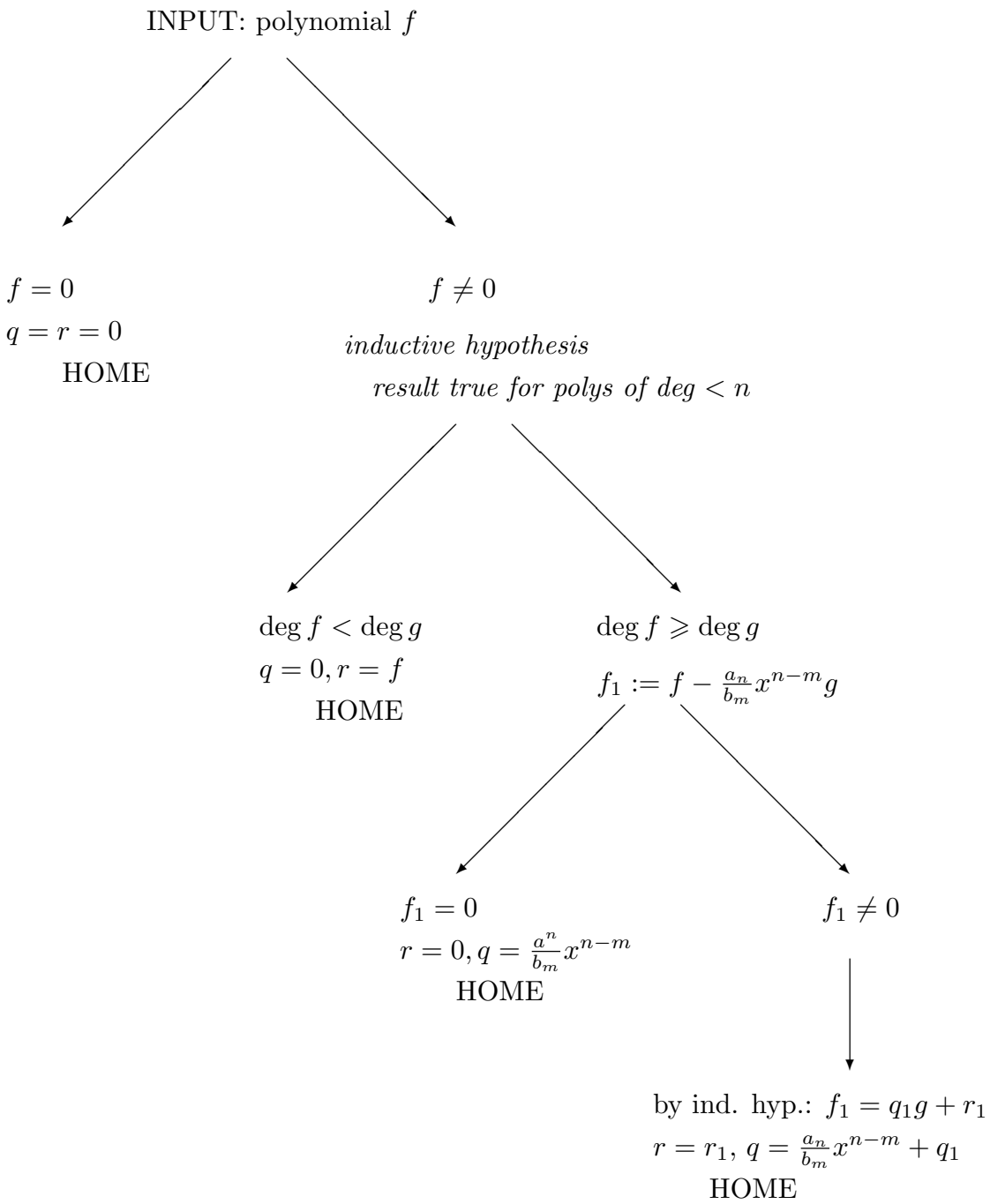
Remark: All the rules for addition and multiplication would work equally well if the coefficients in our polynomials were drawn from $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{C}$, and the above theorem would work too. Indeed, we could replace $\mathbb{R}$ by any integral domain, R say. When it comes to division, though, we shall need to be able to divide coefficients, and then we need R to be a *field*.

Lecture example: Long division of polynomials.

5.4 Theorem. (the division algorithm for polynomials)**.** *Let f and g be polynomials in $\mathbb{R}[x]$, with $g \neq 0$. Then there exist unique polynomials q and r in $\mathbb{R}[x]$ such that*

$$f = q \cdot g + r \text{ where either } r = 0 \text{ or } \deg r < \deg g.$$

(*More generally the above statement holds when $\mathbb{R}$ is replaced by any field K.*)



Proof. **Existence of q and r:** If $f = 0$ we can take $q = r = 0$ so assume $f \neq 0$. We work by strong induction on the degree of f: it suffices to show that if the result is true for all (non-zero) polynomials of degree strictly smaller than $\deg f$, then it is true for f too. [Strong induction was discussed in *Introduction to Pure Mathematics*.]

So let

$$f = a_0 + a_1 x + \cdots + a_n x^n \quad \text{with } a_n \neq 0,$$
$$g = b_0 + b_1 x + \cdots + b_m x^m \quad \text{with } b_m \neq 0$$

(so $\deg f = n$ and $\deg g = m$). If $m > n$ then let $q = 0$ and $r = f$ and we're home. Now assume $m \leqslant n$.

