MTH 310
Lecture Notes
Based on Hungerford, Abstract Algebra

Ulrich Meierfrankenfeld

Department of Mathematics

Michigan State University

East Lansing MI 48824

meier@math.msu.edu

December 3, 2014

Contents

Chapter 0

Set, Relations and Functions

0.1 Logic

In this section we will provide an informal discussion of logic. A statement is a sentence which is either true or false, for example

(1) $1 + 1 = 2$

(2) $\sqrt{2}$ is a rational number.

(3) π is a real number.

(4) Exactly 1323 bald eagles were born in 2000 BC,

all are statements. Statement (1) and (3) are true. Statement (2) is false. Statement (4) is probably false, but verification might be impossible. It nevertheless is a statement.

Let P and Q be statements.

"P and Q" is the statement that P is true and Q is true. We illustrate the statement P and Q in the following *truth table*

P	Q	P and Q
T	T	T
T	F	F
F	T	F
F	F	F

"P or Q" is the statement that at least one of P and Q is true:

P	Q	P or Q
T	T	T
T	F	T
F	T	T
F	F	F

So “P or Q” is false exactly when both P and Q are false.

“not-P” (pronounced ‘not P’ or ‘negation of P’) is the statement that P is false:

P	not-P
T	F
F	T

So not-P is true if P is false. And not-P is false if P is true.

“$P \Longrightarrow Q$” (pronounced “P implies Q”) is the statement “ If P is true, then Q is true”:

P	Q	$P \Longrightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Note here that if P is true, then “$P \Longrightarrow Q$ ” is true if and only if Q is true. But if P is false, then “$P \Longrightarrow Q$” is true, regardless whether Q is true or false. Consider the statement “ Q or not-P” :

P	Q	not-P	Q or not-P
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

(∗) ”Q or not-P” is true if and only ”$P \Longrightarrow Q$” is true.

This shows that one can express the logical operator "$\Longrightarrow$" in terms of the operators " not-" and "or".

"$P \Longleftrightarrow Q$" (pronounced "P is equivalent to Q") is the statement that P is true if and only if Q is true.:

P	Q	$P \Longleftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

So $P \Longleftrightarrow Q$ is true if either both P and Q are true, or both P and Q are false. Hence

(∗∗) "$P \Longleftrightarrow Q$" is true if and only "(P and Q) or (not-P and not-Q)" is true.

To show that P and Q are equivalent often shows that P implies Q and that Q implies P. Indeed the truth table

P	Q	$P \Longrightarrow Q$	$Q \Longrightarrow P$	$(P \Longrightarrow Q)$ and $(Q \Longrightarrow P)$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	F	T	T

shows that

(∗ ∗ ∗) "$P \Longleftrightarrow Q$" is true if and only "($P \Longrightarrow Q$) and ($Q \Longrightarrow P$)" is true.

Often, rather than showing that a statement is true, one shows that the negation of the statement is false (This is called a proof by contradiction). To do this it is important to be able to determine the negation of statement. The negation of not-P is P:

P	not-P	not-(not-P)
T	F	T
F	T	F

The negation of "P and Q" is " not-P or not-Q":

P	Q	P and Q	not-(P and Q)	not-P	not-Q	not-P or not-Q
T	T	T	F	F	F	F
T	F	F	T	F	T	T
F	T	F	T	T	F	T
F	F	F	T	T	F	T

The negation of "P or Q" is " not-P and not-Q":

P	Q	P or Q	not-(P or Q)	not-P	not-Q	not-P and not-Q
T	T	T	F	F	F	F
T	F	T	F	F	T	F
F	T	T	F	T	F	F
F	F	F	T	T	F	T

The statement "not-$Q \Longrightarrow$ not-P" is called the *contrapositive* of the statement "$P \Longrightarrow Q$". It's actually is equivalent to the statement "$P \Longleftrightarrow Q$":

P	Q	$P \Longrightarrow Q$	not-Q	not-P	not-$Q \Longrightarrow$ not-P
T	T	T	F	F	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

The statement " not-$P \Longleftrightarrow$ not-Q" is called the contrapositive of the statement "$P \Longleftrightarrow Q$". It is equivalent to the statement "$P \Longleftrightarrow Q$":

P	Q	$P \Longleftrightarrow Q$	not-P	not-Q	not-$P \Longleftrightarrow$ not-Q
T	T	T	F	F	T
T	F	F	F	T	F
F	T	F	T	F	F
F	F	T	T	T	T

The the statement "$Q \Longrightarrow P$" is called the *converse* of the statement " $P \Longrightarrow Q$". In general the converse is not equivalent to the original statement. For example the statement if $x = 0$ then x is an even integer is true. But the converse (if x is an even integer, then $x = 0$) is not true.

Theorem 0.1.1 (Principal of Substitution)**.** *Let $\Phi(x)$ be formula involving a variable x. For an object d let $\Phi(d)$ be the formula obtained from $\Phi(x)$ by replacing all occurrences of x by d. If a and b are objects with $a = b$, then $\Phi(a) = \Phi(b)$.*

Proof. This should be self evident. For an actual proof and the definition of an formula consult your favorite logic book. □

Example 0.1.2. Let $\Phi(x) = x^2 + 3x + 4$.
If $a = 2$, then

$$a^2 + 3a + 4 = 2^2 + 3 \cdot 2 + 4$$

Notation 0.1.3. *Let $P(x)$ be a statement involving the variable x.*

(a) "for all $x : P(x)$" *is the statement that for objects a the statements $P(a)$ is true. Instead of* "for all $x : P(x)$" *we will also use "$\forall x : P(x)$", "$P(x)$ is true for all x", "$P(x)$ holds for all x" or similar phrases.*

(b) 'there exists $x : P(x)$" *is the statement there exists an object a such that the statements $P(a)$ is true. Instead of* "there exists $x : P(x)$" *we will use "$\exists x : P(x)$", "$P(x)$ is true for some x", "There exists x with $P(x)$" or similar phrases.*

Example 0.1.4. "for all $x : x + x = 2x$" is a true statement.
"for all $x : x^2 = 2$" is a false statement.
"there exists $x : x^2 = 2$" is a true statement.
"$\exists x : x^2 = 2$ and x is an integer" is false statement

Notation 0.1.5. *Let $P(x)$ be a statement involving the variable x.*

(a) *"There exists at most one $x : P(x)$" is the statement*

$$P(x) \text{ and } P(y) \quad \Longrightarrow \quad x = y$$

(b) *"There exists a unique $x : P(x)$" is the statement*

$$\text{there exists } x : \qquad P(y) \quad \Longleftrightarrow \quad y = x$$

Example 0.1.6. "There exists at most one x : ($x^2 = 1$ and x is a real number)" is false since $1^1 = 1$ and $(-1)^1 = 1$, but $1 \neq -1$.

"There exist a unique x : ($x^3 = -1$ and x is a real number)" is true since $x = -1$ is the only elements in $\mathbb{R}$ with $x^3 = 1$.

"There exists at most one x : ($x^2 = -1$ and x is a real number)" is true, since there does not exist any element $x \in \mathbb{R}$ with $x^2 = -1$.

"There exists a unique x : ($x^2 = -1$ and x is a real number)" is false, since there does not exist any element $x \in \mathbb{R}$ with $x^2 = -1$.

Lemma 0.1.7. *Let $P(x)$ be statement involving the variable x. Then*

$$\big(\text{ there exists } x : P(x)\big) \qquad \text{and} \qquad \big(\text{there exists at most one } x : P(x)\big)$$

if and only if

$$\text{there exists a unique } x : P(x)$$

Proof. $\Longrightarrow$: Suppose first that

$$\big(\text{ there exists } x : P(x)\big) \qquad \text{and} \qquad \big(\text{there exists at most one } x : P(x)\big)$$

hold. By definition of "There exists:" we conclude that there exists an object a such that $P(a)$ is true. . Also by definition of "There exists at most one":

$$(*) \qquad P(x) \text{ and } P(y) \quad \Longrightarrow \quad x = y.$$

From $(*)$ and the principal of substitution:

$$(**) \qquad P(a) \text{ and } P(y) \quad \Longrightarrow \quad a = y$$

By A.1.1(LR 7) $P \iff (T \text{ and } P)$ whenever P is a statement and T is a true statement. Since $P(a)$ is a true statement we conclude that

$$(***) \qquad P(y) \quad \Longleftrightarrow \quad P(a) \text{ and } P(y)$$

From $(***)$ and $(**)$ we conclude that

$$(+) \qquad P(y) \quad \Longrightarrow \quad a = y.$$

If $a = y$, then since $P(a)$ is true, we Principal of Substitution shows that $P(y)$ is true. Thus

$$(++) \qquad a = y \quad \Longrightarrow \quad P(y)$$

From (+) and (++) we get

$$P(y) \qquad \Longleftrightarrow \qquad a = y.$$

Hence the definition of "There exists a unique" gives

$$\text{There exists a unique}\, x : P(x).$$

$\Longleftarrow$: Suppose next that

$$\text{There exists a unique}\, x : P(x)$$

holds. Then by definition of "There exists a unique":

$$\text{there exists } x : P(y) \Longleftrightarrow x = y.$$

and so there exists an object a such that

(+++) $$P(y) \qquad \Longleftrightarrow \qquad a = y.$$

Since $a = a$ is true, we conclude that $P(a)$ is true. Thus

(#) $$\text{there exists } x : P(x).$$

holds.

Suppose "$P(x)$ and $P(y)$ " is true. Then $P(x)$ is true and (+++) shows that $x = a$. Also $P(y)$ is true and (+++) gives $y = a$. From $x = a$ and $y = a$ we get $x = y$ (by the Principal of Substitution. We proved that

$$P(x) \text{ and } P(y) \quad \Longrightarrow \quad x = y.$$

and so the definition of "There exists at most one" gives

(##) $$\text{There exists at most one}\, x : P(x).$$

From (#) and (##) we have

$$\text{there exists } x : P(x) \qquad \text{and} \qquad \text{There exists at most one}\, x : P(x).$$

□

Exercises 0.1:

#1. Convince yourself that each of the statement in A.1.1 are true.

#2. Use a truth table to verify the statements LR 17, LR 26, LR 27 and LR 28 in A.1.1.

0.2 Sets

First of all any *set* is a collection of objects.

For example
$$\mathbb{Z} := \{\ldots, -4, -3, -2, -1, -0, 1, 2, 3, 4, \ldots\}$$
is the set of integers. If S is a set and x an object we write $x \in S$ if x is a member of S and $x \notin S$ if x is not a member of S. In particular,

$$(*) \qquad \text{For all } x \text{ exactly one of } \quad x \in S \quad \text{and} \quad x \notin S \quad \text{holds.}$$

Not all collections of objects are sets. Suppose for example that the collection $\mathcal{B}$ of all sets is a set. Then $\mathcal{B} \in \mathcal{B}$. This is rather strange, but by itself not a contradiction. So lets make this example a little bit more complicated. We call a set S nice if $S \notin S$. Let $\mathcal{D}$ be the collection of all nice sets and suppose $\mathcal{D}$ is a set.

Is $\mathcal{D}$ a nice?

Suppose that $\mathcal{D}$ is a nice. Since $\mathcal{D}$ is the collection of all nice sets, $\mathcal{D}$ is a member of $\mathcal{D}$. Thus $\mathcal{D} \in \mathcal{D}$, but then by the definition of nice, $\mathcal{D}$ is not nice.

Suppose that $\mathcal{D}$ is not nice. Then by definition of nice, $\mathcal{D} \in \mathcal{D}$. Since $\mathcal{D}$ is the collection of nice sets, this means that $\mathcal{D}$ is nice.

We proved that $\mathcal{D}$ is nice if and only if $\mathcal{D}$ is not nice. This of course is absurd. So $\mathcal{D}$ cannot be a set.

Theorem 0.2.1. *Let A and B be sets. Then*
$$\Big(A = B\Big) \iff \Big(\text{for all } x : (x \in A) \iff (x \in B)\Big)$$

Proof. Naively this just says that two sets are equal if and only if they have the same members. In actuality this turns out to be one of the axioms of set theory. $\square$

Definition 0.2.2. *Let A and B be sets. We say that A is* subset *of B and write $A \subseteq B$ if*
$$\text{for all } x : (x \in A) \Longrightarrow (x \in B)$$

In other words, A is a subset of B if all the members of A are also members of B.

Theorem 0.2.3. *Let A and B sets. Then $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.*

Proof.
$$\begin{array}{lcl}
 & A = B & \\
\iff & x \in A \iff x \in B & - \text{ 0.2.1} \\
\iff & (x \in A \Longrightarrow x \in B) \text{ and } (x \in B \Longrightarrow x \in A) & - \text{ Rule of Logic: A.1.1(LR 19) } : \Big(P \iff Q\Big) \\
 & & \quad \iff \Big((P \Longrightarrow Q) \text{ and } (Q \Longrightarrow P)\Big) \\
\iff & A \subseteq B \text{ and } B \subseteq A & -\text{definition of subset}
\end{array}$$

□

Theorem 0.2.4. *Let x be an object. Then there exists a set, denote by $\{x\}$ such that*

$$t \in \{x\} \qquad \iff \qquad t = x$$

Proof. This is an axiom of Set Theory. □

Theorem 0.2.5. *Let S be a set and let $P(x)$ be a statement involving the variable x. Then there exists a set, denoted by $\{s \in S \mid P(s)\}$ such that*

$$t \in \{s \in S \mid P(s)\} \qquad \iff \qquad t \in S \text{ and } P(t)$$

Proof. This follows from the so called replacement axiom in set theory. □

Note that an object t is a member of $\{s \in S \mid P(s)\}$ if and only if t is a member of S and the statement $P(t)$ is true.

Example 0.2.6.

$$\{x \in \mathbb{Z} \mid x^2 = 1\} = \{1, -1\}.$$

$$\{x \in \mathbb{Z} \mid x > 0\} \quad \text{is the set of positive integers.}$$

Notation 0.2.7. *Let S be a set and $P(x)$ a statement involving the variable x.*

(a) "for all $x \in S : P(x)$" *is the statement*

$$\text{for all } x : \quad x \in S \Longrightarrow P(x)$$

(b) "there exists $x \in S : P(x)$" *is the statement*

$$\text{there exists } x : \quad x \in S \text{ and } P(x)$$

Example 0.2.8. (1) "for all $x \in \mathbb{R} : x^2 \geq 0$" is a true statement.

(2) "there exists $x \in \mathbb{Q} : x^2 = 2$" is a false statement.

Theorem 0.2.9. *Let S be a set and let $\Phi(x)$ be a formula involving the variable x such that $\Phi(s)$ is defined for all s in S. Then there exists a set, denoted by $\{\Phi(s) \mid s \in S\}$ such that*

$$t \in \{\Phi(s) \mid s \in S\} \qquad \iff \qquad \text{there exists } s \in S : \ t = \Phi(s)$$

Proof. This also follows from the replacement axiom in set theory. □

Note that the members of $\{\Phi(s) \mid s \in S\}$ are all the objects of the form $\Phi(s)$, where s is a member of S.

Example 0.2.10.

$$\{2x \mid x \in \mathbb{Z}\} \quad \text{is the set of even integers}$$

$$\Big\{x^3 \mid x \in \{-1,2,5\}\Big\} = \Big\{-1,8,125\Big\}$$

We now combine the two previous theorems into one:

Theorem 0.2.11. *Let S be a set, let $P(x)$ be a statement involving the variable x and $\Phi(x)$ a formula such that $\Phi(s)$ is defined for all s in S for which $P(s)$ is true. Then there exists a set, denoted by $\Big\{\Phi(s) \mid s \in S \text{ and } P(s)\Big\}$ such that*

$$t \in \Big\{\Phi(s) \mid s \in S \text{ and } P(s)\Big\} \qquad \Longleftrightarrow \qquad \text{there exists } s \in S : \Big(P(s) \text{ and } t = \Phi(s)\Big)$$

Proof. Define

$$(*) \qquad \Big\{\Phi(s) \mid s \in S \text{ and } P(s)\Big\} = \Big\{\Phi(s)) \mid s \in \{r \in S \mid P(r)\}\Big\}$$

Then

$$\begin{array}{cll}
 & t \in \Big\{\Phi(s) \mid s \in S \text{ and } P(s)\Big\} & \\
\Longleftrightarrow & t \in \Big\{\Phi(s) \mid s \in \{r \in S \mid \Phi(r)\}\Big\} & \text{By } (*) \\
\Longleftrightarrow & \text{there exists } s \in \{r \in S \mid P(r)\} \text{ with } t = \Phi(s) & 0.2.9 \\
\Longleftrightarrow & \text{there exists } s \text{ with } \Big(s \in \{r \in S \mid P(r)\} \text{ and } t = \Phi(s)\Big) & \text{definition of `there exists } s \in' \text{ see } 0.2.7 \\
\Longleftrightarrow & \text{there exists } s \text{ with } \Big(\big(s \in S \text{ and } P(s)\big) \text{ and } t = \Phi(s)\Big) & 0.2.5 \\
\Longleftrightarrow & \text{there exists } s \text{ with } \Big(s \in S \text{ and } \big(P(s) \text{ and } t = \Phi(s)\big)\Big) & \text{Rule of Logic: A.1.1(LR 24) :} \\
 & & \Big(P \text{ and } (Q \text{ and } R)\Big) \Longleftrightarrow \Big((P \text{ and } Q) \text{ and } R\Big) \\
\Longleftrightarrow & \text{there exists } s \in S \text{ with } \Big(P(s) \text{ and } t = \Phi(s)\Big) & \text{definition of `there exists } s \in' \text{ see } 0.2.7
\end{array}$$

□

Note that the members of $\{\Phi(s) \mid s \in S \text{ and } P(s)\}$ are all the objects of the form $\Phi(s)$, where s is a member of S for which $P(s)$ is true.

Example 0.2.12.

$$\{2n \mid n \in \mathbb{Z} \text{ and } n^2 = 1\} = \{2, -2\}$$

$$\{-x \mid x \in \mathbb{R} \text{ and } x > 0\} \qquad \text{is the set of negative real numbers}$$

Theorem 0.2.13. *Let A and B be sets.*

(a) *There exists a set, denoted by $A \cup B$ and called 'A union B', such that*

$$x \in A \cup B \qquad \Longleftrightarrow \qquad x \in A \text{ or } x \in B$$

(b) *There exists a set, denoted by $A \cap B$ and called 'A intersect B', such that*

$$x \in A \cap B \qquad \Longleftrightarrow \qquad x \in A \text{ and } x \in B$$

(c) *There exists a set, denoted by $A \setminus B$ and called 'A removed B', such that*

$$x \in A \setminus B \qquad \Longleftrightarrow \qquad x \in A \text{ and } x \notin B$$

(d) *There exists a set, denoted by $\emptyset$ and called empty set, such that*

$$\text{for all } x: \qquad x \notin \emptyset$$

(e) *Let a and b be objects, then there exists a set, denoted by $\{a, b\}$, that*

$$x \in \{a, b\} \qquad \Longleftrightarrow \qquad x = a \text{ or } x = b$$

Proof. (a) This is another axiom of set theory.

(b) Applying 0.2.5 with $P(x)$ being the statement "$x \in B$" we can define

$$A \cap B := \{x \in A \mid x \in B\}$$

(c) Applying 0.2.5 with $P(x)$ being the statement "$x \notin B$" we can define

$$A \setminus B := \{x \in A \mid x \notin B\}$$

(d) One of the axioms of set theory implies the existence of a set A. Then we can define

$$\emptyset := A \setminus A$$

(e) Define $\{a, b\} := \{a\} \cup \{b\}$. Then

$$\begin{array}{clc} & x \in \{a, b\} & \\ \Longleftrightarrow & x \in \{a\} \cup \{b\} & -\text{ definition of } \{a, b\} \\ \Longleftrightarrow & x \in \{a\} \text{ or } x \in \{b\} & -\text{(a)} \\ \Longleftrightarrow & x = a \text{ or } x = b & -0.2.4 \end{array}$$

□

Exercises 0.2:

#1. Let A be a set. Prove that $\emptyset \subseteq A$.

#2. Let A and B be sets. Prove that $A \cap B = B \cap A$.

#3. List all elements of the following sets:

(a) $\{x \in \mathbb{Q} \mid x^2 - 3x + 2 = 0\}$.

(b) $\{x \in \mathbb{Z} \mid x^2 < 5\}$.

(c) $\{x^3 \mid x \in \mathbb{Z} \text{ and } x^2 < 5\}$.

0.3 Relations and Functions

Definition 0.3.1. *Let a, b and c be objects.*

(a) $(a, b) := \big\{\{a\}, \{a, b\}\big\}$. *$(a, b)$ is called the (ordered) pair formed by a and b. a is called the first coordinate of (a, b) and b the second coordinate of (a, b).*

(b) $\big(a, b, c\big) := \big((a, b), c\big)$. *$(a, b, c)$ is called the (ordered) triple formed by a, b and c.*

Theorem 0.3.2. *Let a, b, c,d,e and f be objects.*

(a) $\Big((a, b) = (c, d)\Big) \quad \iff \quad \Big(a = c \text{ and } b = d\Big)$.

(b) $\Big((a, b, c) = (d, e, f)\Big) \quad \iff \quad \Big(a = d \text{ and } b = e \text{ and } c = f\Big)$

Proof. (a): See Exercise 0.3.#1.
(b)

$$\begin{aligned}
& (a, b, c) = (d, e, f) & \\
\iff \quad & \big((a, b), c\big) = \big((d, e), f\big) & \text{– definition of triple} \\
\iff \quad & (a, b) = (d, e) \text{ and } (c, f) & \text{– Part (a) of this theorem} \\
\iff \quad & a = d \text{ and } b = e \text{ and } e = f & \text{– Part (a) of this theorem}
\end{aligned}$$

□

Theorem 0.3.3. *Let A and B be sets. Then there exists a set, denoted by $A \times B$, such that*

$$x \in A \times B \qquad \iff \qquad \text{there exist } a \in A \text{ and } b \in B \textit{ with } x = (a, b)$$

Proof. This can be deduced from the axioms of set theory. □

Example 0.3.4. Let $A = \{1, 2\}$ and $B = \{2, 3, 5\}$. Then

$$A \times B = \big\{(1,2), (1,3), (1,5), (2,2), (2,3), (2,5)\big\}$$

Definition 0.3.5. *Let A and B be sets.*

(a) *A* relation *R from A to B is a triple (A, B, T), such that T is a subset of $A \times B$. Let a and b be objects. We say that a is in R-relation to b and write aRb if $(a, b) \in T$. So aRb is a statement and*

$$aRb \text{ if and only if } (a, b) \in T.$$

(b) *Let $R = (A, B, T)$ be a relation.*

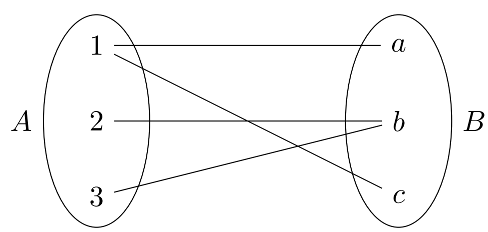
$$\begin{aligned} \operatorname{Dom} R &:= A \\ \operatorname{CoDom} R &:= B \\ \operatorname{Im} R &:= \{b \in B \mid \text{there exists } a \in A \textit{ with } aRb\} \\ \operatorname{CoIm} R &:= \{a \in A \mid \text{there exists } b \in B \textit{ with } aRb\} \end{aligned}$$

(c) *A relation on A is a relation from A to A.*

Example 0.3.6. (1) Using our formal definition of a relation, the familiar relation $\leq$ on the real numbers, would be the triple

$$\Big(\mathbb{R}, \mathbb{R}, \{(a, b) \in \mathbb{R} \times \mathbb{R} \mid a \leq b\}\Big)$$

(2) Let $A = \{1, 2, 3\}$, $B = \{a, b, c\}$, $T = \{(1, a), (1, c), (2, b), (3, b)\}$. Then the relation $\sim := (A, B, T)$ can be visualized by the following diagram:



Also $1 \sim 1$ is a true statement, $1 \sim b$ is a false statement, $2 \sim a$ is false statement, and $2 \sim b$ is a true statement.

Definition 0.3.7. (a) *A function from A to B is a relation F from A to B such that for all $a \in A$ there exists a unique b in B with aFb. We denote this unique b by $F(a)$ (or by Fa). So*

$$\text{for all } a \in A \text{ and } b \in B: \qquad b = F(a) \iff aFb$$

$F(a)$ is called the image of a under F. If $b = F(a)$ we will say that F maps a to b.

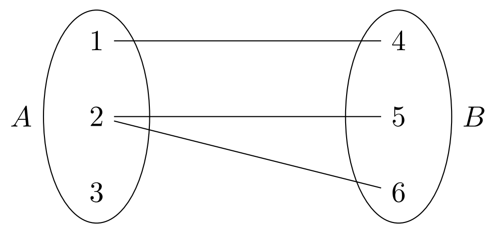
(b) *We write "$F : A \to B$ is function" for "A and B are sets and F is a function from A to B"'.*

(c) *Let $F : A \to B$ be a function and C a subset of A. Then $F[C] := \{F(c) \mid c \in C\}$.*

Example 0.3.8. (a) $F = \big(\mathbb{R}, \mathbb{R}, \{(x, x^2) \mid x \in \mathbb{R}\}\big)$ is a function with $F(x) = x^2$ for all $x \in \mathbb{R}$.

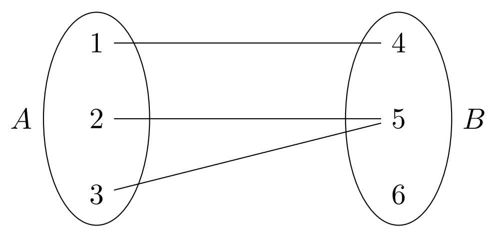
(b) $F = \big(\mathbb{R}, \mathbb{R}, \{(x^2, x^3) \mid x \in \mathbb{R}\}\big)$ is the relation with $x^2 F x^3$ for all $x \in \mathbb{R}$. For $x = 1$ we see that $1F1$ and for $x = -1$ we see that $1F - 1$. So F is not a function.

(c) Let $A = \{1, 2, 3\}$, $B = \{4, 5, 6, \}$, $T = \{(1, 4), (2, 5), (2, 6)\}$ and $R = (A, B, T)$:



Then R is not a function from A to B. Indeed, there does not exist an element b in R with $1Rb$. Also there exists two elements b in B with $2Rb$ namely $b = 5$ and $b = 6$.

(d) Let $A = \{1, 2, 3\}$, $B = \{4, 5, 6, \}$, $S = \{(1, 4), (2, 5), (3, 5)\}$ and $F = (A, B, T)$:



Then F is the function from A to B with $F(1) = 4$, $F(2) = 5$ and $F(3) = 5$.

Notation 0.3.9. *A and B be sets and suppose that $\Phi(x)$ is a formula involving a variable x such that for all x in A*

$$\Phi(a) \textit{ is defined} \quad \textit{and} \quad \Phi(a) \in B.$$

Put $T = \{\big(a, \Phi(a)\big) \mid a \in A\}$ and $F = (A, B, T)$. Then F is a function from A to B. We denote this function by

$$F: \ A \to B, \quad a \ \to \ \Phi(a).$$

So F is a function from A to B and $F(a) = \Phi(a)$ for all $a \in A$.

Example 0.3.10. (1) $F: \ \mathbb{R} \to \mathbb{R}, \ r \to r^2$ denotes the function from $\mathbb{R}$ to $\mathbb{R}$ with $F(r) = r^2$ for all $r \in \mathbb{R}$.

(2) $F:\ \mathbb{R} \to \mathbb{R},\ x \to \frac{1}{x}$ is not a function, since $\frac{1}{0}$ is not defined.

(3) $F:\ \mathbb{R} \setminus \{0\} \to \mathbb{R},\ x \to \frac{1}{x}$ is a function.

Theorem 0.3.11. *Let $f : A \to B$ and $g : C \to D$ be functions. Then $f = g$ if and only if $A = C$, $B = D$ and $f(a) = g(a)$ for all $a \in A$.*

Proof. By definition of a function, $f = (A, B, R)$ and $g = (C, D, S)$ where $R \subseteq A \times B$ and $S \subseteq C \times D$. By 0.3.2(b) :

(∗) *$f = g$ if and only of $A = C$, $B = D$ and $R = S$.*

$\Longrightarrow$: If $f = g$, then the Principal of Substitution implies, $f(a) = g(a)$ for all $a \in A$. Also by (∗), $A = C$ and $B = D$.

$\Longleftarrow$: Suppose now that $A = C$, $B = D$ and $f(a) = g(a)$ for all $a \in A$. By (∗) it suffices to show that $R = S$.

Let $a \in A$ and $b \in B$.

$$\begin{array}{ccl}
 & (a,b) \in R & \\
\Longleftrightarrow & afb & -\text{definition of } afb \\
\Longleftrightarrow & b = f(a) & -\text{the definition of } f(a) \\
\Longleftrightarrow & b = g(a) & -\text{since } f(a) = g(a) \\
\Longleftrightarrow & agb & -\text{definition of } g(a) \\
\Longleftrightarrow & (a,b) \in S & -\text{definition of } agb
\end{array}$$

Since $A = C$ and $B = D$, both R and S are subsets of $A \times B$. Hence each element of R and S is of the form $(a, b), a \in A, b \in B$. It follows that $x \in R$ if and only if $x \in S$ and so $R = S$ by 0.2.1. □

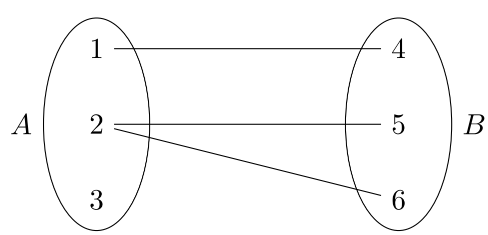
Definition 0.3.12. *Let R be a relation from A to B,*

(a) *R is called* 1-1 *(or* injective*) if for all $b \in B$ there exists at most one a in A with aRb.*

(b) *R is called* onto *(or* surjective*) if for all $b \in B$ there exists at least one $a \in A$ with aRb.*

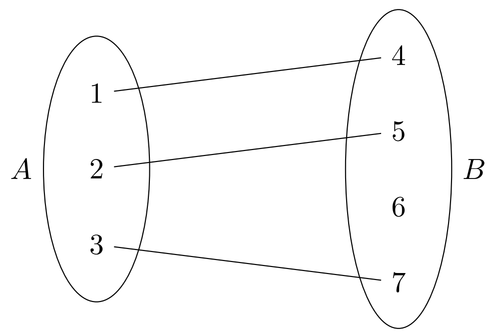
(c) *R is called a* 1-1 correspondence *(or* bijective*) if for all $a \in A$ there exists a unique $b \in B$ with aRb and for all $d \in B$ there exists a unique $c \in A$ with cRd*

Example 0.3.13. (1) The relation



is 1-1 and onto, but its is neither a function nor a 1-1 correspondence.

(2) The relation



is a 1-1 function, but is neither onto nor a 1-1 correspondence.

Lemma 0.3.14. (a) *Let f be a relation from A to B. Then f is a 1-1 correspondence if and only if f is a 1-1 and onto function.*

(b) *Let $f : A \to B$ be a function. Then f is 1-1 if and only*

$$\textit{For all } a, c \in A: \qquad f(a) = f(c) \implies a = c$$

(c) *A relation f from A to B is onto if and only if* $\operatorname{Im} f = B$.

Proof. (a)

	f is a 1-1 correspondence	
$\iff$	for all $a \in A$ there exists a unique $b \in B$ with afb, and for all $d \in B$ there exists a unique $c \in A$ with cfd	- Definition of 1-1 correspondence
$\iff$	f is a function, and for all $d \in B$ there exists a unique $c \in A$ with cfd	- Definition of a function
$\iff$	f is a function, and for all $d \in B$ there exists at most one $c \in A$ with cfd, and for all $d \in B$ there exists at least one $c \in A$ with cfd	- 0.1.7
$\iff$	f is a 1-1 and onto function	- Definition of 1-1 and onto

(b)

$$
\begin{array}{llll}
 & & f \text{ is 1-1} & \\
\iff & \text{for all } b \in B: & \text{there exists at most one } a \in A \text{ with } afb & \text{- definition of 1-1} \\
\iff & \text{for all } b \in B: & \text{there exists at most one } a \in A \text{ with } b = f(a) & \text{- definition of } f(a) \\
\iff & \text{for all } b \in B, a, c \in A: & b = f(a) \text{ and } b = f(c) \Longrightarrow a = c & \text{- definition of ``exists at most one''} \\
\iff & \text{for all } a, c \in A: & f(a) = f(c) \Longrightarrow a = c &
\end{array}
$$

(c) By definition of $\operatorname{Im} f$:

$$\operatorname{Im} f = \{b \in B \mid \text{ there exists } a \in A : afb\}.$$

Hence by 0.2.5

$$(*) \qquad b \in \operatorname{Im} f \quad \iff \quad b \in B \text{ and } \text{ there exists } a \in A : afb$$

Thus $b \in \operatorname{Im} f$ implies $b \in B$ and so $\operatorname{Im} f \subseteq B$. Thus

$(**)$ $B = \operatorname{Im} f$ *if and only if* $B \subseteq \operatorname{Im} f$.

We have

$$
\begin{array}{lll}
 & B = \operatorname{Im} f & \\
\iff & B \subseteq \operatorname{Im} f & \text{- } (**) \\
\iff & b \in B \Longrightarrow b \in \operatorname{Im} f & \text{- Definition of subset} \\
\iff & \text{for all } b \in B : b \in \operatorname{Im} f & \text{- Definition of ''for all } b \in B\text{''} \\
\iff & \text{for all } b \in B : \big(b \in B \text{ and } \text{ there exists } a \in A : afb\big) & \text{- } (*) \\
\iff & \text{for all } b \in B : \text{ there exists } a \in A : afb & \\
\iff & f \text{ is onto} & - \text{ definition of onto}
\end{array}
$$

□

Definition 0.3.15. (a) *Let A be a set. The* identity function id_A *on A is the function*

$$\mathrm{id}_A : A \to A, a \to a$$

So $\mathrm{id}_A(a) = a$ for all $a \in A$.

(b) *Let $f : A \to B$ and $g : B \to C$ be function. Then $g \circ f$ is the function*

$$g \circ f : A \to C, a \to g(f(a)$$

So $(g \circ f)(a) = g(f(a))$ for all $a \in A$.

Exercises 0.3:

#1. Let a, b, c, d be objects. Prove that

$$\Big((a,b) = (c,d)\Big) \iff \Big((a = c) \text{ and } (b = d)\Big)$$

#2. Give an example of an 1-1 and onto relation which is not a function.

#3. Let $F = (A, B, R)$ be a relation. Put

$$S = \{(b,a) \in B \times A \mid (a,b) \in R\} \text{ and } G = (B, A, S)$$

Note that G a relation from B and A. Also, if $a \in A$ and $b \in B$, then bGa if and only if aFb. Show that F is a function if and only if G is 1-1 and onto.

#4. Let A and B be sets. Let A_1 and A_2 be subsets of A and B_1 and B_2 subsets of B such that $A = A_1 \cup A_2, A_1 \cap A_2 = \emptyset$, $B = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$. Let $\pi_1 : A_1 \to B_1$ and $\pi_2 : A_2 \to B_2$ be bijections.(Recall that a bijection is a 1-1 and onto function.) Define

$$\pi : A \to B, a \to \begin{cases} \pi_1(a) & \text{if } a \in A_1 \\ \pi_2(a) & \text{if } a \in A_2 \end{cases}$$

Show that π is a bijection.

#5. Prove that the given function is injective

(a) $f : \mathbb{Z} \to \mathbb{Z}, f(x) = 2x$.

(b) $f : \mathbb{R} \to R, f(x) = x^3$.

(c) $f : \mathbb{Z} \to \mathbb{Q}, f(x) = \frac{x}{7}$.

(d) $f : \mathbb{R} \to \mathbb{R}, f(x) = -3x + 5$.

#6. Prove that the given function is surjective.

(a) $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$.

(b) $f : \mathbb{Z} \to \mathbb{Z}, f(x) = x - 4$.

(c) $f : \mathbb{R} \to \mathbb{R}, f(x) = -3x + 5$.

(d) $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$, $f(a,b) = \frac{a}{b}$ when $b \neq 0$ and $f(a,b) = 0$ when $b = 0$.

#7. (a) Let $f : B \to C$ and $g : C \to D$ be functions such that $g \circ f$ is injective. Prove that f is injective.

(b) Give an example of the situation in part (a) in which g is not injective.

0.4 The Natural Numbers and Induction

A *natural number* is a non-negative integer. $\mathbb{N}$ denotes the set of all natural numbers. So

$$\mathbb{N} = \{0, 1, 2, 3 \ldots\}$$

We do assume that familiarity with the basic properties of the natural numbers, like addition, multiplication and the order relation '$\leq$'.