So assume that the required result is true when f is replaced by a polynomial of degrees strictly less than n. Let

$$f_1 := f - \frac{a_n}{b_m} x^{n-m} \cdot g;$$

this is a well-defined polynomial since $b_m \neq 0$ by assumption, and either $f_1 = 0$ or $\deg f_1 \leqslant n-1 < k = \deg f$. If $f_1 = 0$ then

$$f = \frac{a_n}{b_m} x^{n-m} \cdot g$$

and so $q = (a_n/b_m)x^{n-m}$ and $r = 0$ satisfy the requirements of the statement. If $f_1 \neq 0$, then we can use the strong induction hypothesis on f_1: there exist polynomials q_1 and r_1 such that

$$f_1 = q_1 g + r_1 \qquad \text{with } r_1 = 0 \text{ or } \deg r_1 < \deg g.$$

We deduce that

$$f_1 - \frac{a_n}{b_m} x^{n-m} \cdot g = q_1 g + r_1,$$

and hence

$$f = \left(\frac{a_n}{b_m} x^{n-m} + q_1 \right) g + r_1.$$

We can then take $q = (a_n/b_m)x^{n-m} + q_1$ and $r = r_1$; this satisfies $r = 0$ or $\deg r < \deg g$.

Uniqueness:

This works in the expected manner. Assume there are two pairs of polynomials q, r and $\tilde{q}$, $\tilde{r}$ such that $f = q \cdot g + r$ where either $r = 0$ or $\deg r < \deg g$, and $f = \tilde{q} \cdot g + \tilde{r}$ where either $\tilde{r} = 0$ or $\deg \tilde{r} < \deg g$. Hence we get $(q - \tilde{q})g = \tilde{r} - r$. Then, either $q = \tilde{q}$ and so $\tilde{r} = r$; or $\deg(q - \tilde{q}) \geqslant 0$ and so $\deg[(q - \tilde{q})g] \geqslant \deg g$ by the facts about degrees (see 5.1), while $\tilde{r} - r$ is either 0 or has degree not greater than $\deg g - 1$, which cannot occur. □

5.5 Divisors of polynomials. Notice that everything we did for integers once we had established the division algorithm came directly from that, in particular all the theory of hcf's. So, with minor adaptations, corresponding results hold for polynomials too. For definiteness we work with $\mathbb{R}[x]$, but we could equally well substitute polynomials with coefficients drawn from any field K, for example $K = \mathbb{Q}$ or $\mathbb{C}$.

Definition We say that a non-zero polynomial g divides (or is a factor of) the polynomial f if there exists a polynomial q such that $f = qg$, that is, if $r = 0$ in the division algorithm. Notation: $g|f$.

Divisors Lemma for polynomials *Let f, g and h be non-zero polynomials in $\mathbb{R}[x]$.*

(i) *$f|1$ if and only if f is a non-zero constant.*

(ii) *Assume $f|g$ and $g|f$. Then there exists $\alpha \in \mathbb{R} \setminus \{0\}$ such that $f = \alpha g$.*

(iii) *Assume $h|f$ and $h|g$, then $h|ma + nb$ for all m, $n \in \mathbb{R}[x]$.*

Proof. The proof parallels that of the Divisors Lemma for integers, drawing on Theorem 5.3(ii) and (iii) and the division algorithm for polynomials. The details are left as an exercise. □

5.6 Highest common factors of polynomials. Let p and q be non-zero polynomials in $\mathbb{R}[x]$. A highest common factor of f and g is a non-zero polynomial h such that

(HCF1) h divides both f and g;

(HCF2) for all polynomials k that divide both f and g, we have that $k|h$.

Note: Non-uniqueness does arise here, but it stems solely from the units in the ring $\mathbb{R}[x]$: if non-zero polynomials h and $\tilde{h}$ are both hcf's of f and g, then there exists a non-zero real number α such that $\tilde{h} = \alpha h$ for any non-zero real number α. Nevertheless we shall allow ourselves to write

$\text{hcf}(f, g)$, meaning by this any one of the allowable choices. When two non-zero polynomials f, and g are *coprime*, that is, have no non-constant common factor, it is customary to take $\text{hcf}(f, g) = 1$.

[Lecture example] Consider

$$f = 2x^3 - 5x^2 + x + 2 \text{ and } g = x^2 - 1.$$

We have by the division algorithm

$$2x^3 - 5x^2 + x + 2 = (2x - 5)(x^2 - 1) + (3x - 3).$$

Repeating,

$$(x^2 - 1) = (x/3 + 1/3)(3x - 3) + 0.$$

SO we have $3x - 3$ as a possible $\text{hcf}(f, g)$ or, if we prefer, $x - 1$.

Back to generalities:

Invariance Lemma for hcf's of polynomials *Let f and g be non-zero polynomials in $\mathbb{R}[x]$. Write*

$$f = qg + r \quad \textit{with } q, r \in \mathbb{R}[x] \textit{ and } r = 0 \textit{ or } \deg r < \deg g.$$

Assume $\text{hcf}(g, r)$ *exists. Then* $\text{hcf}(f, g)$ *exists and equals* $\text{hcf}(g, r)$.

Proof. Exactly like that of the Invariance Lemma for hcf's of natural numbers. [Exercise: do it.] □

5.7 Highest Common factor theorem for polynomials.

(i) *Let f and g be non-zero polynomials in $\mathbb{R}[x]$. Then a highest common factor,* $\text{hcf}(f, g)$, *of f and g exists;*

(ii) (*the* **hcf formula**) *there exist polynomials m and n such that*

$$\text{hcf}(f, g) = mf + ng. \qquad \text{(similar to EEA)}$$

Proof. Without loss of generality we may assume that $\deg g \leqslant \deg f$. In what follows, all quantities appearing are polynomials and at each step the division algorithm is invoked. We assume that r_k is the first remainder to be zero and stop as soon as such a zero remainder occurs, so that $\deg r_0, \ldots, \deg r_{k-1}$ are defined. We can write

$$\begin{aligned}
f &= q_0 g + r_0 && \text{where } \deg r_0 < \deg g \\
g &= q_1 r_0 + r_1 && \text{where } \deg r_1 < \deg r_0 \\
r_0 &= q_2 r_1 + r_2 && \text{where } \deg r_2 < \deg r_1 \\
&\cdots \quad \cdots \quad \cdots && \\
r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1} && \text{where } \deg r_{k-1} < \deg r_{k-2} \\
r_{k-2} &= q_k r_{k-1} + r_k && \text{where } 0 = r_k
\end{aligned}$$