A quick remark how to construct the natural numbers:

$$\begin{aligned}
0 &= \emptyset \\
1 &= \{0\} &&= 0 \cup \{0\} \\
2 &= \{0, 1\} &&= 1 \cup \{1\} \\
3 &= \{0, 1, 2\} &&= 2 \cup \{2\} \\
4 &= \{0, 1, 2, 3\} &&= 3 \cup \{3\} \\
&\vdots \\
n + 1 &= \{0, 1, 2, 3, \ldots, n\} &&= n \cup \{n\} \\
&\vdots
\end{aligned}$$

The relation $\leq$ on $\mathbb{N}$ can be defined by $i \leq j$ if $i \subseteq j$.

Definition 0.4.1. *Let S be a subset of $\mathbb{N}$. Then s is called a* minimal element *of S if $s \in S$ and $s \leq t$ for all $t \in S$.*

The following property of the natural numbers is part of our assumed properties of the integers and natural numbers (see Appendix C).

Well-Ordering Axiom*: Let S be a non-empty subset of $\mathbb{N}$. Then S has a minimal element*

Using the Well-Ordering Axiom we now provide an important tool to prove statements which hold for all natural numbers:

Theorem 0.4.2 (Principal Of Mathematical Induction)**.** *Suppose that for each $n \in \mathbb{N}$ a statement $P(n)$ is given and that:*

(i) *$P(0)$ is true.*

(ii) *If $P(k)$ is true for some $k \in \mathbb{N}$, then also $P(k+1)$ is true.*

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. Suppose for a contradiction that $P(n_0)$ is false for some $n_0 \in \mathbb{N}$. Put

$$(*) \qquad S := \{s \in \mathbb{N} \mid P(s) \text{ is false}\}$$

Then $n_0 \in S$ and so S is not empty. The Well-Ordering Axiom C.4.2 now implies that S has a minimal element m. Hence, by definition of a minimal element

$$(**) \qquad m \in S \quad \text{and} \quad m \leq s \text{ for all } s \in S$$

By (i) $P(0)$ is true and so $0 \notin S$ and $m \neq 0$. Thus $k := m - 1$ is a non-negative integer and $k < m$. If $k \in S$, then $(**)$ gives $m \leq k$, a contradiction. Thus $k \notin S$. By definition of S this means that $P(k)$ is true. So by (ii), $P(k+1)$ is true. But $k+1 = (m-1)+1 = m$ and so $P(m)$ is true. But $m \in S$ and so $P(m)$ is false. This contradiction show that $P(n)$ is true for all $n \in \mathbb{N}$. □

Theorem 0.4.3. *Let $n \in \mathbb{N}$ and S be a set with exactly n elements. Then S has exactly 2^n subsets.*

Proof. For $n \in \mathbb{N}$, let $P(n)$ be the statement

$P(n)$: If S is a set with exactly n elements, then S has exactly 2^n subsets.

If $n = 0$, then $S = \emptyset$. So S has exactly one subset, namely $\emptyset$. Since $2^0 = 1$ we see that $P(0)$ holds.

Now suppose that $P(k)$ holds and let S be a set with $k+1$ elements. Fix $s \in S$ and put $T = S \setminus \{s\}$. Then T is a set with k elements.

Let $A \subseteq S$. Then either $s \in A$ or $s \notin A$ but not both.

Suppose that $s \notin A$. Then $A \subseteq T$. By the induction assumption, T has 2^k subsets and so there are 2^k subsets of A with $s \notin A$.

Suppose that $s \in A$. Then $A = \{s\} \cup B$ for a unique subset B of T, namely $B = A \setminus \{s\}$. By the induction assumption there are 2^k choices for B and so there exists 2^k subsets of S with $s \in A$.

Since the number of subsets of A is the number of subsets of A not containing s plus the number of subsets of A containing s we conclude that A has $2^k + 2^k = 2^{k+1}$ subsets. Thus $P(k+1)$ holds.

We proved that $P(0)$ holds and that $P(k)$ implies $P(k+1)$ and so by the Principal Of Induction, $P(n)$ holds for all $n \in \mathbb{N}$. □

Theorem 0.4.4 (Principal Of Complete Induction)**.** *Suppose that for each $n \in \mathbb{N}$ a statement $P(n)$ is given and that*

(i) *If $k \in \mathbb{N}$ and $P(i)$ is true for all $i \in \mathbb{N}$ with $i < k$, then $P(k)$ is true.*

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. Let $Q(n)$ be the statement that $P(i)$ is true for all $i \in \mathbb{N}$ with $i < n$. Since there does not exits $i \in \mathbb{N}$ with $i < 0$ we have

$(*)$ *$Q(0)$ is true.*

Suppose now that $Q(k)$ is true, that is $P(i)$ is a true for all $i \in \mathbb{N}$ with $i < k$. Then by (i), also $P(k)$ is true. Hence $P(i)$ is for all i in $\mathbb{N}$ with $i < k+1$. Thus $Q(k+1)$ is true. We proved

$(**)$ *If $Q(k)$ is true for some $k \in \mathbb{N}$, then also $Q(k+1)$ is true.*

By $(*)$ and $(**)$ the assumptions of the Principal of Mathematical Induction are fulfilled. Hence $Q(n)$ is true for all $n \in \mathbb{N}$. Let $n \in \mathbb{N}$. Then $Q(n+1)$ is true and since $n < n+1$, $P(n)$ is true. □

One last version of the induction principal:

Theorem 0.4.5. *Suppose $r \in \mathbb{Z}$ and for all $n \in \mathbb{Z}$ with $n \geq r$, a statement $P(n)$ is given. Also assume that one of the following statements holds:*

(1) *$P(r)$ is true, and if $k \in \mathbb{Z}$ such that $k \geq r$ and $P(k)$ is true,then $P(k+1)$ is true.*

(2) *If $k \in \mathbb{Z}$ with $k \geq r$ and $P(i)$ holds for all $i \in \mathbb{Z}$ with $r \leq i < k$, then $P(k)$ holds.*

Then $P(n)$ holds for all $n \in \mathbb{Z}$ with $n \geq r$.

Proof. For $n \in \mathbb{N}$ let $Q(n)$ be the statement $P(n+r)$. If (1) holds we can apply 0.4.2 to $Q(n)$ and if (2) holds we can apply 0.4.4 to $Q(n)$. In both cases we conclude that $Q(n)$ holds for all $n \in \mathbb{N}$. So $P(n+r)$ holds for all $n \in \mathbb{N}$ and $P(n)$ holds for all $n \in \mathbb{Z}$ with $n \geq r$. □

Exercises 0.4:

#1. Prove that the sum of the first n positive integers is $\frac{n(n+1)}{2}$.
Hint: Let $P(k)$ be the statement:

$$1 + 2 + \ldots + k = \frac{k(k+1)}{2}.$$

#2. Let r be a real number, $r \neq 1$. Prove that for every integer $n \geq 1$,

$$1 + r + r^2 + \ldots r^{n-1} = \frac{r^n - 1}{r - 1}.$$

#3. Prove that for every positive integer n there exists an integer k with $2^{2n+1} + 1 = 2k$

#4. Let B be a set of n elements.

(a) If $n \geq 2$, prove that the number of two-elements subsets of B is $n(n-1)/2$.

(b) If $n \geq 3$, prove that the number of three-element subsets of B is $n(n-1)(n-2)/3!$.

#5. What is wrong with the following proof that all roses have the same color:

For a positive integer n let $P(n)$ be the statement:

Let A be a set containing n roses. Then all roses in A have the same color.

If $n = 1$, then A only contains on rose and so certainly all roses in A have the same color. Thus $P(1)$ is true.

Suppose now that $P(k)$ is true, that is whenever B is a set of k roses then all roses in B have the same color. We need to show that $P(k+1)$ is true. So let A be any set of $k+1$-roses. Let x and y be distinct roses in A. Consider the set $X = A \setminus \{x\}$ (that is the set of roses in A different from x). Then X is set of k roses. By the induction assumption $P(k)$ is true and so all roses in X have the same color. Similarly let $Y = A \setminus \{y\}$, then all roses in Y have the same color. Now let z be a rose in A distinct from x and y. Since z is distinct from x, $z \in X$; and since z is distinct from y, $z \in Y$. We will show that all roses in A have the same color as z. Indeed let a be any rose in A. If $a \neq x$, then both a and z are in X and so a has the same color as z. If $a = x$ then both a and z are in Y and so again a and z have the same color. We proved that all roses in A have the same color as z. Thus $P(k+1)$ is true.

We proved that $P(1)$ is true and that $P(k)$ implies $P(k+1)$. Hence by the Principal of Mathematical Induction, $P(n)$ is true for all n. Thus in any finite set of roses all the roses have the same color. So all roses have the same color.

#6. Let x be a real number greater than -1. Prove that for every positive integer n, $(1+x)^n \geq 1+nx$.

0.5 Equivalence Relations

Definition 0.5.1. *Let $\sim$ be a relation on a set A (that is a relation from A and A). Then*

(a) $\sim$ *is called* reflexive *if $a \sim a$ for all $a \in A$.*

(b) $\sim$ *is called* symmetric *if $b \sim a$ for all $a, b \in A$ with $a \sim b$, that is if*

$$a \sim b \qquad \Longrightarrow \qquad b \sim a.$$

(c) $\sim$ *is called* transitive *if $a \sim c$ for all $a, b, c \in A$ with $a \sim b$ and $b \sim c$, that is if*

$$(a \sim b \quad \text{and} \quad b \sim c) \qquad \Longrightarrow \qquad a \sim c$$

(d) $\sim$ *is called an* equivalence relation *if $\sim$ is reflexive,symmetric and transitive.*

Example 0.5.2. (1) Consider the relation " $\leq$ " on the real numbers:

$a \leq a$ for all real numbers a and so " $\leq$ " is reflexive.

$1 \leq 2$ but $2 \not\leq 1$ and so " $\leq$ " is not symmetric.

If $a \leq b$ and $b \leq c$, then $a \leq c$ and so " $\leq$ " is transitive.

Since " $\leq$ " is not symmetric, " $\leq$ " is not an equivalence relation.

(2) Consider the relation " $=$ " on any set A.

$a = a$ and so " $=$ " is reflexive.

If $a = b$, then $b = a$ and so " $=$ " is symmetric.

If $a = b$ and $b = c$, then $a = c$ and so " $=$ " is transitive.

" $=$ " is reflexive, symmetric and transitive and so an equivalence relation.

(3) Consider the relation " $\neq$ " on any set A.

$a \neq a$ and so if $A \neq \emptyset$, " $\neq$ " is not reflexive.

Suppose A has at least two distinct elements a, b. Then

$$a \neq b \quad \text{and} \quad b \neq a \qquad \text{but} \quad \text{not-}(a \neq a)$$

So " $\neq$ " is not transitive.

Definition 0.5.3. (a) *Let a, b be integers, then we say that a* divides *b and write $a|b$ if there exists an integer k with $b = ak$.*

(b) *Let n be an integers. Then the relation '$\equiv \pmod{n}$' on $\mathbb{Z}$ is defined by*

$$a \equiv b \pmod{n} \quad \iff \quad n \mid a - b$$

If $a \equiv b \pmod{n}$ we say that a is congruent *to b* modulo *n.*

Example 0.5.4. (1) $2|6$, since $6 = 2 \cdot 3$. But $7 \nmid 31$,

(2) $6 \equiv 4 \pmod{2}$ is true since 2 divides $6 - 4$.

But $3 \equiv 8 \pmod{2}$ is false since 2 does not divide $3 - 8$. Thus $3 \not\equiv 8 \pmod{2}$.

If a and b are integers, then $a \equiv b \pmod{2}$ if and only if $b - a$ is even and so if and only if either both a and b are even, or both a and b are odd.

Hence $a \not\equiv b \pmod{2}$ if and only if one of a and b is even and the other is odd.

(3) Let a, b be integers. Then

$$\begin{aligned}
& a \equiv b \pmod{0} \\
\iff\quad & 0 \mid a - b \\
\iff\quad & a - b = 0 \cdot k \qquad \text{for some } k \in \mathbb{Z} \\
\iff\quad & a - b = 0 \\
\iff\quad & a = b
\end{aligned}$$

So congruent modulo 0 is the equality relation.

(4) Since $m = m \cdot 1$, 1 divides all integers. Thus $1 \mid b - a$ for all integers a and b and so

$$a \equiv b \pmod 1 \text{ for all } a, b \in \mathbb{Z}$$

Lemma 0.5.5. *Let $n \in \mathbb{Z}$. Then the relation "$\equiv \pmod n$" is an equivalence relation on $\mathbb{Z}$.*

Proof. We have to show that "$\equiv \pmod n$" is reflexive, symmetric and transitive. Let $a, b, c \in \mathbb{Z}$.

Reflexive: Since $a - a = 0 = 0 \cdot n$ we see that $n \mid a - a$ and so $a \equiv a \pmod n$. Thus "$\equiv \pmod n$" is reflexive.

Symmetric: Suppose that $a \equiv b \pmod n$. Then $n \mid (a - b)$ and so $a - b = nk$ for some $k \in \mathbb{Z}$. Thus $b - a = -(a - b) = -(nk) = n(-k)$. So $n \mid b - a$ and $b \equiv a \pmod n$. Thus "$\equiv \pmod n$" is symmetric.

Transitive: Suppose that $a \equiv b \pmod n$ and $b \equiv c \pmod n$. Then $n \mid a - b$ and $n \mid b - c$ and so there exist $k, l \in \mathbb{Z}$ with $a - b = nk$ and $b - c = nl$. Thus

$$a - c = (a - b) + (b - c) = nk + nl = n(k + l).$$

Hence $n \mid a - c$ and $a \equiv c \pmod n$. Thus "$\equiv \pmod n$" is transitive. □

Definition 0.5.6. *Let $\sim$ be an equivalence relation on the set A and let $n \in \mathbb{Z}$.*

(a) *For $a \in A$ we define $[a]_\sim := \{b \in A \mid a \sim b\}$. We often just write $[a]$ for $[a]_\sim$. $[a]_\sim$ is called the* equivalence class *of a with respect to $\sim$.*

(b) *$A/\sim := \{[a]_\sim \mid a \in A\}$. So $A/\sim$ is the set of equivalence classes with respect to $\sim$.*

(c) *Let $a \in \mathbb{Z}$. Then $[a]_n$ is the equivalence class a with respect to '$\equiv \pmod n$'. $[a]_n$ is called the congruence class of a modulo n.*

(d) *$\mathbb{Z}_n := \mathbb{Z}/'a \equiv b \pmod n'$. So $\mathbb{Z}_n = \{[a]_n \mid a \in \mathbb{Z}\}$ is the set of congruence classes modulo n.*

Example 0.5.7. (1) Consider the relation '$\equiv \pmod 2$':

$$[1]_2 = \{b \in \mathbb{Z} \mid 1 \equiv b \pmod 2\} = \{b \in \mathbb{Z} \mid b \text{ is odd}\}$$

and so $[1]_2$ is the set of odd integers.

$$[0]_2 = \{b \in \mathbb{Z} \mid 0 \equiv b \pmod 2\} = \{b \in \mathbb{Z} \mid b \text{ is even }\}$$

and so $[0]_2$ is the set of odd integers.

In general:

$$[a]_2 = \{b \in \mathbb{Z} \mid a \equiv b \pmod 2\} = \begin{cases} \{b \in \mathbb{Z} \mid b \text{ is even}\} & \text{if } a \text{ is even} \\ \{b \in \mathbb{Z} \mid b \text{ is odd }\} & \text{if } a \text{ is odd} \end{cases}.$$

So

$$\mathbb{Z}_2 = \Big\{\{n \in \mathbb{Z} \mid \text{n is even}\}, \{n \in \mathbb{Z} \mid \text{n is odd}\}\Big\} = \Big\{[0]_2, [1]_2\Big\}.$$

(2) Consider the relation '$\equiv \pmod 5$': We have

$$0 \equiv b \pmod 5 \iff 5 \mid b - 0 \iff 5 \mid b \iff b = 5k \text{ for some } k \in \mathbb{Z}$$

so

$$[0]_5 = \{b \in \mathbb{Z} \mid 0 \equiv b \pmod 5\} = \{5k \mid k \in \mathbb{Z}\} = \{0, 5, 10, 15, 20, \ldots, -5, -10, -15, -20, \ldots\}$$

Also

$$1 \equiv b \pmod 5 \iff 5 \mid b{-}1 \iff b{-}1 = 5k \text{ for some } k \in \mathbb{Z} \iff b = 5k{+}1 \text{ for some } k \in \mathbb{Z}$$

and so

$$[1]_5 = \{b \in \mathbb{Z} \mid 1 \equiv b \pmod 5\} = \{5k{+}1 \mid k \in \mathbb{Z}\} = \{1, 6, 11, 16, 21, \ldots, -4, -9, -14, -19, \ldots\}$$

Similarly,

$$\begin{aligned}
[2]_5 &= \{b \in \mathbb{Z} \mid 2 \equiv b \pmod 5\} = \{5k + 2 \mid k \in \mathbb{Z}\} = \{2, 7, 12, 17, 22, \ldots, -3, -8, -13, -18, \ldots\} \\
[3]_5 &= \{b \in \mathbb{Z} \mid 3 \equiv b \pmod 5\} = \{5k + 3 \mid k \in \mathbb{Z}\} = \{3, 8, 13, 18, 23, \ldots, -2, -7, -12, -17, \ldots\} \\
[4]_5 &= \{b \in \mathbb{Z} \mid 4 \equiv b \pmod 5\} = \{5k + 4 \mid k \in \mathbb{Z}\} = \{4, 9, 14, 19, 24, \ldots, -1, -6, -11, -16, \ldots\} \\
[5]_5 &= \{b \in \mathbb{Z} \mid 5 \equiv b \pmod 5\} = \{5k + 5 \mid k \in \mathbb{Z}\} = \{5, 10, 15, 20, 25, \ldots, 0, -5, -10, -15, \ldots\} = [0]_5 \\
[6]_5 &= \{b \in \mathbb{Z} \mid 6 \equiv b \pmod 5\} = \{5k + 6 \mid k \in \mathbb{Z}\} = \{6, 11, 16, 21, 26, \ldots, 1, -4, -9, -14, \ldots\} = [1]_5
\end{aligned}$$

So it seems that

$$\mathbb{Z}_5 = \{[0]_5, [1]_5, [2]_5, [3]_5, [4]_5\}.$$

Later (see 2.1.2(b)) we will give a rigorous proof for this.