Lemma 2 in 4.4, applied to the degrees $\deg r_0, \deg r_1, \ldots$, ensures that the process does indeed terminate in a finite number of steps. The Invariance Lemma tells us that

$$\text{hcf}(f, g) = \text{hcf}(g, r_0) = \ldots = \text{hcf}(r_{k-2}, r_{k-1}) = r_{k-1}.$$

To obtain the hcf formula we retrace our steps: write r_{k-1} in terms of r_{k-2}, then substitute the resulting formula for r_{k-1} into the equation for r_{k-3} and so on until a formula in f and g of the required form is obtained. □

5.8 Irreducible polynomials. A non-zero polynomial $p \in \mathbb{R}[x]$ of degree k is said to be *irreducible* if its only divisors are polynomials of degree 0 (ie non-zero constants) and degree k. Irreducible polynomials play the role for $\mathbb{R}[x]$ that primes do for $\mathbb{Z}$ and there is a unique factorisation theorem for $\mathbb{R}[x]$ (or $K[x]$, where K is any field). This is a parallel to the Fundamental Theorem of Arithmetic for $\mathbb{Z}$.

Lecture examples

- $2x^3 - 5x^2 + x + 2$ can be factorised as $(x-1)(x-2)(2x+1)$ or $2(x-1)(x+1/2)(x-2)$ or $2(1-x)(x+1/2)(2-x)$ or ....
- In $\mathbb{R}[x]$ we can factorise $x^3 - 1$ as $(x-1)(x^2+x+1)$ but we cannot write it as a product of factors of degree 1 in $\mathbb{R}[x]$.

We shan't pursue existence and uniqueness of factorisation into irreducibles any further in this course. We conclude this section with a very useful, and probably familiar, theorem we can use to test for factors of degree 1.

5.9 The Remainder Theorem. *Let $f \in \mathbb{R}[x]$ and let $a \in \mathbb{R}$. Then there exists $qin\mathbb{R}[x]$ such that*

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

In particular, $(x-a)$ is a factor of f if and only if $f(a) = 0$.

Proof. The division theorem for polynomials tells us that we can find polynomials q and r such that $f = q \cdot (x-a) + r$ with either $r = 0$ or $\deg r < \deg(x-a) = 1$. Hence either $r = 0$ or r is a constant. Evaluating at $x = a$ we get $r = 0$ if and only if $f(a) = 0$. □

6. Equivalence relations, and modular arithmetic

The first part of this section considers equivalence relations on a set and establishes their relationship to partitions of that set. [Some of this material was also covered in Introduction to Pure Mathematics in MT, but for completeness and ease of reference a self-contained treatment is included here.]

In the second part of the section we focus on a special equivalence relation, congruence mod n, on $\mathbb{Z}$ and use it to manufacture important examples of finite rings and fields.

6.1 Relations: recap. [web notes only] Fix a non-empty set Ω. A *binary relation* on Ω is (officially) a subset $\mathcal{R}$ of $\Omega \times \Omega$. So the elements of $\sim$ are certain pairs $(a, b) \in \Omega \times \Omega$. A relation $\mathcal{R}$ gives a true/false split (a *predicate*): for $(a, b) \in \Omega \times \Omega$ we may have

- $(a, b) \in \mathcal{R}$ is **true**: we write $a\mathcal{R}b$;
- $(a, b) \in \mathcal{R}$ is **false**: we write $a \not\mathcal{R} b$.

We henceforth use the symbol $\sim$ ('twiddles') rather than $\mathcal{R}$ to denote a general relation, and adopt the more usual notation $a \sim b$ in place of $(a, b) \in \sim$.

Relations of special types. We say $\sim$ is

(R) *reflexive* if $a \sim a$ for any $a \in \Omega$;

(S) *symmetric* if $a \sim b$ implies $b \sim a$ for any $a, b \in \Omega$;

(T) *transitive* if $a \sim b$ and $b \sim c$ imply $a \sim c$ for any $a, b, c \in \Omega$.

Examples

Ω	$\sim$	(R)	(S)	(T)
$\mathbb{R}$	$\leqslant$	Y	N	Y
$\mathbb{R}$	$<$	N	N	Y
lines in $\mathbb{R}^2$	"is parallel to"	Y	Y	Y
lines in $\mathbb{R}^2$	"is orthogonal to"	N	Y	N
$\mathbb{Z}$	$a \sim b$ iff $3 \mid (a-b)$	Y	Y	Y
$\mathbb{N} \setminus \{0\}$	"divides"	Y	N	Y
any Ω	$=$	Y	Y	Y
any Ω	$a \sim b\ \forall a, b$	Y	Y	Y

6.2 Equivalence relations and equivalence classes: definitions and examples. Let Ω be a non-empty set and $\sim$ a relation on Ω. Then $\sim$ is an *equivalence relation* if $\sim$ satisfies (R), (S) and (T). When $\sim$ is an equivalence relation we shall call

$$[a] := \{\, b \in \Omega \mid a \sim b \,\}$$

the *equivalence class* of a.

6.3 Examples of equivalence relations. [mainly web notes only]

(1) Equality relation, $=$, on any non-empty set: $[a] = \{a\}$ for any a.

- $\Omega = \mathbb{R}^2$: $(a, b) \sim (c, d)$ if and only iff $a^2 + b^2 = c^2 + d^2$. Notice that (R), (S) and (T) hold for this $\sim$ precisely because they hold for the $=$ relation on $\mathbb{R}$. Watch out for other examples which work in the same way.

$$[(a, b)] = \{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = c^2 \,\}, \quad \text{where } c^2 = a^2 + b^2.$$

(2) **Fractions** Let $\Omega = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ with $\sim$ given by

$$(m, n) \sim (p, q) \iff \frac{m}{n} = \frac{p}{q}.$$

So $[(m, n)]$ contains all pairs (p, q) which correspond to the same rational number as does (m, n).