(3) Consider the relation '$\equiv \pmod 0$). By 0.5.4 $a \equiv b \pmod 0$ if and only if $a = b$. So

$$[a]_0 = \{a\}$$

and

$$\mathbb{Z}_0 = \Big\{\{a\} \mid a \in \mathbb{Z}\Big\}$$

(4) By 0.5.4 $a \equiv b \pmod 1$ for all a, b. Thus
So

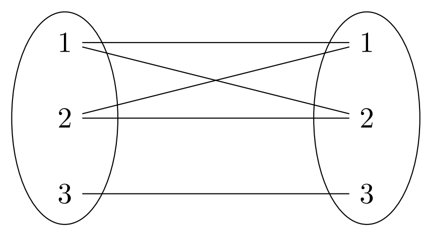
$$[a]_0 = \mathbb{Z}$$

and

$$\mathbb{Z}_1 = \left\{\mathbb{Z}\right\}$$

(5) Consider the relation

$\sim$:



on the set $A = \{1, 2, 3\}$. Then $\sim$ is an equivalence relation. Also

$$[1]_\sim = \{a \in A \mid 1 \sim a\} = \{1, 2\}, \qquad [2]_\sim = \{a \in A \mid 2 \sim a\} = \{1, 2\} \qquad [3]_\sim = \{a \in A \mid 3 \sim a\} = \{3\}$$

and so

$$A/\!\sim = \left\{\{1, 2\}, \{3\}\right\}$$

Theorem 0.5.8. *Let $\sim$ be an equivalence relation on the set A and $a, b \in A$. Then the following statements are equivalent:*

(a) $a \sim b$.

(b) $b \in [a]$.

(c) $[a] \cap [b] \neq \emptyset$.

(d) $[a] = [b]$.

(e) $a \in [b]$

(f) $b \sim a$.

Proof. (a) $\Longrightarrow$ (b): Suppose that $a \sim b$. Since $[a] = \{b \in A \mid a \sim b\}$ we conclude that $b \in [a]$.

(b) $\Longrightarrow$ (c): Suppose that $b \in [a]$. Since $\sim$ is reflexive, we get $b \sim b$ and so $b \in [b]$. Thus $b \in [a] \cap [b]$ and $[a] \cap [b] \neq \emptyset$.

(c) $\Longrightarrow$ (d): Suppose $[a] \cap [b] \neq \emptyset$. Then there exists $c \in [a] \cap [b]$.

We will first show that $[a] \subseteq [b]$. So let $d \in [a]$. Then $a \sim d$. Since $c \in [a]$ and $[a] = \{e \in A \mid a \sim e\}$ we have $a \sim c$ and since $\sim$ is symmetric we conclude that $c \sim a$. As $a \sim d$ and $\sim$ is transitive, this gives $c \sim d$. From $c \in [b]$ we get $b \sim c$. Since $c \sim d$ and $\sim$ is transitive, we infer that $b \sim d$ and so $d \in [b]$. Thus $[a] \subseteq [b]$.

A similar argument shows that $[b] \subseteq [a]$. We proved that $[a] \subseteq [b]$ and $[b] \subseteq [a]$ and so $[a] = [b]$ by 0.2.3

(d) $\implies$ (e): Since a is reflexive, $a \sim a$ and so $a \in [a]$. As $[a] = [b]$ we get $a \in [b]$.

(e) $\implies$ (f): By definition $[b] = \{e \in A \mid b \sim e\}$. Since $a \in [b]$ we conclude that $b \sim a$.

(f) $\implies$ (a): Since $\sim$ is symmetric, $b \sim a$ implies $a \sim b$. $\square$

Exercises 0.5:

#1. Let $f : A \to B$ be a function and define a relation $\sim$ on A by

$$u \sim v \iff f(u) = f(v).$$

Prove that $\sim$ is an equivalence relation.

#2. Let $A = \{1, 2, 3\}$. Use the definition of a relation (see 0.3.5(b)) to exhibit a relation on A with the stated properties.

(a) Reflexive, not symmetric, not transitive.

(b) Symmetric, not reflexive, not transitive.

(c) Transitive, not reflexive, not symmetric.

(d) Reflexive and symmetric, not transitive.

(e) Reflexive and transitive, not symmetric.

(f) Symmetric and transitive, not reflexive.

#3. Let $\sim$ be the relation on the set $\mathbb{R}^*$ of non-zero real numbers defined by

$$a \sim b \iff \frac{a}{b} \in \mathbb{Q}.$$

Prove that $\sim$ is an equivalence relation.

#4. Let $\sim$ be a symmetric and transitive relation on a set A. What is wrong with the following 'proof' that $\sim$ is reflexive.:

$a \sim b$ implies $b \sim a$ by symmetry; then $a \sim b$ and $b \sim a$ imply that $a \sim a$ by transitivity.

Chapter 1

Arithmetic in $\mathbb{Z}$

1.1 The Division Algorithm

Theorem 1.1.1 (The Division Algorithm)**.** *Let a and b be integers with $b > 0$. Then there exist unique integers q and r such that*

$$a = bq + r \qquad \textit{and} \qquad 0 \leq r < b.$$

Proof. We will first show that q and r exist. Put

$$S := \{a - bx \mid x \in \mathbb{Z} \text{ and } a - bx \geq 0\}$$

We would like to apply the well-ordering Axiom to S, so we need to verify that S is not empty. That is we need to find $x \in \mathbb{Z}$ such that $a - bx \geq 0$.

If $a \geq 0$, then $a - b0 = a > 0$ and we can choose $x = 0$.

So suppose $a < 0$. Let's try $x = a$. Then $a - bx = a - ba = (1 - b)a$. Since $b > 0$ and b is an integer, $b \geq 1$ and so $1 - b \leq 0$. Since $a < 0$, this implies $(1 - b)a \geq 0$ and so $a - bx \geq 0$. So we can indeed choose $x = a$.

We have proved that S is non-empty. Note that every element of S is a natural number and so $S \subseteq \mathbb{N}$. Hence by the Well-ordering Axiom C.4.2 S has a minimal element r. Thus

$$r \in S \qquad \text{and} \qquad r \leq s \text{ for all } s \in S.$$

Since $r \in S$, the definition of S implies that there exists $q \in \mathbb{Z}$ with $r = a - bq$. Then $a = bq + r$ and it remains to show $0 \leq r < b$. Since $r \in S$, $r \geq 0$. Suppose for a contradiction that $r \geq b$. Then $r - b \geq 0$. Hence

$$a - b(q + 1) = (a - bq) - b = r - b \geq 0$$

and $q + 1 \in \mathbb{Z}$. Thus $r - b \in S$. Since $b > 0$ we have $r - b < r$, but this is a contradiction since r is a minimal element of S.

This shows the existence of q and r. To show the uniqueness let $q, r, \tilde{q}$ and $\tilde{r}$ be integers with

$$\Big(a = bq + r \text{ and } 0 \le r < b\Big) \qquad \text{and} \qquad \Big(a = b\tilde{q} + \tilde{r} \text{ and } 0 \le \tilde{r} < b\Big).$$

We need to show that $q = \tilde{q}$ and $r = \tilde{r}$.

From $a = bq + r$ and $a = b\tilde{q} + \tilde{r}$ we have

$$bq + r = b\tilde{q} + \tilde{r}$$

and so

$$(*) \qquad b(q - \tilde{q}) = \tilde{r} - r.$$

Multiplying the equation $0 \le r < b$ with -1 gives $0 \ge -r > -b$ and so

$$-b < -r \le 0.$$

Adding the inequality

$$0 \le \tilde{r} < b$$

yields

$$-b < \tilde{r} - r < b$$

Using (*) we conclude

$$-b < -b(q - \tilde{q}) < b.$$

Since $b > 0$ we can divide by b and get

$$-1 < q - \tilde{q} < 1.$$

The only integer strictly between -1 and 1 is 0. Hence $q - \tilde{q} = 0$ and so $q = \tilde{q}$. Hence (*) gives $\tilde{r} - r = b(q - \tilde{q}) = b0 = 0$ and so also $\tilde{r} = r$. □

Corollary 1.1.2 (Division Algorithm). *Let a and c be integers with $c \neq 0$. Then there exist unique integers q and r such that*

$$a = cq + r \text{ and } 0 \le r < |c|.$$

Proof. See Exercise 1.1. #1 □

Definition 1.1.3. *Let a and b be integers with $b \neq 0$. Let q, r be the unique integers with $a = bq + r$ and $0 \le r < |b|$. Then r is called the* remainder *of a when divided by b and q is called the integral quotient of a when divided by b.*

Example 1.1.4. (1) $42 = 8 \cdot 5 + 2$ and $0 \le 2 < 8$. So the remainder of 42 when divided by 8 is 2.

(2) $-42 = 8 \cdot -6 + 6$ and $0 \le 6 < 8$. So the remainder of -42 when divided by 8 is 6.

Exercises 1.1:

#1. Let a and c be integers with $c \neq 0$. Proof that there exist unique integers q and r such that

$$a = cq + r \text{ and } 0 \leq r < |c|.$$

#2. Prove that the square of an integer is either of the form $3k$ or the form $3k + 1$ for some integer k.

#3. Use the Division Algorithm to prove that every odd integer is of the form $4k + 1$ or $4k + 3$ for some integer k.

#4. (a) Divide 5^2, 7^2, 11^2, 15^2 and 27^2 by 8 and note the remainder in each case.

(b) Make a conjecture about the remainder when the square of an odd number is divided by 8.

(c) Prove your conjecture.

#5. Prove that the cube of any integer has be exactly one of these forms: $9k$, $9k + 1$ or $9k + 8$ for some integer k.

1.2 Divisibility

Lemma 1.2.1. *Let a and b be integers.*

(a)

$$b \mid a \iff b \mid -a \iff -b \mid a \iff -b \mid -a$$

(b) *a and $-a$ have the same divisors.*

(c) *If $b \mid a$ and $a \neq 0$, then $1 \leq |b| \leq |a|$.*

(d) *If $a \neq 0$, then a has only finitely many divisors.*

Proof. (a) We will first show

$$(*) \qquad b \mid a \implies b \mid -a.$$

For this suppose that b divides a. Then by definition of "divide" there exists $k \in \mathbb{Z}$ with $a = kb$. Thus $-a = -(kb) = (-k)b$. Since $k \in \mathbb{Z}$ also $-k \in \mathbb{Z}$. Thus the definition of "divide" shows that b divides $-a$. So $(*)$ holds.

$$(**) \qquad b \mid -a \implies -b \mid a.$$

Suppose that b divides $-a$. Then by definition of "divide" there exists $k \in \mathbb{Z}$ with $-a = kb$. Thus $a = -(-a) = -(kb) = k(-b)$. Thus the definition of "divide" shows that $-b$ divides a. So $(**)$ holds.

$$(***) \qquad -b \mid a \quad \Longrightarrow \quad -b \mid -a$$

This is $(*)$ applied with $-b$ in place of b.

$$(+) \qquad -b \mid -a \quad \Longrightarrow \quad b \mid a.$$

By $(**)$ applied with $-b$ in place of b, if $-b \mid -a$ then $-(-b) \mid a$ and so $b \mid a$.

We proved

$$b \mid a \quad \Longrightarrow \quad b \mid -a \quad \Longrightarrow \quad -b \mid a \quad \Longrightarrow \quad -b \mid -a \quad \Longrightarrow \quad b \mid a$$

and so (a) holds.

(b) By (a) $b \mid a$ if and only if $b \mid -a$. So b is a divisor of a if and only if b is a divisor of $-a$.

(c) Suppose $a \neq 0$ and that $b \mid a$. Then $a = kb$ for some k in $\mathbb{Z}$. Since $0b = 0$ and $a \neq 0$ we have $k \neq 0$ and since k is an integer $|k| \geq 1$. Since $|b| \geq 0$ this gives $|k||b| \geq 1|b| = |b|$. Hence

$$b \leq |b| \leq |k||b| = |kb| = |a|.$$

Also since $a = kb$ and $a \neq 0$, $b \neq 0$ and so $|b| \geq 1$. Thus (c) is proved.

(d) Suppose $a \neq 0$ and let b be divisor of a. By (c), $|b| \leq |a|$ and so $-|a| \leq b \leq |a|$. Thus b is one of $-|a|, -|a|+1, -|a|+2, \ldots, -1, 0, 1, \ldots, |a|-1, |a|$ and so a has at most $2|a|+1$ divisors. □

Definition 1.2.2. *Let a, b and d be integers.*

(a) *d is called a common divisor of a and b provided that $d \mid a$ and $d \mid b$.*

(b) *d is called a greatest common divisor of a and b provided that*

  (i) *d is a common divisor of a and b; and*

  (ii) *if c is a common divisor of a and b then $c \leq d$.*

Example 1.2.3. (1) The largest integer dividing both 24 and 42 is 6. So 6 is the greatest common divisor of 24 and 42.

(2) All integers divide 0 and 0. So there does not exist a greatest common divisor of 0 and 0.

Lemma 1.2.4. *Let a and b be integers, not both 0. Then a and b have a unique greatest common divisor. We denote the unique greatest common divisor of a and b by* $\gcd(a, b)$.

Proof. We may assume that $a \neq 0$. Then by 1.2.1(d), a has only finitely many divisors. Thus a and b have only finitely many common divisors. Let $c_1, c_2, \ldots, c_n$ be the common divisors of a and b such that

$$c_1 < c_2 < c_3 < \ldots < c_n.$$

Then c_n is the unique greatest common divisor. □

Lemma 1.2.5. *Let a, b, c, u and v be integers and suppose that c is a common divisor of a and b. Then c divides $au + bv$. In particular, c divides $a + b$, au, $-au$, $a + bv$, $au - bv$ and $a - bv$.*

Proof. Since c is a common divisor of a and b we have $c \mid a$ and $c \mid b$. So by definition of 'divide' there exist $k, l \in \mathbb{Z}$ with $a = kc$ and $b = lc$. Thus

$$au + bv = (kc)u + (lcv) = (ku + lv)c$$

Since k, l, u and v are integers, also $ku + lv$ is an integer. So the definition of 'divide' shows that $c \mid au + bv$.

Choosing special values for u and v proves the second statement:

u	v	$au + bv$
1	1	$a + b$
u	0	au
$-u$	0	$-au$
1	v	$a + bv$
u	$-v$	$au - bv$
1	$-v$	$a - bv$

□

Lemma 1.2.6. *Let a, b, q and r be integers with $a \neq 0$ or $b \neq 0$ and $a = bq + r$. Then $b \neq 0$ or $r \neq 0$. Moreover, $\gcd(a, b) = \gcd(b, r)$.*

Proof. If $b = 0$ and $r = 0$ then also $a = bq + r = 0q + 0$, a contradiction to the hypothesis that $a \neq 0$ or $b \neq 0$. Thus $b \neq 0$ or $r \neq 0$.

In particular, both $\gcd(a, b)$ and $\gcd(b, r)$ exists. Put $d := \gcd(a, b)$ and $e := \gcd(b, r)$. Then d divides a and b and so by 1.2.5 d divides $r = a - bq$. Hence d is a common divisor of b and r. Thus $d \leq e$ by the definition of *gcd*.

Since $e = \gcd(b, r)$, e divides b and r. So by 1.2.5 e divides $a = bq + r$. Thus e is a common divisor of a and b and so $e \leq d$. We have proved $d \leq e$ and $e \leq d$ and so $e = d$. □

Theorem 1.2.7 (Euclidean Algorithm). *Let a and b be integers not both 0 and let E_{-1} and E_0 be the equations*

$$\begin{array}{lcccccc} E_{-1} & : & a & = & a1 & + & b0 \\ E_0 & : & b & = & a0 & + & b1 \end{array},$$

Let $i \in \mathbb{N}$ and suppose inductively we already defined equation $E_k, -1 \leq k \leq i$ of the form

$$E_k \;:\; r_k \;=\; ax_k \;+\; by_k \;.$$

Suppose $r_i \neq 0$ and let $t_{i+1}, q_{i+1} \in \mathbb{Z}$ with

$$r_{i-1} = r_i q_{i+1} + t_{i+1} \quad \text{and} \quad |t_{i+1}| < |r_i|.$$

(Note here that such t_{i+1}, q_{i+1} exist by the division algorithm 1.1.2)

Let E_{i+1} be the equation of the form $r_{i+1} = ax_{i+1} + by_{i+1}$ obtained by subtracting q_{i+1}-times equation E_i from E_{i-1}. Then there exists $m \in \mathbb{N}$ with $r_{m-1} \neq 0$ and $r_m = 0$. Put $d = |r_{m-1}|$. Then

(a) *$r_k, x_k, y_k \in \mathbb{Z}$ for all $k \in \mathbb{Z}$ with $-1 \leq k \leq m$.*

(b) *d is the greatest common divisor of a and b.*

(c) *$r_{m-1} = ax_{m-1} + by_{m-1}$ and $d = ax + by$ for some $x, y \in \mathbb{Z}$.*

Proof. For $k \in \mathbb{Z}$ with $k \geq -1$, let $P(k)$ be the statement that r_k, x_k and y_k are integers and if $k \geq 1$, then $|r_k| < |r_{k-1}|$.

By the definition of E_0 and E_1 we have $r_{-1} = a, x_{-1} = 1, y_{-1} = 0, r_0 = b, x_0 = 0$ and $y_0 - 1$. Thus $P(-1)$ and $P(0)$ hold. Suppose now that $i \in \mathbb{N}$, that $P(k)$ holds for all $k \in \mathbb{Z}$ with $-1 \leq k \leq i$ and that $r_i \neq 0$. We have

$$\begin{array}{lcccccc} E_{i-1} & : & r_{i-1} & = & ax_{i-1} & + & by_{i-1} \\ E_i & : & r_i & = & ax_i & + & by_i. \end{array}$$

and subtracting q_{i+1} times E_i from E_{i-1} we obtain

$$E_{i+1} \;:\; r_{i-1} - r_i q_{i+1} \;=\; a(x_{i-1} - x_i q_{i+1}) \;+\; b(y_{i-1} - x_i q_{i+1}).$$

Hence

$$\begin{aligned} r_{i+1} &= r_{i-1} - r_i q_{i+1} \\ x_{i+1} &= x_{i-1} - x_i q_{i+1} \\ y_{i+1} &= y_{i-1} - x_i q_{i+1}. \end{aligned}$$

By choice, q_{i+1} is an integer. By the induction assumption, x_i, x_{i-1}, y_{i-1} and y_i are integers. Hence also r_{i+1}, x_{i+1} and y_{i+1} are integers. By choice of q_{i+1} and t_{i+1}

$$r_{i-1} = r_i q_{i+1} + t_{i+1} \quad \text{and} \quad |t_{i+1}| < |r_i|.$$

So

$$t_{i+1} = r_i q_{i+1} - r_{i-1} = r_{i+1} \text{ and } \quad |r_{i+1}| < |r_i|.$$

Hence $P(i+1)$ holds. So by the principal of complete induction, $P(n)$ holds for all $n \in \mathbb{Z}$ with $n \geq -1$ (for which E_n is defined).

In particular, (a) holds and

$$|r_0| > |r_1| > |r_2| > |r_3| > \ldots > |r_i| > \ldots.$$

Since the r_i's are integers, we conclude that there exists $m \in \mathbb{N}$ with $r_{m-1} \neq 0$ and $r_m = 0$.