(3) On **matrices**: We have various useful equivalence relations On $\mathrm{M}_n(K)$ (K any field, $n \geqslant 2$). For example, consider $\sim$ given as follows:

- **Row equivalence**: $A \sim B$ if and only if there exists an invertible $P \in \mathrm{M}_n(K)$ such that $PA = B$—this says that B is obtainable from A by elementary row operations;
- **Equivalence**: $A \sim B$ if and only if there exist invertible $P, Q \in \mathrm{M}_n(K)$ such that $PAQ = B$—this says that B is obtainable from A by elementary row and column operations or, equivalently, A and B have the same reduced Echelon form. **LA theorem**:

$$A \sim B \iff A \text{ and } B \text{ have the same rank.}$$

- **Similarity**: $A \sim B$ if and only if there exists an invertible P such that $PAP^{-1} = B$—A and B represent the same linear transformation with respect to possibly different bases.

The first two of these can be extended to non-square matrices: we can define row equivalence, and equivalence, on $\mathrm{M}_{m,n}(K)$.

(4) [Lecture example] The torus.

(5) **integers modulo n.** [Recap from Introduction to Pure Mathematics] Let $\Omega = \mathbb{Z}$ and let $n > 1$ be a fixed integer. Then

$$a \sim_n b \iff n | (a - b) \iff \exists k \in \mathbb{Z} \text{ such that } a - b = kn$$

defines an equivalence relation on $\mathbb{Z}$.

- **The case** $n = 2$: there are two equivalence classes

$$E = \{\, 2k \mid k \in \mathbb{Z} \,\} = [0],$$
$$O = \{\, 2k + 1 \mid k \in \mathbb{Z} \,\} = [1].$$

Notice that $[2] = [4] = \ldots = [-2] = [-4] = \ldots = [0]$ and $[3] = \ldots = [1]$.

- **The case** $n = 3$: there are three equivalence classes

$$[0] = \{\, 3k \mid k \in \mathbb{Z} \,\},$$
$$[1] = \{\, 3k + 1 \mid k \in \mathbb{Z} \,\},$$
$$[2] = \{\, 3k + 2 \mid k \in \mathbb{Z} \,\}.$$

[Lecture: pictures of decomposition into equivalence classes for some familiar examples.]

6.4 Partitions. Let $\sim$ be an equivalence relation on Ω. In all our examples: each element of Ω belongs to one, and only one, equivalence class. In every case the equivalence classes form a partition of Ω in the sense of the following definition. Let Ω be a non-empty set and $\mathcal{U} = \{U_i\}_{i \in I}$ be a non-empty family of subsets of Ω. Then $\mathcal{U}$ is called a *partition* of Ω if

(part1) $U_i \neq \varnothing$ for all $i \in I$;

(part2) $U_i \cap U_j = \varnothing$ for all indices i and j in I such that $i \neq j$;

(part3) $\bigcup_{i\in I} U_i = \Omega$.

In words: Ω is the disjoint union of the non-empty sets U_i.

Examples (for familiarisation with the notation in the definition of partition) [web notes only]

(1) Evens/odds: $\{E, O\}$ is a partition of $\mathbb{Z}$; here we may take $I = \{0, 1\}$, $U_0 = E$ and $U_1 = O$.

(2) $\{U_i\}_{i\in I=\{0,1,\dots,n-1\}}$, where $U_i = [i]_{\sim_n}$ is a partition of $\mathbb{Z}$, for any n.

(3) $\{U_i\}_{i\in\mathbb{N}}$, where $U_i = \{n \in \mathbb{N} \mid 2^i - 1 \leqslant n < 2^{i+1} - 1\}$ is a partition of $\mathbb{N}$.

(4) $\{\{x\}\}_{x\in\mathbb{R}}$ is a partition of the real numbers, with indexing set $\mathbb{R}$.

The indexing set I can be finite or infinite, countable or uncountable. The subsets U_i in the partition may be all the same size, or of different sizes.

We shall reveal that there is a one-to-one correspondence between equivalence relations on a set Ω and partitions of Ω: The main work is done in Theorems 'equiv-part' and 'part2equiv', and the correspondence is set out in Theorem 'part-equiv'.

6.5 Theorem (from an equivalence relation to a partition)**.** *Let Ω be a non-empty set and $\sim$ an equivalence relation on Ω. Then there exists a partition $\mathcal{U}_\sim$ in which the sets are the (distinct) equivalence classes for $\sim$.*

Proof. (part1) and (part3): For any $a \in \Omega$, we have $a \in [a]$, by (R). Hence each equivalence class is non-empty and the union of all the equivalence classes is Ω.

(part2): We need to show that, for $a, b \in \Omega$,

$$[a] \cap [b] = \varnothing \quad \text{or} \quad [a] = [b].$$

Suppose $[a] \cap [b] \neq \varnothing$. We shall prove $[a] = [b]$. Let $c \in [a] \cap [b]$. Then $a \sim c$ and $b \sim c$. By (S), $c \sim b$ and then by (T) we have $a \sim b$. Hence $b \in [a]$. Take $d \in [b]$ so $b \sim d$. By (T) again, $d \in [a]$. So $[b] \subseteq [a]$. Likewise $[a] \subseteq [b]$. So the equivalence classes of two given elements are either equal or disjoint. □

As an application we revisit Theorem 1.3 from Linear Algebra II on decomposing permutations into disjoint cycles. Recall that for a permutation $\sigma \in \mathrm{Sym}(n)$ and $k \in \mathbb{Z}$,

$$\sigma^k = \begin{cases} \sigma \circ \cdots \circ \sigma \quad (k \text{ times}) & \text{if } k > 0, \\ \mathrm{id} \quad (\text{the identity}) & \text{if } k = 0, \\ \sigma^{-1} \circ \cdots \circ \sigma^{-1} \quad (|k| \text{ times}) & \text{if } k < 0, \end{cases}$$

and the usual rules of exponents apply, so that, for $k, \ell \in \mathbb{Z}$,

$$\sigma^{k+\ell} = \sigma^k \sigma^\ell \quad \text{and} \quad \sigma^{k\ell} = (\sigma^k)^\ell.$$

6.6 Theorem (cycle decomposition of permutation)**.** *Let $\sigma \in \mathrm{Sym}(n)$, the permutations of $\{1, 2, \dots, n\}$. Then σ is expressible as a finite product of disjoint cycles.*

Proof. Let $\Omega = \{1, \dots, n\}$. For $a, b \in \Omega$ define $a \sim b$ if and only if there exists $k \in \mathbb{Z}$ such that $b = a\sigma^k$.

Claim: $\sim$ is an equivalence relation. The properties (R), (S) and (T) follow directly from properties of exponents.

By Theorem 6.5, Ω is the disjoint union of the distinct $\sim$-equivalence classes (also called *orbits*). Let these equivalence classes be $[c_1], \ldots, [c_r]$. Let $[c_i]$ have size n_i $(i = 1, \ldots, r)$. Then

$$[c_i] = \{c_i, c_i\sigma, c_i\sigma^{n_i-1}\}$$

and σ acts on $[c_i]$ as the cycle

$$\sigma_i := (c_i \ c_i\sigma \ \ldots \ c_i\sigma^{n_i-1}).$$

The cycles $\sigma_1, \ldots, \sigma_r$ are disjoint and $\sigma = \sigma_1\sigma_2\cdots\sigma_r$ (note that disjoint cycles commute, so the order in which they are performed is immaterial). $\square$

We now return to the general theory.

6.7 Theorem (from a partition to an equivalence relation)**.** *Assume that $\mathcal{U} = \{U_i\}_{i\in I}$ is a partition of Ω. Let $\sim_{\mathcal{U}}$ be the relation*

$$a \sim_{\mathcal{U}} b \iff \big(\exists i \text{ such that } a \in U_i \text{ and } b \in U_i\big).$$

Then $\sim_{\mathcal{U}}$ is an equivalence relation and its (distinct) equivalence classes are exactly the sets U_i of the partition $\mathcal{U}$.

Proof. Properties (part1) and (part2) for a partition guarantee that $\sim_{\mathcal{U}}$ is an equivalence relation. **Exercise:** check this, and also that $U_i = [x]$ for any $x \in U_i$ ((part3) is used here).

6.8 Theorem (the correspondence between equivalence relations and partitions)**.** *Let Ω be a non-empty set. Then the maps*

$$\Phi\colon \sim \longmapsto \mathcal{U}_{\sim} \quad \text{and} \quad \Psi\colon \mathcal{U} \longmapsto \sim_{\mathcal{U}}$$

set up a bijective correspondence between the set of all equivalence relations on Ω and the set of all partitions of Ω.

Proof. By construction,

$$\sim_{\mathcal{U}_\sim} = \sim \quad \text{and} \quad \mathcal{U}_{\sim_{\mathcal{U}}} = \mathcal{U}.$$

This is saying that Φ and Ψ are mutually inverse maps, and so set up the required correspondence. $\square$

Application: counting equivalence relations. [Lecture example]

6.9 Notation: the set of equivalence classes for an equivalence relation. Consider an equivalence relation $\sim$ on a set Ω. We shall denote the set of equivalence classes by $\Omega/\sim$ and call it the *quotient of Ω by $\sim$*. Often, as with the example on fractions in 6.3(2), we do not want to distinguish between elements which are related, and passing from Ω to $\Omega/\sim$ via the map $x \mapsto [x]$ $(x \in \Omega)$ allows us to think of such elements as being identified. [Informal introductory discussion on quotient structures in lecture.]

We now look more closely at the partition of $\mathbb{Z}$ induced by $\sim_n$ and the associated quotient $\mathbb{Z}_n := \mathbb{Z}/\sim_n$.

6.10 $\sim_n$ revisited. Consider the equivalence relation $\sim_n$ on $\mathbb{Z}$ given by

$$a \sim_n b \iff n|(a-b).$$

The relation $\sim_n$ is so widely used that it has a special name, *congruence mod n* (where *mod* stands for *modulo*), and a special symbol: if integers a and b are such that $a \sim_n b$ we write $a \equiv b \mod n$ or simply $a \equiv b$ if there is no ambiguity.

We next show that congruence mod n is compatible, in a natural sense, with the arithmetic operations on $\mathbb{Z}$.

6.11 Congruence Lemma. *Fix a positive integer n. Then, for $s, t, u, v \in \mathbb{Z}$,*

$$s \equiv u \mod n \quad \textit{and} \quad t \equiv v \mod n$$

imply

$$s + t \equiv u + v \mod n \quad \textit{and} \quad st \equiv uv \mod n.$$

Proof. By assumption there exist integers p and q such that $s - u = pn$ and $t - v = qn$. Hence

$$(s+t) - (u+v) = (s-u) + (t-v) = pn + qn = (p+q)n,$$

so $s + t \equiv u + v \mod n$ and similarly

$$st - uv = (s-u)t + (t-v)u = ptn + qvn = (pt + qv)n,$$

so $st \equiv uv \mod n$. □

6.12 Theorem (elements of $\mathbb{Z}_n$)**.** *Let n be a positive integer. There are precisely n equivalence classes for $\sim_n$, namely $[0], [1], \ldots, [n-1]$.*

Proof. By the division algorithm any $a \in \mathbb{Z}$ can be uniquely expressed as

$$a = qn + r \quad \text{where } 0 \leqslant r < n.$$

Then $a \in [r]$ and so $[a] = [r]$. Also if we take $r, s \in \{0, 1, \ldots, n-1\}$, with $r \geqslant s$, then $r \sim_n s$ would imply that there exists $k \in \mathbb{N}$ such that $0 \leqslant r - s = kn$ and so $kn \leqslant r < n$. This forces $k = 0$ and hence $r = s$. Therefore there are exactly n equivalence classes and these may be represented as $[0], [1], \ldots, [n-1]$. □

starting to be more relevant

We can now 'transfer' arithmetic operations from $\mathbb{Z}$ to $\mathbb{Z}_n$.