From $r_{i-1} = r_i q_{i+1} + t_{i+1} = r_i q_{i+1} + r_{i+1}$ and 1.2.6 we have $\gcd(r_{i-1}, r_i) = \gcd(r_i, r_{i+1})$ and so

$$\gcd(a,b) = \gcd(r_{-1}, r_0) = \gcd(r_0, r_1) = \ldots = \gcd(r_{m-1}, r_m) = \gcd(r_{m-1}, 0) = |r_{m-1}| = d.$$

So (b) holds.

The first statement in (c) is the equation E_{m-1}. If $r_{m-1} > 0$, then $d = r_{m-1} = a x_{m-1} + b y_{m-1}$ and if $r_{m-1} < 0$, then $d = -r_{m-1} = a(-x_{m-1}) + b(-y_{m-1})$ and so (c) holds. $\square$

Example 1.2.8. Let $a = 1492$ and $b = 1066$. Then

$$\begin{array}{lrclcrclcrlcrcr}
E_{-1}: & 1492 & = & 1492 & \cdot & 1 & + & 1066 & \cdot & 0 & & & & \\
E_0: & 1066 & = & 1492 & \cdot & 0 & + & 1066 & \cdot & 1 & & & & \\
E_1: & 426 & = & 1492 & \cdot & 1 & + & 1066 & \cdot & -1 & & \mid E_{-1} & - & E_0 \\
E_2: & 214 & = & 1492 & \cdot & -2 & + & 1066 & \cdot & 3 & & \mid E_0 & - & 2E_1 \\
E_3: & 212 & = & 1492 & \cdot & 3 & + & 1066 & \cdot & -4 & & \mid E_1 & - & E_2 \\
E_4: & 2 & = & 1492 & \cdot & -5 & + & 1066 & \cdot & 7 & & \mid E_2 & - & E_3 \\
E_5: & 0 & & & & & & & & & & \mid E_3 & - & 106E_4
\end{array}$$

So $\gcd(1492, 1066) = 2$ and $2 = 1492 \cdot -5 + 1066 \cdot 7$.

Theorem 1.2.9. *Let a and b be integers not both zero and $d := \gcd(a,b)$. Then d is the smallest positive integer of the form $au + bv$ with $u, v \in \mathbb{Z}$.*

Proof. By the Euclidean Algorithm 1.2.7 d is of the form $au + bv$ with $u, v \in \mathbb{Z}$. Now let e be any positive integer of the form $e = au + bv$ for some $u, v \in \mathbb{Z}$. Since $d = \gcd(a,b)$, d divides a and b. Thus by 1.2.5, d divides $au + bv = e$. Hence 1.2.1(c) shows that $d \leq |d| \leq |e| = e$. Thus d is the smallest possitive integer of the form $au + bv$ with $u, v \in \mathbb{Z}$. $\square$

Corollary 1.2.10. *Let a and b be integers not both 0 and d a positive integer. Then d is the greatest common divisor of a and b if and only if*

(I) *d is a common divisor of a and b; and*

(II) *if c is a common divisor of a and b, then $c \mid d$.*

Proof. $\Longrightarrow$: Suppose first that $d = \gcd(a, b)$. Then (I) holds by the definition of gcd. By 1.2.7 $d = ax + by$ for some $x, y \in \mathbb{Z}$. So if c is a common divisor of a and b, then 1.2.5 shows that $c \mid d$. Thus (II) holds.

$\Longleftarrow$: Suppose next that (I) and (II) holds. Then d is a common divisor of a and b by (I). Let c be a common divisor of a and b. Then by (II), $c \mid d$. Thus by 1.2.1, $c \leq |d| = d$. Hence by definition, d is a greatest common divisor of a and b. $\square$

Theorem 1.2.11. *Let a, b integers not both 0 with $\gcd(a, b) = 1$. Let c be an integer with $a \mid bc$. Then $a \mid c$.*

Proof. Since $\gcd(a, b) = 1$, 1.2.7 shows that $1 = ax + by$ for some $x, y \in \mathbb{Z}$. Hence

$$c = 1c = (ax + by)c = a(xc) + (bc)y.$$

Note that a divides a and bc, and that xc and y are integers. So by 1.2.5, a also divides $a(xc) + (cb)y$. Thus $a \mid c$. $\square$

Exercises 1.2:

#1. If $a \mid b$ and $b \mid c$, prove that $a \mid c$.

#2. If $a \mid c$ and $b \mid c$, must ab divide c? What if $\gcd(a, b) = 1$?

#3. Let a and b be integers, not both zero. Show that $\gcd(a, b) = 1$ if and only if there exist integers u and v with $ua + vb = 1$.

#4. Let a and b be integers, not both zero. Let $d = \gcd(a, b)$ and let e be a positive common divisor of a and b.

(a) Show that $\gcd(\frac{a}{e}, \frac{b}{e}) = \frac{d}{e}$.

(b) Show that $\gcd(\frac{a}{d}, \frac{b}{d}) = 1$.

#5. Prove or disprove each of the following statements.

(a) If $2 \nmid a$, then $4 \mid (a^2 - 1)$.

(b) If $2 \nmid a$, then $8 \mid (a^2 - 1)$.

#6. Let n be a positive integers and a and b integers with $\gcd(a, b) = 1$. Use induction to show that $\gcd(a, b^n) = 1$.

#7. Let a, b, c be integers with a, b not both zero. Prove that the equation $ax + by = c$ has integer solutions if and only if $\gcd(a, b) \mid c$.

#8. Prove that $\gcd(n, n + 1) = 1$ for any integer n.

#9. Prove or disprove each of the following statements.

(a) If $2 \nmid a$, then $24 \mid (a^2 - 1)$.

(b) If $2 \nmid a$ and $3 \nmid a$, then $24 \mid (a^2 - 1)$.

#10. Let n be an integer. Then $\gcd(n + 1, n^2 - n + 1) = 1$ or 3.

#11. Let a, b, c be integers with $a \mid bc$. Show that there exist integers $\tilde{b}, \tilde{c}$ with $\tilde{b} \mid b, \tilde{c} \mid c$ and $a = \tilde{b}\tilde{c}$.

1.3 Integral Primes

Definition 1.3.1. *An integer p is called a prime if $p \notin \{0, 1, -1\}$ and the only divisors of p are 1, -1, p and $-p$.*

Lemma 1.3.2. (a) *Let p be an integer. Then p is a prime if and only if $-p$ is prime.*

(b) *Let p be a prime and a an integer. Then either ($p \mid a$ and $\gcd(a, p) = |p|$) or ($p \nmid a$ and $\gcd(a, p) = 1$).*

(c) *Let p and q be primes with $p \mid q$. Then $p = q$ or $p = -q$.*

Proof. (a) Note that

$$(*) \qquad p \notin \{0, \pm 1\} \quad \text{if and only if} \quad -p \notin \{0, \pm 1\},$$

By 1.2.1

$(**)$ *p and $-p$ have the same divisor.*

Moreover,

$$(***) \qquad \pm p = \pm(-p)$$

Thus the following statements are equivalent:

	p is a prime		
$\Longleftrightarrow$	$p \notin \{0, \pm 1\}$ and the only divisors of p are ± 1 and $\pm p$	-	Definition of a prime.
$\Longleftrightarrow$	$-p \notin \{0, \pm 1\}$ and the only divisors of $-p$ are ± 1 and $\pm(-p)$	-	$(*)$,$(**)$ and $(***)$
$\Longleftrightarrow$	$-p$ is a prime.	-	Definition of a prime.

So (a) holds.

(b): Put $d := \gcd(a, p)$. Then $d \mid p$ and since d is prime, $d \in \{\pm 1, \pm p\}$. Since d is positive we conclude

$$(+) \qquad d = 1 \qquad \text{or} \qquad d = |p|.$$

Case 1: Suppose $p \mid a$.

Since $p \mid p$, p is a common divisor of a and p. Thus (by 1.2.1(c), also $|p|$ is a common divisor of a and p. Since $d = \gcd(a, p)$ this gives and so $d \geq |p|$. As $p \notin \{0, \pm 1\}$ we have $|p| > 1$. Hence also $d > 1$ and so $d \neq 1$. Thus by (+) $d = |p|$. So $p \mid a$ and $\gcd(a, p) = |p|$. Thus (b) holds in this case.

Case 2: Suppose $p \nmid a$.

Then also $|p| \nmid a$. As $d = \gcd(a, p)$,we have $d \mid a$ and so $d \neq |p|$. Hence by (+) $dab = 1$. Thus $p \nmid a$ and $\gcd(a, b) = 1$. So (b) also holds in this case.

(c): Suppose p and q are primes with $p \mid q$. Since q is a prime we get $p \in \{\pm 1, \pm q\}$. Since p is prime, $p \notin \{\pm 1\}$ and so $p \in \{\pm q\}$. □

Theorem 1.3.3. *Let p be an integer with $p \notin \{0, \pm 1\}$. Then the following two statements are equivalent:*

(a) *p is a prime.*

(b) *If a and b are integers with $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. Suppose p is prime and $p \mid ab$ for some integers a and b. If $p \nmid a$, then by 1.3.2, $\gcd(p, a) = 1$. Since $p \mid ab$, 1.2.11 implies $p \mid b$. So $p \mid a$ or $p \mid b$.

For the converse, see Exercise 1.3#2. □

Corollary 1.3.4. *Let p be a prime integer, n a positive integer and $a_1, a_2, \ldots a_n$ integers with $p \mid a_1 a_2 \ldots a_n$. Then $p \mid a_i$ for some $i \in \mathbb{Z}$ with $1 \leq i \leq n$.*

Proof. The proof is by induction on n. If $n = 1$, then $p \mid a_1$ and so the Corollary holds with $i = 1$. Suppose now that the Corollary holds for $n = k$ and let $a_1, a_2 \ldots a_{k+1}$ be integers with $p \mid a_1 a_2 \ldots a_k a_{k+1}$. Put $a = a_1 \ldots a_k$ and $b = a_{k+1}$. Then $p \mid ab$ and so by 1.3.3, $p \mid a$ or $p \mid b$. If $p \mid a$, then $p \mid a_1 \ldots a_k$ and so by the induction assumption, $p \mid a_i$ for some $i \in \mathbb{Z}$ with $1 \leq i \leq k$. If $p \mid b$, then $p \mid a_{k+1}$. In either case $p \mid a_i$ for some $i \in \mathbb{Z}$ with $1 \leq i \leq k + 1$. Thus the Corollary holds for $n = k + 1$.

The Principal of Induction now shows that the Corollary holds for all positive integers n. □

Lemma 1.3.5. *Let n be an integer with $n > 1$. Then the following statements are equivalent:*

(a) *n is not a prime.*

(b) *There exists $a \in \mathbb{Z}$ with $a \mid n$ and $1 < a < n$.*

(c) *There exist $a, b \in \mathbb{Z}$ with $n = ab$, $1 < a < n$ and $1 < b < n$.*

(d) *There exist* $a, b \in \mathbb{Z}$ *with* $n = ab$*,* $a > 1$ *and* $b > 1$*.*

(e) *There exist* $a, b \in \mathbb{Z}$ *with* $n = ab$*,* $a < n$ *and* $b < n$*.*

Proof. We will first prove

$(*)$ *Let* a *and* b *be positive integers with* $n = ab$*, then*

$$\left(1 < a \iff b < n\right) \quad \text{and} \quad \left(1 < b \iff a < n\right)$$

Since a is positive, we have $1 < a$ if and only if $\frac{1}{a} < 1$, if and only if $\frac{n}{a} < n$ and if and only if $b < n$. By symmetry, $1 < b$ if and only of $a < n$.

(a) $\Longrightarrow$ (b): Suppose that n is not a prime. Since $n > 1$, $n \notin \{0, \pm 1\}$ and the definition of a prime shows that there exists a divisor m of n with $m \notin \{\pm 1, \pm n\}$. Put $a = |m|$. Then also a is a divisor of n, a is positive and $a \neq 1$ and $a \neq n$. Since a divides n, 1.2.1 implies $1 \leq |a| \leq |n|$. As a and n are positive this gives $1 \leq a \leq n$. Together with $a \neq 1$ and $a \neq n$ we get $1 < a < n$.

(b) $\Longrightarrow$ (c): Suppose $a \in \mathbb{Z}$ with $a \mid n$ and $1 < a < n$. Then by definition of divide, $n = ab$ for some $b \in \mathbb{Z}$. Since n and a are positive also b is positive. By $(*)$, since $1 < a$ we have $b < n$ and since $a < n$ we have $1 < b$. So (c) holds.

(c) $\Longrightarrow$ (d): If (c) holds, then (d) holds for the same a and b.

(d) $\Longrightarrow$ (e): Suppose there exist $a, b \in \mathbb{Z}$ with $n = ab$, $a > 1$ and $b > 1$. Then $(*)$ gives $a < n$ and $b < n$. So (e) holds.

(e) $\Longrightarrow$ (a): Suppose now that $n = ab$ with $a, b \in \mathbb{Z}$ and $a < n$ and $b < n$. Then a is a divisor of n and $a \neq n$. Since $b < n$, $(*)$ gives $a > 1$ and so $a \neq 1$, Since a and n are positive also $a \neq -1$ and $a \neq -n$. So a is a divisor of n other than ± 1, $\pm n$ and the definition of a prime shows that n is not a prime. □

Theorem 1.3.6. *Let* n *be integer with* $n > 1$*. Then there exist a positive integer* k *and positive primes* $p_1, p_2, \ldots, p_k$ *with*

$$n = p_1 p_2 \ldots p_k.$$

Proof. The proof is by complete induction on n. So let m be an integer with $m > 1$ and suppose that the theorem is true for all integers n with $1 < n < m$.

Case 1. *Suppose* m *is a prime.*

Put $k = 1$ and $p_1 = m$. Then $m = p_1$ and theorem holds for $n = m$ in this case.

Case 2. *Suppose* m *is not a prime prime.*

Then by 1.3.5 there exist integers a and b with $n = ab$, $1 < a < n$ and $1 < b < n$. By the induction assumption there exist positive integer i and j and primes $p_1, \ldots, p_i$, $q_1 \ldots . q_j$ with

$$a = p_1 \ldots p_i \qquad \text{and} \qquad b = q_1 \ldots q_j.$$

Thus

$$m = ab = p_1 \ldots p_i q_1 \ldots q_j.$$

Put $k := i + j$ and for $1 \leq l \leq j$ define $p_{i+l} := q_l$. Then

$$m = p_1 \ldots p_i p_{i+1} \ldots p_{i+j} = p_1 \ldots p_k$$

So again the theorem for $n = m$.

By the Principal of Complete Induction, the theorem now holds for all integers n with $n \geq 2$. □

Theorem 1.3.7 (Fundamental Theorem of Arithmetic,FTA)**.** *Let n be an integer with $n > 1$. Then n is a product of positive primes. Moreover, if*

$$n = p_1 p_2 \ldots p_k \qquad \textit{and} \qquad n = q_1 q_2 \ldots q_l,$$

where k, l are positive integers and $p_1, \ldots p_k, q_1, \ldots q_l$ are positive primes. Then $k = l$ and (possibly after reordering the $q_i's$)

$$p_1 = q_1, \quad p_2 = q_2, \quad \ldots, \quad p_k = q_k.$$

In more precise terms: There exists a bijection $\pi : \{1, 2 \ldots, k\} \to \{1, 2, \ldots, l\}$ with $p_i = q_{\pi(i)}$ for all $1 \leq i \leq k$.

Proof. By 1.3.6 n is a product of positive primes. The proof of the second statement is by complete induction on n. So let m be an integer with $m > 1$ and suppose that the FTA holds for all integers n with $1 < n < m$. Suppose also that

$$(*) \qquad m = p_1 p_2 \ldots p_k \qquad \text{and} \qquad m = q_1 q_2 \ldots q_l.$$

where k, l are positive integers and $p_1, \ldots p_k, q_1, \ldots q_l$ are positive primes.

Since p_i and q_j are primes, $p_i \neq 1$ and $q_j \neq 1$. Since p_i and q_j are positive we conclude

$$(**) \qquad p_i > 1 \text{ for all } 1 \leq i \leq k \quad \text{and} \quad q_j > 1 \text{ for all } 1 \leq j \leq l.$$

Case 1. *Suppose that m is a prime.*

Assume for a contradiction, that $k > 1$. Then by $(*)$ $m = p_1(p_2 \ldots p_k)$ and by $(**)$, $p_1 > 1$ and $p_2 \ldots p_k > 1$. Thus 1.3.5 shows that m is not a prime, contrary to the assumption. Thus $k = 1$ and by symmetry also $l = 1$. Also $p_1 = m = q_1$ and the FTA holds for $n = m$.

Case 2. *Suppose that m is not a prime.*

Then $p_1 \neq m \neq q_1$ and so $k \geq 2$ and $l \geq 2$.

Since $m = (p_1 \dots p_{k-1})p_k$ we see that p_k divides m. As $m = q_1 \dots q_l$ we conclude that p_k divides $q_1 \dots q_l$ and thus by 1.3.4, $p_k \mid q_j$ for some $1 \leq j \leq l$. Since p_k and q_j are primes, 1.3.2, gives $p_k = q_j$ or $p_k = -q_j$. Since p_k and q_j are positive, $p_k = q_j$. Reordering the q_j's we may assume that $j = l$. So

$$(***) \qquad p_k = q_l$$

Put $u := \frac{m}{p_k} = \frac{m}{q_l}$. Dividing the first equation in $(*)$ by p_k and the second by q_l gives

$$(+) \qquad u = p_1 p_2 \dots p_{k-1} \qquad \text{and} \qquad u = q_1 q_2 \dots q_{l-1}.$$

By $(**)$ $p_k > 1$ and so $u = \frac{m}{p_k} < m$. Also $p_1 > 1$ so $u = p_1 \dots p_{k-1} > 1$. Hence $1 < u < m$ and so by the induction assumption the FTA holds for $n = u$. Thus $(+)$ implies $k - 1 = l - 1$ and, possibly after reordering $q_1, \dots, q_{k-1}$,

$$p_1 = q_1, \quad p_2 = q_k, \quad \dots, \quad p_{k-1} = q_{k-1}.$$

From $k - 1 = l - 1$ we get $k = l$ and so by $(***)$ $p_k = q_l = q_k$. So the FTA holds for $n = m$.

The Principal of Complete Induction now shows that the FTA holds for any integer n with $n > 1$. □

Exercises 1.3:

#1. Let p be an integer other than $0, \pm 1$. Prove that p is a prime if and only if it has this property: Whenever r and s are integers such that $p = rs$, then $r = \pm 1$ or $s = \pm 1$.

#2. Let p be an integer other than $0, \pm 1$ with this property

(*) Whenever b and c are integers with $p \mid bc$, then $p \mid b$ or $p \mid c$. Prove that p is a prime.