6.13 Addition and multiplication on $\mathbb{Z}_n$. Observe that the numbers $0, 1, \ldots, n-1$ labelling the distinct equivalence classes of $\sim_n$ (see Theorem 6.12 are exactly the possible remainders that we can get when we divide an integer by n. Furthermore, we can alternatively specify $\sim_n$ by saying that $a \sim_n b$ if and only if a and b leave the same remainder when they are divided by n. And we have a natural map $x \mapsto \overline{x}$ on $\mathbb{Z}$ which sends x to its remainder after division by n.

We can therefore think of the members of $\mathbb{Z}_n = \mathbb{Z}/\sim_n$ as being identified with $\overline{0}, \overline{1}, \ldots, \overline{n-1}$. We may define binary operations of sum and product on $\mathbb{Z}_n$ as follows: $a, b \in \{0, \ldots, n-1\}$,

$$\overline{a} + \overline{b} := \overline{a+b} \quad \text{and} \quad \overline{a} \cdot \overline{b} := \overline{ab}.$$

For example, in $\mathbb{Z}_8$ we have

$$\overline{2} + \overline{3} = \overline{5}, \qquad \overline{4} + \overline{7} = \overline{11} = \overline{3},$$
$$\overline{2} \cdot \overline{3} = \overline{6}, \qquad \overline{4} \cdot \overline{7} = \overline{28} = \overline{4}.$$

Exercise: Write out addition and multiplication tables for $\mathbb{Z}_3$ and $\mathbb{Z}_4$.

Remarks: There are no issues here of the operations being well defined, since we have assigned a unique label k in the range $0, 1, \ldots, n-1$ to each element $\overline{k}$ in $\mathbb{Z}_n$. But to represent the sum of $\overline{a}$ and $\overline{b}$ using a standard label we will usually need to calculate the remainder left by $a + b$, and likewise for product.

6.14 Passing from $\mathbb{Z}$ to $\mathbb{Z}_n$. We have a natural map π from $\mathbb{Z}$ to $\mathbb{Z}_n$:

$$\pi\colon x \mapsto \overline{x}.$$

Then π is surjective, but not injective).

Now consider $s, t \in \mathbb{Z}$ leaving remainders a, b respectively on division by n. Then $\overline{s} = \overline{a}$ and $\overline{t} = \overline{b}$. Therefore

$$\begin{aligned}
\pi(s+t) &= \overline{s+t} && \text{(by defn of } \pi\text{)}\\
&= \overline{a+b} && \text{(by the Congruence Lemma)}\\
&= \overline{a} + \overline{b} && \text{(by defn of } + \text{ in } \mathbb{Z}_n\text{)}\\
&= \overline{s} + \overline{t}\\
&= \pi(s) + \pi(t).
\end{aligned}$$

and likewise $\pi(st) = \pi(s) \cdot \pi(t)$. Thus the map $\pi\colon \mathbb{Z} \to \mathbb{Z}_n$ preserves addition and multiplication.

6.15 Theorem ($\mathbb{Z}_n$ as a ring)**.** *Let $n \geqslant 2$. Endow $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}$ with the binary operations of addition and multiplication* mod n*: for $a, b \in \{0, \dots, n-1\}$,*

$$\begin{aligned}
\overline{a} + \overline{b} &:= \overline{a+b},\\
\overline{a} \cdot \overline{b} &:= \overline{ab}.
\end{aligned}$$

(1) *With these operations $\mathbb{Z}_n$ is a commutative ring with identity. The zero element for addition is $\overline{0}$ and $-\overline{a} = \overline{n-a}$. The multiplicative identity is $\overline{1}$.*

(2) *The following are equivalent:*

(i) *n is prime;*

(ii) *$\mathbb{Z}_n$ is an integral domain;*

(iii) *$\mathbb{Z}_n$ is a field.*

Proof. (1) We can use the structure-preserving properties of the map $\pi\colon \mathbb{Z} \to \mathbb{Z}_n$ to verify the associativity, commutativity and distributivity properties. For example, for $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}_n$ we have $a(b+c) = (a+b)+c$ in $\mathbb{Z}$, so that

$$\begin{aligned}
\overline{a} + (\overline{b} + \overline{c}) = \overline{a} + \overline{b+c} = \pi(a) + \pi(b+c) = \pi(a + (b+c))\\
= \pi((a+b)+c) = \overline{a+b} + \overline{c} = (\overline{a} + \overline{b}) + \overline{c}.
\end{aligned}$$

Certainly $\overline{a} + \overline{0} = \overline{a+0} = \overline{a}$ and $\overline{a}\overline{1} = \overline{a1} = \overline{a}$. Also $\overline{a} + \overline{n-a} = \overline{a + (n-a)} = \overline{n} = \overline{0}$. Hence (1) holds.

(2) Since we have assumed $n \geqslant 2$, we have $\overline{0} \neq \overline{1}$. We claim that $\mathbb{Z}_n$ has no zero divisors if and only if n is prime. If n is not prime we may write $n = ab$ where $1 < a < n$ and $1 < b < n$. Then $\overline{a} \neq \overline{0}$, $\overline{b} \neq \overline{0}$ but $\overline{a}\overline{b} = \overline{ab} = \overline{0}$. Conversely, assume n is prime and that $\overline{a}\overline{b} = \overline{0}$. Then $n|(ab)$, and hence, because n is prime, $n|a$ or $n|b$ (recall 4.10), whence $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$.

(ii) implies (iii) always, and (iii) implies (ii) because $\mathbb{Z}_n$ is finite (see Theorem 2.8). □

7. Groups in general and permutation groups in particular

In this section we embark on a study of groups, looking in particular at the group of permutations of a finite set. Recall from Section 2:

1.4 axioms for a group and an abelian group, and initial examples;

1.5 the theorem that $(\mathrm{Sym}(X); \circ)$ is a group.

Recall also

Linear Algebra II web notes, Section 1, on permutations of a finite set.

7.1 Groups: preliminaries Assume $(G; \star)$ is a group. If the set G is finite, then we call its size, $|G|$, the *order* of G.