#3. (a) List all the positive divisors of $3^s 5^t$ where $s, t \in \mathbb{Z}$ and $s, t > 0$.

(b) If $r, s, t \in \mathbb{Z}$ are positive, how many positive divisors does $2^r 3^s 5^t$ have?

#4. Prove that $\gcd(a, b) = 1$ if and only if there is no prime p such that $p \mid a$ and $p \mid b$.

#5. Prove or disprove each of the following statements:

(a) If p is a prime and $p \mid a^2 + b^2$ and $p \mid c^2 + d^2$, then $p \mid (a^2 - c^2)$

(b) If p is a prime and $p \mid a^2 + b^2$ and $p \mid c^2 + d^2$, then $p \mid (a^2 + c^2)$

(c) If p is a prime and $p \mid a$ and $p \mid a^2 + b^2$, then $p \mid b$

#6. Let a and b be integers. Then $a \mid b$ if and only if $a^3 = b^3$.

#7. Prove or disprove: Let n be a positive integer, then there exists $p, a \in \mathbb{Z}$ such that $n = p + a^2$ and either $p = 1$ or p is a prime.

Chapter 2

Congruence in $\mathbb{Z}$ and Modular Arithmetic

2.1 Congruence and Congruence Classes

Let a,b and n be integers. Recall that the relation '$\equiv \pmod{n}$' on $\mathbb{Z}$ is defined by

$$a \equiv b \pmod{n} \iff n|a-b$$

By 0.5.5 '$\equiv \pmod{n}$' is an equivalence relation on Z. Recall also that $[a]_n$ is the equivalence class of '$\equiv \pmod{n}$' with respect to a. So

$$[a]_n = \{b \in \mathbb{Z} \mid a \equiv b \pmod{n}\}.$$

Theorem 2.1.1. *Let a, b, n be integers with $n \neq 0$. Then the following statements are equivalent*

(a) $a = b + nk$ *for some integer* k.

(b) $a - b = nk$ *for some integer* k.

(c) $n \mid a - b$.

(d) $a \equiv b \pmod{n}$.

(e) $b \in [a]_n$.

(f) $[a]_n \cap [b]_n \neq \emptyset$.

(g) $[a]_n = [b]_n$.

(h) $a \in [b]_n$.

(i) $b \equiv a \pmod{n}$.

(j) $n \mid b - a$.

(k) $b - a = nl$ *for some integer* l.

(l) $b = a + nl$ *for some integer* l.

(m) *a and b have the same remainder when divided by n.*

Proof. (a) $\iff$ (b): Add b to both sides of (b).

(b) $\iff$ (c): Follows from the definition of 'divide'.

(c) $\iff$ (d): Follows from the definition of '$\equiv \pmod{n}$'.

By 0.5.5 '$\equiv \pmod n$' is an equivalence relation. So Theorem 0.5.8 implies that (d)-(i) are equivalent. Since we already proved that (a)-(d) are equivalent we conclude that (a) to (i) are equivalent.

Note that (g) is symmetric in a and b. Since (a)-(c) are equivalent to (g), we can interchange a and b in (a)-(c) and conclude that (j) to (l) are equivalent to (g). Thus (a)-(l) are equivalent.

By the division algorithm there exists integers q_1, r_1, q_2, r_2 with

$$a = nq_1 + r_1 \qquad \text{and} \qquad 0 \leq r_1 < |n|$$

and

$$b = nq_2 + r_2 \qquad \text{and} \qquad 0 \leq r_2 < |n|.$$

So r_1 and r_2 are remainders of a and b, respectively when divided by n.

(m) $\Longrightarrow$ (a): Suppose (m) holds. Then $r_1 = r_2$ and

$$a - b = (nq_1 + r_1) - (nq_2 + r_2) = n(q_1 - q_2) + (r_1 - r_2) = n(q_1 - q_2).$$

Hence $a = b + n(q_1 - q_2)$. Since $q_1 - q_2 \in \mathbb{Z}$ we see that (a) holds with $k = q_1 - q_2$.

(a) $\Longrightarrow$ (m): Suppose (a) holds. Then $a = b + nk$ for some integer k. Then

$$a = (nq_2 + r_2) + nk = n(q_2 + k) + r_2.$$

Since $q_2 + k \in \mathbb{Z}$ and $0 \leq r_2 < |n|$, we conclude that r_2 is the remainder of a when divided by n. So $r_1 = r_2$ and (m) holds. □

Corollary 2.1.2. *Let n be positive integer.*

(a) *Let $a \in \mathbb{Z}$. Then there exists a unique $r \in \mathbb{Z}$ with $0 \leq r < n$ and $[a]_n = [r]_n$, namely r is the remainder of a when divided by n.*

(b) *There are exactly n distinct congruence classes modulo n, namely*

$$[0], [1], [2], \ldots, [n-1].$$

(c) *$|\mathbb{Z}_n| = n$, that is $\mathbb{Z}_n$ has exactly n elements.*

Proof. (a) Let $a \in \mathbb{Z}$, let s be the remainder of a when divided by n and let $r \in \mathbb{Z}$ with $0 \leq r < n$. Since $r = 0n + r$ and $0 \leq r < n$, r is the remainder of r when divided by n. By 2.1.1, $[a]_n = [r]_n$ if and only a and r have the same remainder when divided by n, and so if and only if $r = s$.

(b) By definition each congruence class modulo n is of the form $[a]_n$, with $a \in \mathbb{Z}$. By (a), $[a]_n$ is equal to exactly one of

$$[0], [1], [2], \ldots, [n-1].$$

So (b) holds.

(c) Since $\mathbb{Z}_n$ is the set of congruence classes modulo n, (c) follows from (b). □

Example 2.1.3. Determine $\mathbb{Z}_5$.

$$\mathbb{Z}_5 = \Big\{[0]_5, [1]_5, [2]_5, [3]_5, [4]_5\Big\} = \Big\{[0]_5, [1]_5, [2]_5, [-2]_5, [-1]_5\Big\}$$

Exercises 2.1:

#1. (a) Let k be an integer with $k \equiv 1 \pmod 4$. Compute the remainder of $6k+5$ when divided by 4.

(b) Let r and s be integer with $r \equiv 3 \pmod{10}$ and $s \equiv -7 \pmod{10}$. Compute the remainder of $2r + 3s$ when divided by 10.

#2. If $a, m, n \in \mathbb{Z}$ with $m, n > 0$, prove that $[a^m]_2 = [a^n]_2$

#3. If $p \geq 5$ and p is a prime, prove that $[p] = [1]$ or $[p] = [5]$ in $\mathbb{Z}_6$.

#4. Find all solutions of each congruence:

(a) $2x \equiv 3 \pmod 5$ (b) $3x \equiv 1 \pmod 7$

(c) $6x \equiv 9 \pmod{15}$ (d) $6x \equiv 10 \pmod{15}$

#5. If $a \equiv 2 \pmod 4$, prove that there are no integers c and d with $a = c^2 - d^2$.

#6. If $[a] = [1]$ in $\mathbb{Z}_n$, prove that $\gcd(a, n) = 1$. Show by example that the converse is not true.

#7. (a) Show that $10^n \equiv 1 \pmod 9$ for every positive integer n.

(b) Prove that every positive integer is congruent to the sum of its digits mod 9. [for example, $38 \equiv 11 \pmod 9$].

2.2 Modular Arithmetic

Theorem 2.2.1. *Let $a, \tilde{a}, b, \tilde{b}$ and n be integers with $n \neq 0$. Suppose that*

$$[a]_n = [\tilde{a}]_n \qquad \textit{and} \qquad [b]_n = [\tilde{b}]_n.$$

or that

$$a \equiv \tilde{a} \pmod n \qquad \textit{and} \qquad b \equiv \tilde{b} \pmod n$$

Then

$$[a+b]_n = [\tilde{a}+\tilde{b}]_n \qquad \textit{and} \qquad [ab]_n = [\tilde{a}\tilde{b}]_n.$$

and

$$a + b \equiv \tilde{a} + \tilde{b} \pmod n \qquad \textit{and} \qquad ab \equiv \tilde{a}\tilde{b} \pmod n$$

Proof. Since

$$[a]_n = [\tilde{a}]_n \qquad \text{and} \qquad [b]_n = [\tilde{b}]_n.$$

or

$$a \equiv \tilde{a} \pmod{n} \qquad \text{and} \qquad b \equiv \tilde{b} \pmod{n}$$

we conclude from 2.1.1 that

$$\tilde{a} = a + nk \qquad \text{and} \qquad \tilde{b} = b + nl$$

for some $k, l \in \mathbb{Z}$. Hence

$$\tilde{a} + \tilde{b} = (a + nk) + (b + nl) = (a + b) + n(k + l).$$

Since $k + l \in \mathbb{Z}$, 2.1.1 gives

$$[a + b]_n = [\tilde{a} + \tilde{b}]_n \qquad \text{and} \qquad a + b \equiv \tilde{a} + \tilde{b} \pmod{n}$$

Also

$$\tilde{a} \cdot \tilde{b} = (a + nk)(b + nl) = ab + n(al + kb + knl),$$

and, since $al + kb + knl \in \mathbb{Z}$, 2.1.1 implies

$$[ab]_n = [\tilde{a}\tilde{b}]_n \qquad \text{and} \qquad ab \equiv \tilde{a}\tilde{b} \pmod{n}.$$

□

In view of 2.2.1 the following definition is well-defined.

Definition 2.2.2. *Let a, b and n be integers with $n \neq 0$. Then*

$$[a]_n \oplus [b]_n = [a + b]_n \quad \text{and} \quad [a]_n \odot [b]_n = [ab]_n.$$

The function

$$\mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n, (A, B) \to A \oplus B$$

is called the addition *on $\mathbb{Z}_n$, and the function*

$$\mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n, (A, B) \to A \odot B$$

is called the multiplication *on $\mathbb{Z}_n$.*

Example 2.2.3. (1) Compute $[3]_8 \odot [7]_8$.

$$[3]_8 \odot [7]_8 = [3 \cdot 7]_8 = [21]_8 = [8 \cdot 2 + 5]_8 = [5]_8$$

Note that $[3]_8 = [11]_8$ and $[7]_8 = [-1]_8$. So we could also have used the following computation:

$$[11]_8 \odot [-1]_8 = [11 \cdot -1]_8 = [-11]_8 = [-11 + 8 \cdot 2]_8 = [5]_8$$

Theorem 2.2.1 ensures that we will always get the same answer, not matter what representative we pick for the congruence class.

(2) Compute $[123]_{212} \oplus [157]_{212}$.

$$[123]_{212} \oplus [157]_{212} = [123 + 157]_{212} = [280]_{212} = [280 - 212]_{212} = [68]_{212}$$

Note that $[123]_{212} = [123 - 212]_{212} = [-89]_{212}$ and $[157]_{212} = [157 - 212]_{212} = [-55]_{212}$. Also

$$[-89]_{212} \oplus [-55]_{212} = [-89 - 55]_{212} = [-144]_{212} = [-144 + 212]_{212} = [68]_{212}$$

(3) **Warning:** Congruence classes can not be used as exponents:

We have

$$[2^4]_3 = [16]_3 = [1]_3 \quad \text{and} \quad [2^1]_3 = [2]_3$$

So

$$[2^4]_3 \neq [2^1]_3 \quad \text{even though} \quad [4]_3 = [1]_3$$

Theorem 2.2.4. *Let n be a non-zero integer and $A, B, C \in \mathbb{Z}_n$. Then*

(1) $A \oplus B \in \mathbb{Z}_n$ *[closure for addition].*

(2) $A \oplus (B \oplus C) = (A \oplus B) \oplus C$. *[associative addition]*

(3) $A \oplus B = B \oplus A$. *[commutative addition]*

(4) $A \oplus [0]_n = A = [0]_n \oplus A$. *[additive identity]*

(5) *There exists $X \in \mathbb{Z}_n$ with $A \oplus X = [0]_n$.* *[additive inverse]*

(6) $A \odot B \in \mathbb{Z}_n$. *[closure for multiplication]*

(7) $A \odot (B \odot C) = (A \odot B) \odot C$. *[associative multiplication]*

(8) $A \odot (B \oplus C) = (A \odot B) \oplus (A \odot C)$ *and* $(A \oplus B) \odot C = (A \odot C) \oplus (B \odot C)$. *[distributive laws]*

(9) $A \odot B = B \odot A$. *[commutative multiplication]*

(10) $[1]_n \odot A = A = A \odot [1]_n$ *[multiplicative identity]*

Proof. If $d \in \mathbb{Z}$ we will just write $[d]$ for $[d]_n$. By definition of $\mathbb{Z}_n$ there exists integers a, b and c with $A = [a]$, $B = [b]$ and $C = [c]$.

(1) We have $A \oplus B = [a] \oplus [b] = [a + b]$. Since $a + b \in \mathbb{Z}$ we conclude that $A \oplus B \in \mathbb{Z}_n$.

(2) Using the definition of $\oplus$ and the fact that addition in $\mathbb{Z}$ is associative we compute

$$\begin{aligned} A \oplus (B \oplus C) &= [a] \oplus ([b] \oplus [c]) = [a] \oplus [b + c] = [a + (b + c)] = [(a + b) + c] \\ &= [a + b] \oplus [c] = ([a] \oplus [b]) \oplus [c] = (A \oplus B) \oplus C. \end{aligned}$$

(3) Using the definition of $\oplus$ and the fact that addition in $\mathbb{Z}$ is commutative we compute

$$A \oplus B = [a] \oplus [b] = [a + b] = [b + a] = [b] \oplus [a] = B \oplus A.$$

(4) Using the definition of $\oplus$ and the fact that 0 is an additive identity in $\mathbb{Z}$ we compute

$$A \oplus [0] = [a] \oplus [0] = [a + 0] = [a] = A,$$

and

$$[0] \oplus A = [0] \oplus [a] = [0 + a] = [a] = A.$$

(5) Put $X = [-a]$. Then $X \in \mathbb{Z}_n$. Using the definition of $\oplus$ and the fact that $-a$ is an additive inverse for a in $\mathbb{Z}$ we compute

$$A \oplus X = [a] \oplus [-a] = [a + (-a)] = [0].$$

(6) Similarly to (1) we have $A \odot B = [a] \odot [b] = [ab]$ and so $A \odot B \in \mathbb{Z}_n$.

(7) Similarly to (2) we can use the definition of $\odot$ and the fact that addition in $\mathbb{Z}$ is associative to compute

$$\begin{aligned} A \odot (B \odot C) &= [a] \odot ([b] \odot [c]) = [a] \odot [bc] = [a(bc)] = [(ab)c] \\ &= [ab] \odot [c] = ([a] \odot [b]) \odot [c] = (A \odot B) \odot C. \end{aligned}$$

(8) Using the definition of $\oplus$ and $\odot$ and the distributive law in $\mathbb{Z}$ we compute

$$\begin{aligned} A \odot (B \oplus C) &= [a] \odot ([b] \oplus [c]) = [a] \odot [b + c] = [a(b + c)] \\ &= [ab + bc] = [ab] \oplus [ac] = ([a] \odot [b]) \oplus ([a] \odot [c]) \\ &= (A \odot B) \oplus (A \odot C), \end{aligned}$$

and similarly

$$
\begin{aligned}
(A \oplus B) \odot C &= ([a] \oplus [b]) \odot [c] = [a+b] \odot [c] = [(a+b)c] \\
&= [ac+bc] = [ac] \oplus [bc] = ([a] \odot [c]) \oplus ([b] \odot [c]) \\
&= (A \odot C) \oplus (B \odot C).
\end{aligned}
$$

(9) Similarly to (3) we can use the definition of $\odot$ and the fact that multiplication in $\mathbb{Z}$ is commutative to compute

$$A \odot B = [a] \odot [b] = [ab] = [ba] = [b] \odot [a] = B \odot A.$$

(10) Similarly to (4) we can use the definition of $\odot$ and the fact that 1 is a multiplicative identity in $\mathbb{Z}$ to compute

$$A \odot [1] = [a] \odot [1] = [a1] = [a] = A,$$

and

$$[1] \odot A = [1] \odot [a] = [1a] = [a] = A$$

$\square$

Notation 2.2.5. *Let a, b, n be integers with $n \neq 0$. We will often just write a for $[a]_n$, $a+b$ for $[a]_n \oplus [b]_n$ and ab (or $a \cdot b$) for $[a]_n \odot [b]_n$. This notation is only to be used if it clear from the context that the symbols represent congruence classes modulo n. Exponents are always integers and never congruences class.*

Example 2.2.6. (1) Compute $4+5$ and $4 \cdot 5$ in $\mathbb{Z}_7$.

$$4+5 = 9 = 2 \qquad \text{and} \qquad 4 \cdot 5 = 20 = 6$$

(2) Determine the addition and multiplication table of $\mathbb{Z}_5$.

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	5
2	2	3	4	5	6
3	3	4	5	6	7
4	4	5	6	7	8

and

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	6	8
3	0	3	6	9	12
4	0	4	8	12	16

and after computing remainders when divided by 5:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

and

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Definition 2.2.7. *Let n be a non-zero integer, $A \in \mathbb{Z}_n$ and $k \in \mathbb{N}$. Then A^k is inductively defined by*

$$A^0 = [1]_n \quad \text{and} \quad A^{k+1} = A^k \odot A.$$

So

$$A^k = \underbrace{\Big(\big((A \odot A) \odot A\big) \dots \odot A\Big) \odot A}_{k-\text{times}}$$

Lemma 2.2.8. *Let n be a non-zero integer and $k, l \in \mathbb{N}$.*

(a) *Let $a \in \mathbb{Z}$. Then $[a]_n^k = [a^k]_n$.*

(b) *Let $A, B \in \mathbb{Z}_n$. Then $(A \odot B)^k = A^k \odot B^k$, $A^{k+l} = A^k \odot A^l$ and $A^{kl} = (A^k)^l$.*

Proof. (a) The proof is by induction on k. For $k = 0$, $[a]^0 = [1] = [a^0]$ and so (a) holds for $k = 0$. Suppose (a) holds for k, then

$$[a]^{k+1} = [a]^k \odot [a] = [a^k] \odot [a] = [a^k a] = [a^{k+1}],$$

$$\begin{array}{rcll} [a]^{k+1} & = & [a]^k \odot [a] & - \text{Definition of } [a]^{k+1}, 2.2.7 \\ & = & [a^k] \odot [a] & - \text{Induction assumption} \\ & = & [a^k a] & - \text{Definition of } \odot \\ & = & [a^{k+1}] & - \text{Definition of } a^{k+1}, \end{array}$$

and so (a) holds for $k + 1$. So by the Principal of Induction, (a) holds for all $k \in \mathbb{N}$.

(b) Choose $a, b \in \mathbb{Z}$ with $A = [a]$ and $B = [b]$. Using (a) and the fact that (b) holds for integers in place of congruence classes we compute:

$$(A \odot B)^k = ([a] \odot [b])^k = [ab]^k = [(ab)^k] = [a^k b^k] = [a^k] \odot [b^k] = [a]^k \odot [b^k] = A^k \odot B^k,$$

$$A^{k+l} = [a]^{k+l} = [a^{k+l}] = [a^k a^l] = [a^k] \odot [a^l] = [a]^k \odot [a]^l = A^k \odot A^l,$$

and

$$A^{kl} = [a]^{kl} = [a^{kl}] = [(a^k)^l] = [a^k]^l = ([a]^k)^l = (A^k)^l$$

$\square$

Remark 2.2.9. *Consider the expression*

$$2^5 + 3 \cdot 7 \qquad \textit{in} \quad \mathbb{Z}_n$$

It is not clear which element of $\mathbb{Z}_n$ this represents, indeed it could be any of the following for elements:

$$[2^5 + 3 \cdot 7]_n$$
$$[2^5]_n \oplus [3 \cdot 7]_n$$
$$[2^5]_n \oplus ([3]_n \odot [7]_n)$$
$$[2]_n^5 \oplus [3 \cdot 7]_n$$
$$[2]_n^5 \oplus ([3]_n \odot [7]_n)$$

But thanks to Theorem 2.2.1 and Theorem 2.2.8 all these elements are actually equal. So our simplified notation is not ambiguous. In other words, our use of the simplified notation is only justified by Theorem 2.2.1 and Theorem 2.2.8.

Example 2.2.10. (1) Compute $[13^{34567}]_{12}$.

$$[13^{34567}]_{12} = [13]_{12}^{34567} = [1]_{12}^{34567} = [1^{34567}]_{12} = [1]_{12}$$

In simplified notation this becomes: In $\mathbb{Z}_{12}$, $13 = 1$ and so

$$13^{34567} = 1^{34567} = 1$$

Why is the calculation shorter? In simplified notation the expression

$$[13^{34567}]_{12} \quad \text{and} \quad [13]_{12}^{34567}$$

are both written as

$$13^{34567}$$

So the step

$$[13^{34567}]_{12} = [13]_{12}^{34567}$$

is invisibly performed by the simplified notation. Similarly, the step

$$[1]_{12}^{34567} = [1^{34567}]_{12}$$

disappears through our use of the simplified notation.

(2) Compute $[7]_{50}^{198}$.

In $\mathbb{Z}_{50}$:

$$7^{198} = (7^2)^{99} = 49^{99} = (-1)^{99} = -1 = 49.$$

(3) Determine the remainder of $53 \cdot 7^{100} + 47 \cdot 7^{71} + 4 \cdot 7^3$ when divided by 50.

In $\mathbb{Z}_{50}$:

$$\begin{aligned} 53 \cdot 7^{100} + 47 \cdot 7^{71} + 4 \cdot 7^3 &= 3 \cdot (7^2)^{50} - 3 \cdot (7^2)^{35} \cdot 7 + 4 \cdot 7^2 \cdot 7 \\ &= 3 \cdot (-1)^{50} - 3 \cdot (-1)^{35} \cdot 7 + 4 \cdot -1 \cdot 7 \\ &= 3 + 21 - 28 = 3 - 7 = -4 = 46 \end{aligned}$$

Thus $[53 \cdot 7^{100} + 47 \cdot 7^{73} + 4 \cdot 7^3]_{50} = [46]_{50}$. Since $0 \leq 46 < 50$, 2.1.2(a) shows that the remainder in question is 46.

Example 2.2.11. Find all solutions of $x^3 + 2x + 3 = 0$ in $\mathbb{Z}_5$.

All computation below are in $\mathbb{Z}_5$.

By Corollary 2.1.2 $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Since $3 = -2$ and $4 = -1$, $\mathbb{Z}_5 = \{0, 1, 2, -2, -1\}$. We compute

$$\begin{array}{c|cccccccccc} x & x^3 & + & 2x & + & 3 & & & & \\ \hline 0 & 0 & + & 0 & + & 3 & = & & & 3 \\ 1 & 1 & + & 2 & + & 3 & = & 6 & = & 1 \\ 2 & 8 & + & 4 & + & 3 & = & 15 & = & 0 \\ -2 & -8 & - & 4 & + & 3 & = & -9 & = & 1 \\ -1 & -1 & - & 2 & + & 3 & & & = & 0 \end{array}$$

So the solution of $x^3 + 2x + 3 = 0$ in $\mathbb{Z}_5$ are $x = 2$ and $x = -1 = 4$.

Exercises 2.2:

#1. Let n be a non-zero integer and $A \in \mathbb{Z}_n$. Show that $A \odot [0]_n = [0]_n$.

#2. (a) Solve the equation $x^2 + x = 0$ in $\mathbb{Z}_5$.

(b) Solve the equation $x^2 + x = 0$ in $\mathbb{Z}_6$.

(c) If p is a prime, prove that the only solutions of $x^2 + x = 0$ in $\mathbb{Z}_p$ are $[0]$ and $[p - 1]$.

#3. Solve the equations:

(a) $x^2 = 1$ in $\mathbb{Z}_2$

(b) $x^4 = 1$ in $\mathbb{Z}_5$

(c) $x^2 + 3x + 2 = 0$ in $\mathbb{Z}_6$

(d) $x^2 + 1 = 0$ in $\mathbb{Z}_{12}$

#4. (a) Find an element a in $\mathbb{Z}_7$ such that every non-zero element of $\mathbb{Z}_7$ is a power of a.

(b) Do part (a) in $\mathbb{Z}_5$

(c) Can you do part (a) in $\mathbb{Z}_6$?

#5. (a) Solve the equation $x^2 + x = 0$ in $\mathbb{Z}_5$.

(b) Solve the equation $x^2 + x = 0$ in $\mathbb{Z}_6$.

(c) If p is a prime, prove that the only solutions of $x^2 + x = 0$ in $\mathbb{Z}_p$ are $[0]$ and $[p-1]$.

2.3 Cogruence classes modulo primes

Lemma 2.3.1. *Let $n, m \in \mathbb{Z}$ with $n \neq 0$. Then $n \mid m$ if and only if $[m]_n = [0]_n$.*

Proof. $n \mid m$ if and only if $n \mid m - 0$ and so by 2.1.1 if and only $[m]_n = [0]_n$. □

Theorem 2.3.2. *Let n be an integer with $|n| > 1$. Then the following statements are equivalent:*

(a) *n is a prime.*

(b) *For any $A \in \mathbb{Z}_n$ with $A \neq [0]_n$ there exists $X \in \mathbb{Z}_n$ with $AX = [1]_n$.*

(c) *Whenever A and B are elements in $\mathbb{Z}_n$ with $AB = [0]_n$, then $A = [0]_n$ or $B = [0]_n$.*

Proof. Let $m \in \mathbb{Z}$. We will write $[m]$ for $[m]_n$.

(a) $\Longrightarrow$ (b): Suppose n is a prime and let $A \in \mathbb{Z}_n$ with $A \neq [0]$. Then $A = [a]$ for some $a \in \mathbb{Z}$. Since $[a] \neq [0]$, 2.3.1 implies $n \nmid a$. Since n is prime, 1.3.2 shows $\gcd(a, n) = 1$ and so by the Euclidean Algorithm 1.2.7 there exist $u, v \in \mathbb{Z}$ with $au + nv = 1$. Hence 2.1.1(a)(g) implies $[au] = [1]$. By the definition of multiplication in $\mathbb{Z}_n$, $[a][u] = [au]$ and so $[a][u] = [1]$. Put $X = [u]$. Then $X \in \mathbb{Z}_n$ and $AX = [1]$.

(b) $\Longrightarrow$ (c): Suppose (b) holds and let $A, B \in \mathbb{Z}_n$ with $AB = [0]$. Assume that $A \neq [0]$. Then by (b) there exists $X \in \mathbb{Z}_n$ with $AX = [1]$. We compute

$$\begin{aligned}
[0] &= X[0] && \text{– See Exercise 2.2.\#1}\\
&= X(AB) && \text{– Since } AB=[0]\\
&= (XA)B && \text{– associative multiplication, 2.2.4(7)}\\
&= (AX)B && \text{– commutative multiplication, 2.2.4(9)}\\
&= [1]B && \text{– Since } AX=[1]\\
&= B && \text{– Since [1] is a multiplicative identity, 2.2.4(10)}
\end{aligned}$$

We have proven that $A \neq [0]$ implies $B = [0]$. So $A = [0]$ or $B = [0]$ and (c) holds.

(c) $\Longrightarrow$ (a): We will use Theorem 1.3.3, namely n is a prime if and only if $n \mid a$ or $n \mid b$ whenever a and b are integers with $n \mid ab$.

So suppose (c) holds and let a and b be integers with $n \mid ab$. Then $[ab] = [0]$ by 2.3.1 and thus $[a][b] = [ab] = [0]$. (b) implies $[a] = [0]$ or $[b] = [0]$. Hence by 2.3.1 $n \mid a$ or $n \mid b$. Thus by 1.3.3, n is a prime. □

Example 2.3.3. Use multiplication tables to verify Theorem 2.3.2 for $n = 4$ and $n = 5$.

Note first that Condition 2.3.2(b) in Theorem 2.3.2 says that every row of the multiplication table of $\mathbb{Z}_n$ other than Row 0 (that is the row corresponding to 0) contains 1.

Condition 2.3.2(b) in Theorem 2.3.2 says that 0 only appears in Row 0 and in Column 0 of the multiplication table.

The multiplication table for $\mathbb{Z}_4$ and $\mathbb{Z}_5$ are :

$\mathbb{Z}_4$:

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

$\mathbb{Z}_5$:

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Row 2 of the table for $\mathbb{Z}_4$ does not contain a 1. Also the entry in Row 2, Column 2 is 0. Moreover 4 is not a prime. So for $n = 4$ none of the three statements in Theorem 2.3.2 holds.

Each row, other than Row 0 of the table for $\mathbb{Z}_5$ contains a 1. Also 0 only appears in Row 0 and in Column 0. Moreover, 5 is a prime. So for $n = 5$ all of the three statements in Theorem 2.3.2 hold.

Corollary 2.3.4 (Multiplicative Cancellation Law)**.** *Let p be a prime and $A, B, C \in \mathbb{Z}_p$ with $A \neq [0]_p$. Then $AB = AC$ if and only if $B = C$.*

Proof. $\Longleftarrow$: If $B = C$ then $AB = AC$ by the principal of substitution.

$\Longrightarrow$: Now suppose that $AB = AC$. By 2.3.2 there exists $X \in \mathbb{Z}_p$ with $AX = [1]_p$. We compute

$$\begin{aligned}
& AB = AC \\
\Longrightarrow\quad & X(AB) = X(AC) && \text{– Principal Of Substitution} \\
\Longrightarrow\quad & (XA)B = (XA)C && \text{– associative multiplication 2.2.4(7) ,twice} \\
\Longrightarrow\quad & (AX)B = (AX)C && \text{– commutative multiplication 2.2.4(7),twice} \\
\Longrightarrow\quad & [1]_pB = [1]_pC && \text{– Since } AX = [1]_p \\
\Longrightarrow\quad & B = C && \text{– Since } [1]_p \text{ is a multiplicative identity 2.2.4(10)}
\end{aligned}$$

□

Example 2.3.5. Verify that the Cancellation Law holds in $\mathbb{Z}_5$, but does not hold in $\mathbb{Z}_4$.

Let $A, D \in \mathbb{Z}_p$ with $A \neq [0]_p$. The Cancellation law says if $B, C \in \mathbb{Z}_p$ with $D = AB$ and $D = AC$, then $B = C$. So there exists at most one $C \in \mathbb{Z}_p$ with $AC = D$. In terms of the multiplication table this means that no entry appears more than once in Row A of the multiplication table.

Note that 2 appears twice in Row 2 of the multiplication table of $\mathbb{Z}_4$, namely in Column 1 and Column 3. Indeed $2 \cdot 1 = 2 = 2 = 6 = 2 \cdot 3$ in $\mathbb{Z}_4$ but $1 \neq 3$ in $\mathbb{Z}_4$. So the Cancellation Law does not hold for $\mathbb{Z}_4$.

Except for Row 0, each row of the multiplication table of $\mathbb{Z}_5$ contains each of the congruence classes 0,1,2,3 and 4 exactly once. So the Cancellation law holds in $\mathbb{Z}_5$.

Corollary 2.3.6. *Let p be a prime and A and B in $\mathbb{Z}_p$ with $A \neq [0]_p$.*

(a) *There exists a unique $X \in \mathbb{Z}_p$ with $AX = [1]_p$.*

(b) *There exists a unique $Y \in \mathbb{Z}_p$ with $AY = B$, namely $Y = XB$.*

Proof. By 2.3.2 there exists $X \in \mathbb{Z}_p$ with $AX = [1]_p$. Thus $AX \neq [0]_p$. Since $A[0]_p = [0]_p$ by exercise 2.2.#1 we conclude $X \neq [0]_p$. Let $Y \in \mathbb{Z}_p$. Then

$$\begin{aligned}
& AY = B \\
\Longleftrightarrow\quad & X(AY) = XB && \text{– Multiplicative Cancellation Law, 2.3.4} \\
\Longleftrightarrow\quad & (XA)Y = XB && \text{– associative multiplication, 2.2.4(7)} \\
\Longleftrightarrow\quad & (AX)Y = XB && \text{– commutative multiplication, 2.2.4(9)} \\
\Longleftrightarrow\quad & [1]_pY = AB && \text{– Since } AX = [1]_p \\
\Longleftrightarrow\quad & Y = AB && \text{– Since 1 is a multiplicative identity, 2.2.4(10)}
\end{aligned}$$

So $Y = XB$ is the unique element in $\mathbb{Z}_p$ with $AX = Y$. Thus (b) holds.

The case $B = [1]_p$ shows that $X[1]_p = X$ is the unique element in $\mathbb{Z}_p$ with $AX = [1]_p$. So (a) holds. $\square$

Example 2.3.7. (a) Solve the equation $2x = 1$ in $\mathbb{Z}_5$.

(b) Solve the equation $2x = 1$ in $\mathbb{Z}_6$.

(c) Solve the equation $2x = 4$ in $\mathbb{Z}_6$.

(a): In $\mathbb{Z}_5$: $2 \cdot 3 = 1$. So 3 is a solution on $2x = 1$. By 2.3.6(a),$2x = 1$ has a unique solution and so 3 is the unique solution of $2x = 1$ in $\mathbb{Z}_5$.

(b) and (c): By 2.1.2 $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$. We compute

x	0	1	2	3	4	5
$2x$	0	2	4	6	8	10
$2x$	0	2	4	0	2	4

So $2x = 1$ has no solution in $\mathbb{Z}_6$, but $2x = 4$ has two solutions, namely $x = 2$ and $x = 5$. The second solution is explained by the fact that $2 \cdot 3 = 6 = 0$ and so

$$2 \cdot 5 = 2 \cdot (2 + 3) = 2 \cdot 2 + 2 \cdot 3 = 2 \cdot 2 + 0 = 2 \cdot 2.$$

Exercises 2.3:

#1. How many solutions does the equation $6x = 4$ have in

(a) $\mathbb{Z}_7$ (b) $\mathbb{Z}_8$ (c) $\mathbb{Z}_9$ (d) $\mathbb{Z}_{10}$

#2. Let a, b and n be integers with $n \neq 0$ and $\gcd(a, n) = 1$. Let u and v be integers with $au+nv = 1$. Put $A = [a]_n$ and $B = [b]_n$.

(a) Show that $[a]_n \odot [u]_n = [1]_n$.

(b) Show that there exists a unique X in $\mathbb{Z}_n$ with $A \odot X = B$, namely $X = [ub]_n$.

(c) Show that there exists $Y \in \mathbb{Z}_n$ with $B \odot Y = [1]_n$ if and only if $\gcd(b, n) = 1$.

#3. Let $a, b, n, m \in \mathbb{Z}$ with $n \neq 0$ and $m \neq 0$. Prove each of the following statements:

(a) $[a]_n = [b]_n$ if and only if $[ma]_{mn} = [mb]_{mn}$.

(b) $[a]_n = [b]_n$ if and only if there exists $r \in \mathbb{Z}$ with $0 \leq r < |m|$ and $[a]_{nm} = [b + rn]_{nm}$.

(c) Suppose that $[a]_n = [b]_n$, $m \mid a$ and $m \mid n$. Then $m \mid b$.

Remark 2.3.8. *Let n be a non-zero integer and A,$B \in \mathbb{Z}_n$. The preceding two exercises give rise to a method to solve the equation $A \odot X = B$ in $\mathbb{Z}_n$:*

(Step 1) *Choose* $a, b \in \mathbb{Z}$ *with* $A = [a]_n$ *and* $B = [b]_n$. *Also let* $X = [x]_n$ *with* $x \in \mathbb{Z}$. *So the equation* $A \odot X = B$ *becomes* $[ax]_n = [b]_n$.

(Step 2) *Use the Euclidean Algorithm to compute* $d = \gcd(a, n)$ *and* $u, v \in \mathbb{Z}$ *with* $au + nv = d$.

(Step 3) *If* $d \nmid b$, *then* $A \odot X = B$ *does not have a solution. Indeed, if* $X = [x]_n$ *were a solution, then* $[ax]_n = [b]_n$. *Note that* $d \mid a$ *and* $d \mid n$. *So also* $d \mid ax$ *and thus by Exercise 3(c)* $d \mid b$, *a contradiction.*

(Step 4) *Suppose now that* $d \mid b$. *Put* $\tilde{a} = \frac{a}{d}$, $\tilde{b} = \frac{b}{d}$ *and* $\tilde{n} = \frac{n}{d}$. *Then* $a = \tilde{a}d$,$ax = \tilde{a}xd$, $b = \tilde{b}d$ *and* $n = \tilde{n}d$. *Thus by Exercise 3(a)* $[\tilde{a}x]_{\tilde{n}} = [\tilde{b}]_{\tilde{n}}$ *if and only if* $[ax]_n = [b]_n$.

(Step 5) *Dividing* $ua + vb = d$ *by* d *gives* $u\tilde{a} + v\tilde{b} = 1$. *So by Exercise 2(b),* $[\tilde{a}x]_{\tilde{n}} = [\tilde{b}]_{\tilde{n}}$ *has a unique solution in* $\mathbb{Z}_{\tilde{n}}$, *namely* $[x]_{\tilde{n}} = [u\tilde{b}]_{\tilde{n}}$.