Cancellation rules in a group. For all a, x, $y \in G$,

$$a \star x = a \star y \implies x = y;$$
$$x \star a = y \star a \implies x = y$$

(Problem sheet 4, Question 1).

As is customary, we shall henceforth usually suppress $\star$ in products of group elements and write ab for the product $a \star b$ in a generic group $(G; \star)$.

Cayley Tables. You have seen several instances of binary operations being specified by a 'multiplication table' (see 1.2). In the context of groups such a table is known as a *Cayley table*. This is a 'brute force' way of presenting a group, but can be useful when working with *small* finite groups. In a Cayley table, each element occurs once and once only in each row and each column. (This is so because the cancellation rules tell us that $r_a : x \mapsto x \star a$ and $\ell_a : x \mapsto a \star a$ are injective maps, for any fixed a; by the Pigeonhole Principle, r_a and ℓ_a are bijections.)

Advice: use Cayley tables sparingly.

7.2 Subgroups: definitions and the Subgroup Test. With groups, as with vector spaces and rings, we can considerably enlarge our range of examples by looking at substructures.

Let $(G, \star)$ be a group and let H be a non-empty subset of G. If H is a group when equipped with the restriction of $\star$ then H is said to be a *subgroup* of G. and we write $H \leqslant G$.

Extreme cases: In any group G (with identity denoted e) we have

- $\{e\}$ is always a subgroup of G;
- G is always a subgroup of G.

We say that a subgroup H of a group G is *non-trivial* if $H \neq \{e\}$ and *proper* if $H \subsetneq G$.

The Subgroup Test: *Let $(G; \star)$ be a group and H a non-empty subset of G.*

(i) *$(H; \star)$ is a subgroup of $(G; \star)$ if and only if*

(SG1) *$h \star k \in H$ for all h, $k \in H$;*

(SG2) *$h^{-1} \in H$ for all $h \in H$.*

Alternatively:

(ii) *$(H; \star)$ is a subgroup of $(G; \star)$ if and only if it satisfies the single condition*

(SG) *for all h, $k \in H$, $h\,k^{-1} \in H$.*

Proof. [web notes only] $\Longleftarrow$ is an immediate consequence of the definition of a subgroup.

$\Longleftarrow$ Suppose $\varnothing \neq H \subseteq G$ and that (SG1) and (SG2) hold. Note first that (SG1) is just the requirement (bop) that $\star$ defines a binary operation on H. Associativity is inherited from G so axiom (G1) is satisfied by $(H; \star)$.

Now, because G is non-empty, we may take some fixed element $c \in H$. Apply (SG2) with h as c to get $c^{-1} \in H$. Now apply (SG1) with $h = c$ and $k = c^{-1}$ to get $e_G \in H$. Thus (see the

note above) axiom (G2) (identity element) is satisfied. Finally, (G3) (inverses) is just the condition (SG2). We have proved that H is a subgroup.

Finally we need to show that (SG1) and (SG2) together are equivalent to the condition (SG). Clearly, (SG1) and (SG2) taken together imply (SG). Conversely, if (SG) holds, we can take $h = k = c$ in (SG), where c is some element of the non-empty set H. We get $e_G \in H$. Then put $h = e_G$ in (SG) to show that (SG2) holds. Now, for any $h, \tilde{h} \in H$, apply (SG) with k as $(\tilde{h}^{-1}$ to get $h\tilde{h} \in H$. $\square$

Warning: Observe how the fact that $H \neq \varnothing$ is used in the above proof. When applying the Subgroup Test, don't forget to check/note that your candidate subgroup is indeed a non-empty set. One way to do this is to verify that $e_G \in H$.

Some examples of subgroups: (SG) can easily be verified for the following non-empty subsets of the groups specified.

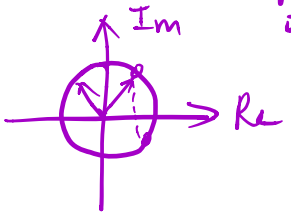
- $\{ x \in \mathbb{R} \mid x > 0 \}$ is a subgroup of $\mathbb{R} \setminus \{0\}$ under multiplication.
- $S^1 := \{ z \in \mathbb{C} \mid |z| = 1 \}$ is a subgroup of $\mathbb{C} \setminus \{0\}$ under multiplication.

(always ends up on unit circle)

$\vec{z}_1 \cdot \vec{z}_2 = \|\vec{z}_1\| \cdot \|\vec{z}_2\| \cdot \cos(\vec{z}_1, \vec{z}_2)$

OR $(r_1 \cdot \cos\theta_1)(r_2 \cdot \cos\theta_2) = r_1 \cdot r_2 \cdot \cos(\theta_1 + \theta_2)$

$= 1 \cdot \cos(\theta_1 + \theta_2)$



z_1^{-1} → inverse conjugate

7.3 Powers of group elements. Let $(G;\star)$ be a group (finite or infinite). In the same way as for a permutation, we can define integer powers of an element $g \in G$ as follows: for $k \in \mathbb{Z}$,

$$g^k = \begin{cases} \mathrm{id} & \text{if } k = 0, \\ \underbrace{g \star \cdots \star g}_{k \text{ times}} & \text{if } k > 0, \\ (g^{|k|})^{-1} & \text{if } k < 0. \end{cases}$$

The usual rules of exponents apply: for example, for $k, \ell \in \mathbb{Z}$,

$$g^{k+\ell} = g^k g^\ell = g^\ell g^k \quad \text{and} \quad \left(g^k\right)^\ell = g^{k\ell}.$$

7.4 Cyclic groups. Let $(G;\star)$ be a group (finite or infinite) and let $g \in G$. It is easy to use the Subgroup Test and the rules for exponents to check that

$$\langle g \rangle := \{\, g^k \mid k \in \mathbb{Z} \,\}.$$

is always a subgroup of G, and is the smallest subgroup of G which contains g. We call $\langle g \rangle$ the *cyclic subgroup generated by* g.

We say that the group $(G;\star)$ is *cyclic* if there exists $g \in G$ such that $G = \langle g \rangle$. Note that every cyclic group is abelian.

Exercise: for $k = 2, 3.4$, write out the Cayley table for a cyclic group $\{e, g, \ldots, g^{k-1}\}$, where $g^k = e$.

Examples: cyclic groups:

- $(\mathbb{Z};+)$ is an infinite cyclic group.
- $(\mathbb{Z}_n;+)$ is a cyclic group of order n, generated by $\overline{1}$. So a cyclic group of order n exists for each n.
- Consider
$$\{\, \mathrm{e}^{2k\pi \mathrm{i}/n} \in \mathbb{C} \mid k = 0, \ldots, n-1 \,\}$$
be the set of nth roots of 1 in $\mathbb{C}$. This is a group under multiplication, and is cyclic, generated by $\mathrm{e}^{2k\pi \mathrm{i}/n}$.

Examples: non-cyclic groups:

- $(\mathbb{R};+)$ is not cyclic. Note that an infinite cyclic group is automatically countable (since $\mathbb{Z}$ is); but $\mathbb{R}$ is uncountable.
Further examples below, when we consider subgroups of permutation groups.

7.5 Theorem. *A subgroup of a cyclic group is cyclic.*

Proof. Let H be a subgroup of the cyclic group $G = \langle g \rangle$. Let

$$S := \{k \in \mathbb{Z} \mid g^k \in H\}.$$

Because H is a subgroup, the rules for exponents in 7.3 tell us that S is a non-empty subset of $\mathbb{Z}$ closed under addition and subtraction. By Theorem 4.2(1), there exists $m \in \mathbb{Z}$ such that $S = m\mathbb{Z}$. It follows easily that $H = \langle g^m \rangle$. $\square$

7.6 Permutation groups. We shall henceforth denote by S_n, rather than $\mathrm{Sym}(n)$, the group of permutations of the finite set $\{1, 2, \ldots, n\}$. Recall from LAII that we have $|S_n| = n!$.

We can view S_n as sitting inside S_{n+1} as the subgroup of permutations of $\{1, \ldots, n+1\}$ having $n+1$ as a fixed point. The group S_3 is non-abelian (see detailed discussion of S_3 below). Hence S_n is non-abelian for $n > 3$ too.

Cycle decomposition: Every element of S_n is expressible in an essentially unique way as a product of disjoint cycles (LAII, Theorem 1.3 and accompanying discussion, and GRF 6.6), and has an associated **cycle-type** (see LAII notes, p. 3), specifying the lengths of the constituent cycles. When talking about permutations and cycle-types we shall suppress 1-cycles.

Parity: An element σ of S_n is classified as

even if σ is expressible as a product of an even number of transpositions (i.e. 2-cycles),

odd if σ is expressible as a product of an odd number of transpositions

(see LAII, Theorems 1.5 and 1.6, and Definition 1.7). Here the transpositions are not required to be disjoint. Recall that an m-cycle $(a_1\ a_2\ \ldots\ a_m)$ is even iff m is odd.

7.7 Small permutation groups. For $n-1$,and $n = 2$, the group S_n has, respectively, 1-element and 2 elements. The cases $n = 3, 4$ are more interesting.

The group S_3: We can list the six elements of S_3 as

$$\mathrm{id},\quad (1\ 2),\quad (2\ 3),\quad (3\ 1),\quad (1\ 2\ 3),\quad (1\ 3\ 2).$$

The Cayley table is

$\circ$	id	(1 2)	(2 3)	(3 1)	(1 2 3)	(1 3 2)
id	id	(1 2)	(2 3)	(3 1)	(1 2 3)	(1 3 2)
(1 2)	(1 2)	id	(1 3 2)	(1 2 3)	(3 1)	(2 3)
(2 3)	(2 3)	(1 2 3)	id	(1 3 2)	(1 2)	(3 1)
(3 1)	(3 1)	(1 3 2)	(1 2 3)	id	(1 2 3)	(1 2)
(1 2 3)	(1 2 3)	(2 3)	(3 1)	(1 2)	(1 3 2)	id
(1 3 2)	(1 3 2)	(3 1)	(1 2)	(2 3)	id	(1 2 3)

Note in particular that

$$(1\ 2)(1\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1\ 2\ 3) \text{ whereas } (1\ 3)(1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1\ 3\ 2).$$

So S_3 is **non-abelian**.

The three transpositions are odd permutations. The other elements, id, $(1\ 2\ 3) = (1\ 2)(1\ 3)$ and $(1\ 3\ 2) = (1\ 3)(1\ 2)$, are even permutations.

The following are the proper non-trivial subgroups of S_3:

$$\begin{aligned} A_3 :=& \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\} \quad (\text{note } (1\ 3\ 2) = (1\ 2\ 3)^2 = (1\ 2\ 3)^{-1}), \\ & \{\mathrm{id}, (1\ 2)\}, \\ & \{\mathrm{id}, (2\ 3)\}, \\ & \{\mathrm{id}, (3\ 1)\}. \end{aligned}$$

Note that A_3 is cyclic,

The group S_4: This has 24 elements. You were asked on LAII problem sheet 1 (Question 2) to classify the elements according to their cycle-types. We have

1	identity	even
6	2-cycles	odd
8	3-cycles	even
6	4-cycles	odd
3	double transpositions	even

The following are important subgroups of S_4:

$$V_4 := \{\, \mathrm{id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3) \,\},$$
$$A_4 := \{\, \text{all even permutations in } S_4 \,\}.$$

We note that V_4 is a subgroup of A_4 and that V_4 is not a cyclic group (why?).

Advice: get to know S_3 and S_4 well.

7.8 Alternating groups. Generalising what we noted above for $n = 3, 4$, we define, for any n,

$$A_n := \{\, \sigma \in S_n \mid \sigma \text{ is even} \,\};$$

this forms a subgroup of S_n. We have $|A_n| = n!/2$.

[Lecture: why study permutation groups?—informal remarks]