(Step 6) *By Exercise 3(b),* $[x]_{\tilde{n}} = [u\tilde{b}]_{\tilde{n}}$ *if and only if* $[x]_n = [u\tilde{b} + r\tilde{n}]_n$ *for some* $r \in \mathbb{Z}$ *with* $0 \leq r < d$. *So* X *in* $\mathbb{Z}_n$ *is a solution of* $A \odot X = B$ *if and only if* $X = [u\tilde{b} + r\tilde{n}]_n$ *for some* $r \in \mathbb{Z}$ *with* $0 \leq r < d$. *In other words, the solutions of* $A \odot X = B$ *are*

$$[u\tilde{b}]_n \quad , \quad [u\tilde{b} + \tilde{n}]_n \quad , \quad [u\tilde{b} + 2\tilde{n}]_n \quad , \quad \ldots \quad , \quad [u\tilde{b} + (d-2)\tilde{n}]_n \quad , \quad [u\tilde{b} + (d-1)\tilde{n}]_n$$

#4. Solve the following equations:

(a) $12x = 2$ in $\mathbb{Z}_{19}$.

(b) $31x = 1$ in $\mathbb{Z}_{50}$.

(c) $27x = 2$ in $\mathbb{Z}_{40}$.

(d) $7x = 2$ in $\mathbb{Z}_{24}$.

(e) $34x = 1$ in $\mathbb{Z}_{97}$.

(f) $15x = 9$ in $\mathbb{Z}_{18}$.

(g) $25x = 10$ in $\mathbb{Z}_{65}$.

(h) $21x = 17$ in $\mathbb{Z}_{33}$.

Chapter 3

Rings

3.1 Definitions and Examples

Definition 3.1.1. *A* ring *is a triple* $(R, +, \cdot)$ *such that*

(i) *R is a set;*

(ii) *$+$ is a function (called* ring addition*) and $R \times R$ is a subset of the domain of $+$. For $(a, b) \in R \times R$, $a + b$ denotes the image of (a, b) under $+$;*

(iii) *$\cdot$ is a function (*called *ring multiplication) and $R \times R$ is a subset of the domain of $\cdot$. For $(a, b) \in R \times R$, $a \cdot b$ (and also ab) denotes the image of (a, b) under $\cdot$;*

and such that the following eight statement hold:

(Ax 1) $a + b \in R$ *for all $a, b \in R$;* *[closure of addition]*

(Ax 2) $a + (b + c) = (a + b) + c$ *for all $a, b, c \in R$;* *[associative addition]*

(Ax 3) $a + b = b + a$ *for all $a, b \in R$.* *[commutative addition]*

(Ax 4) *there exists an element in R, denoted by 0_R and called 'zero R',* *[additive identity]*
such that $a = a + 0_R = a$ *and* $a = 0_R + a$ *for all $a \in R$;*

(Ax 5) *for each $a \in R$ there exists an element in R, denoted by $-a$* *[additive inverses]*
and called 'negative a', such that $a + (-a) = 0_R$;

(Ax 6) $ab \in R$ *for all $a, b \in R$;* *[closure for multiplication]*

(Ax 7) $a(bc) = (ab)c$ *for all $a, b, c \in R$;* *[associative multiplication]*

(Ax 8) $a(b + c) = ab + ac$ *and* $(a + b)c = ac + bc$ *for all $a, b, c \in R$.* *[distributive laws]*

In the following we will usually say "Let R be a ring" for " Let $(R, +, \cdot)$ be a ring."

Definition 3.1.2. *Let R be a ring. Then R is called* commutative *if*

(Ax 9) $ab = ba$ *for all* $a, b \in R$. *[commutative multiplication]*

Definition 3.1.3. *Let R be a ring. We say that R is a ring with* identity *if there exists an element, denoted by 1_R and called 'one R', such that*

(Ax 10) $a = 1_R \cdot a$ *and* $a = a \cdot 1_R$ *for all* $a \in R$. *[multiplicative identity]*

Example 3.1.4. (a) $(\mathbb{Z}, +, \cdot)$ is a commutative ring with identity.

(b) $(\mathbb{Q}, +, \cdot)$ is a commutative ring with identity.

(c) $(\mathbb{R}, +, \cdot)$ is a commutative ring with identity.

(d) $(\mathbb{C}, +, \cdot)$ is a commutative ring with identity.

(e) Let n be a non-zero integer. Then $(\mathbb{Z}_n, \oplus, \odot)$ is a commutative ring with identity.

(f) $(2\mathbb{Z}, +, \cdot)$ is a commutative ring without a multiplicative identity.

(g) Let n be integer with $n > 1$. The set $\mathrm{M}_n(\mathbb{R})$ of $n \times n$ matrices with coefficients in $\mathbb{R}$ together with the usual addition and multiplication of matrices is a non-commutative ring with identity.

Example 3.1.5. Let $R = \{0, 1\}$ and $a, b \in R$. Define an addition and multiplication on R by

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & a \end{array} \quad \text{and} \quad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & b \end{array}$$

For which values of a and b is $(R, +, \cdot)$ a ring?

Note first that 0 is additive identity, so $0_R = 0$.

Suppose that $a = 1$. Then $1 + x = 1 \neq 0_R$ for all $x \in R$ and so 1 does not have a additive inverse. Hence R is not a ring.

Suppose now that $a = 0$.
Assume that $b = 1$. Then hen $(R, +, \cdot)$ is $(\mathbb{Z}_2, \oplus, \odot)$ and so R is ring.
Assume that $b = 0$. Then then $xy = 0$ for all $x, y \in R$. Note also that $0 + 0 = 0$. It follows that Axioms 6-8 hold, indeed all expressions evaluate to 0. Axiom 1-5 hold since the addition is the same as in $\mathbb{Z}_2$. So R is a ring.
In both cases R is commutative. If $b = 1$, then 1 is an identity. If $b = 0$, R does not have an identity.

Example 3.1.6. Let $R = \{0, 1\}$ Define an addition and multiplication on R by

$\boxplus$	0	1
0	1	0
1	0	1

and

$\boxdot$	0	1
0	0	1
1	1	1

Is $(R, \boxplus, \boxdot)$ a ring?

Note that 1 is an additive identity, so $0_R = 1$. Also 0 is a multiplicative identity. So $1_R = 0$. Using the symbols 0_R and 1_R we can write the addition and multiplication table as follows:

$\boxplus$	0_R	1_R
0_R	0_R	1_R
1_R	1_R	0_R

and

$\boxdot$	0_R	1_R
0_R	0_R	0_R
1_R	0_R	1_R

Indeed, most entries in the tables are determined by the fact that 0_R and 1_R are the additive and multiplicative identity, respectively. Also $1_R \boxplus 1_R = 0 \boxplus 0 = 1 = 0_R$ and $0_R \boxdot 0_R = 1 \boxdot 1 = 1 = 0_R$.

Observe now that the new tables are the same as for $\mathbb{Z}_2$. So $(R, \boxplus, \boxdot)$ is a ring.

Theorem 3.1.7. *Let R and S be rings. Recall from 0.3.3 that $R \times S = \{(r, s) \mid r \in R, s \in S\}$. Define an addition and multiplication on $R \times S$ by*

$$\begin{aligned} (r, s) + (r', s') &= (r + r', s + s') \\ (r, s)(r', s') &= (rr', ss') \end{aligned}$$

for all $r, r' \in R$ and $s, s' \in S$. Then

(a) *$R \times S$ is a ring;*

(b) *$0_{R\times S} = (0_R, 0_S)$;*

(c) *$-(r, s) = (-r, -s)$ for all $r \in R, s \in S$;*

(d) *if R and S are both commutative, then so is $R \times S$;*

(e) *if both R and S have an identity, then $R \times S$ has an identity and $1_{R\times S} = (1_R, 1_S)$.*

Proof. See Exercise 3.1.#2. □

Example 3.1.8. Determined the addition table of the ring $\mathbb{Z}_2 \times \mathbb{Z}_3$.

Recall from 2.1.2(b) that $\mathbb{Z}_2 = \{0, 1\}$ and $\mathbb{Z}_3 = \{0, 1, 2\}$. So

$$\mathbb{Z}_2 \times \mathbb{Z}_3 = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2).\}$$

and

$+$	$(0,0)$	$(0,1)$	$(0,2)$	$(1,0)$	$(1,1)$	$(1,2)$
$(0,0)$	$(0,0)$	$(0,1)$	$(0,2)$	$(1,0)$	$(1,1)$	$(1,2)$
$(0,1)$	$(0,1)$	$(0,2)$	$(0,0)$	$(1,1)$	$(1,2)$	$(1,0)$
$(0,2)$	$(0,2)$	$(0,0)$	$(0,1)$	$(1,2)$	$(1,0)$	$(1,1)$
$(1,0)$	$(1,0)$	$(1,1)$	$(1,2)$	$(0,0)$	$(0,1)$	$(0,2)$
$(1,1)$	$(1,1)$	$(1,2)$	$(1,0)$	$(0,1)$	$(0,2)$	$(0,0)$
$(1,2)$	$(1,2)$	$(1,0)$	$(1,1)$	$(0,2)$	$(0,0)$	$(0,1)$

Exercises 3.1:

#1. Let $E = \{0, e, b, c\}$ with addition and multiplication defined by the following tables. Assume associativity and distributivity and show that R is a ring with identity. Is R commutative?

$+$	0	e	b	c
0	0	e	b	c
e	e	0	c	b
b	b	c	0	e
c	c	b	e	0

$\cdot$	0	e	b	c
0	0	0	0	0
e	0	e	b	c
b	0	b	b	0
c	0	c	0	c

#2. Prove Theorem 3.1.7.

3.2 Elementary Properties of Rings

Lemma 3.2.1. *Let R be ring and $a, b \in R$. Then $(a + b) + (-b) = a$.*

Proof.

$$\begin{aligned} (a+b)+(-b) &= a+(b+(-b)) && -\text{Ax } 2 \\ &= a + 0_R && -\text{Ax } 5 \\ &= a && -\text{Ax } 4 \end{aligned}$$

□

Theorem 3.2.2 (Additive Cancellation Law)**.** *Let R be ring and $a, b, c \in R$. Then*

$$\begin{array}{rc} & a = b \\ \iff & c+a = c+b \\ \iff & a+c = b+c \end{array}$$

Proof. "First Statement $\Longrightarrow$ Second Statement': Suppose that $a = b$. Then $c + a = c + b$ by the Principal of Substitution 0.1.1.

"Second Statement $\Longrightarrow$ Third Statement': Suppose that $c + a = c + b$. Then Ax 3 applied to each side of the equation gives $a + c = b + c$.

"Third Statement $\Longrightarrow$ First Statement': Suppose that $a + c = b + c$. Adding $-c$ to both sides of the equation gives $(a + c) + (-c) = (b + c) + (-c)$. Applying 3.2.1 to both sides gives $a = b$. □

Definition 3.2.3. *Let R be a ring and $c \in R$. Then c is called an* additive identity *of R if*

$$a + c = a \qquad \text{and} \qquad c + a = a$$

for all $a \in R$.

Corollary 3.2.4 (Additive Identity Law)**.** *Let R be a ring and $a, c \in R$. Then the following three statements are equivalent:*

$$\begin{array}{rrcl} & a & = & 0_R \\ \iff & c+a & = & c \\ \iff & a+c & = & c \end{array}$$

In particular, 0_R is the unique additive identity of R.

Proof. Put $b = 0_R$. Then by Ax 4 $c + b = c$ and $b + c = c$. Thus by the Principal of Substitution:

$$\begin{array}{rclcrcl} a & = & 0_R & \iff & a & = & b \\ c+a & = & c & \iff & c+a & = & c+b \\ a+c & = & c & \iff & a+c & = & b+c \end{array}$$

So the Corollary follows from the Cancellation Law 3.2.2. □

Definition 3.2.5. *Let R be a ring and $c \in R$. An* additive inverse *of c is an element a in R with $c + a = 0_R$.*

Corollary 3.2.6 (Additive Inverse Law). *Let R be a ring and $a, c \in R$. Then*

$$\begin{array}{rrcl} & a & = & -c \\ \iff & c+a & = & 0_R \\ \iff & a+c & = & 0_R \end{array}$$

In particular, $-c$ is the unique additive inverse of c.

Proof. Put $b = -c$. By Ax 5, $c + b = 0_R$ and so by Ax 3, $b + c = 0_R$. Thus by the Principal of Substitution:

$$\begin{array}{rclcrcl} a & = & -c & \iff & a & = & b \\ c+a & = & 0_R & \iff & c+a & = & c+b \\ a+c & = & 0_R & \iff & a+c & = & b+c \end{array}$$

So the Corollary follows from the Cancellation Law 3.2.2. □

Definition 3.2.7. *Let $(R, +, \cdot)$ be a ring and S a subset of R. Then $(S, +, \cdot)$ is called a* subring *of $(R, +, \cdot)$ provided that $(S, +, \cdot)$ is a ring.*

Theorem 3.2.8 (Subring Theorem). *Suppose that R is a ring and S a subset of R. Then S is a subring of R if and only if the following four conditions hold:*

(I) $0_R \in S$.

(II) *S is closed under addition (that is : if $a, b \in S$, then $a + b \in S$);*

(III) *S is closed under multiplication (that is: if $a, b \in S$, then $ab \in S$);*

(IV) *S is closed under negatives (that is: if $a \in S$, then $-a \in S$)*

Proof. $\Longrightarrow$: Suppose first that S is a subring of R.

By Ax 4 for S there exists $0_S \in S$ with $0_S + a = a$ for all $a \in S$. In particular, $0_S + 0_S = 0_S$. So by 3.2.4

$$(*) \qquad 0_S = 0_R.$$

Since $0_S \in S$, this gives $0_R \in S$ and (I) holds.

By Ax 1 for S, $a + b \in S$ for all $a, b \in S$. So (II) holds.

By Ax 6 for S, $ab \in S$ for all $a, b \in S$. So (III) holds.

Let $s \in S$. Then by Ax 5 for S, there exists $t \in S$ with $s + t = 0_S$. By $(*)$ $0_S = 0_R$ and so $s + t = 0_R$. Thus by 3.2.6 $t = -s$. Since $t \in S$ this gives $-s \in S$ and (IV) holds.

$\Longleftarrow$: Suppose now that (I)-(IV) hold.

Since S is a subset of R, S is a set. Hence Condition (i) in the definition of a ring holds for S.

Since S is a subset of R, $S \times S$ is a subset $R \times R$. By Conditions (ii) and (iii) in the definition of a ring, $R \times R$ is a subset of the domains of $+$ and $\cdot$. Hence also $S \times S$ is a subset of the domains of $+$ and $\cdot$. Thus Conditions (ii) and (iii) in the definition of a ring hold for S.

By (II) $a + b \in S$ for all $a, b \in S$ and so Ax 1 holds for S.

By Ax 2 $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$. Since $S \subseteq R$ we conclude that $(a + b) + c = a + (b + c)$ for all $a, b, c \in S$. Thus Ax 2 holds for S.

Similarly, since Ax 3 for all elements in R it also holds for all elements of S.

Put $0_S := 0_R$. Then (I) implies $0_S \in S$. By Ax 4 for R, $a = 0_R + a$ and $a = a + 0_R$ for all $a \in R$. Thus $a = 0_S + a$ and $a = a + 0_S$ for all $a \in S$ and so Ax 4 holds for S.

Let $s \in S$. Then $s + (-s) = 0_R$ and since $0_S = 0_R$, $s + (-s) = 0_S$. By (IV) $-s \in S$ and so Ax 5 holds for S.

By (III) $ab \in S$ for all $a, b \in S$ and so Ax 6 holds for S.

Since Ax 7 and Ax 8 hold for all elements of R they also holds for all elements of S. Thus Ax 7 and Ax 8 holds for S.

So Ax 1-Ax 8 hold for S and thus S is a ring. Hence, by definition, S is a subring of R. □

Example 3.2.9. (1) Show that $\mathbb{Z}$ is a subring of $\mathbb{Q}$, $\mathbb{Q}$ is a subring of $\mathbb{R}$ and $\mathbb{R}$ is a subring of $\mathbb{C}$.

By example 3.1.4 $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are rings. So by definition of a subring, $\mathbb{Z}$ is a subring of $\mathbb{Q}$, $\mathbb{Q}$ is a subring of $\mathbb{R}$ and $\mathbb{R}$ is a subring of $\mathbb{C}$.

(2) Let $n \in \mathbb{Z}$ and put $n\mathbb{Z} := \{nk \mid k \in \mathbb{Z}\}$. Show that $n\mathbb{Z}$ is subring of $\mathbb{Z}$.

We will verify the four conditions of the Subring Theorem for $S = n\mathbb{Z}$.

Observe first that since $n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$,

$$(*) \qquad a \in n\mathbb{Z} \iff \text{there exists } k \in \mathbb{Z} \text{ with } a = nk.$$

Let $a, b \in n\mathbb{Z}$. Then by $(*)$

$$(**) \qquad a = nk \quad \text{and} \quad b = nl,$$

for some $k, l \in \mathbb{Z}$.

(I): $0 = n0$ and so $0 \in n\mathbb{Z}$ by $(*)$

(II): $a + b \overset{(**)}{=} nk + nl = n(k + l)$. Since $k + l \in \mathbb{Z}$, $(*)$ shows $a + b \in \mathbb{Z}$. So $n\mathbb{Z}$ is closed under addition.

(III): $ab \overset{(**)}{=} (nk)(nl) = n(knl)$. Since $nkl \in \mathbb{Z}$, $(*)$ shows $ab \in \mathbb{Z}$. So $n\mathbb{Z}$ is closed under multiplication.

(IV): $-a \overset{(**)}{=} -(nk) = n(-k)$. Since $-k \in \mathbb{Z}$, $(*)$ shows $-a \in \mathbb{Z}$. So $n\mathbb{Z}$ is closed under negatives.

(3) Show that $\{[0]_4, [2]_4\}$ is a subring of $\mathbb{Z}_4$.

We compute in $\mathbb{Z}_4$: $0_{\mathbb{Z}_4} = 0 \in \{0, 2\}$ and so Condition (I) of the Subring Theorem holds. We compute :

$$\begin{array}{c|cc} + & 0 & 2 \\ \hline 0 & 0 & 2 \\ 2 & 2 & 0 \end{array} \qquad \begin{array}{c|cc} \cdot & 0 & 2 \\ \hline 0 & 0 & 0 \\ 2 & 0 & 0 \end{array} \quad \text{and} \quad \begin{array}{c|cc} x & 0 & 2 \\ \hline -x & 0 & 2 \end{array}$$

So $\{0, 2\}$ is closed under addition, multiplication and negatives. Thus $\{0, 2\}$ is a subring of $\mathbb{Z}_4$ by Subring Theorem.

Definition 3.2.10. *Let R be a ring and $a, b \in R$. Then $a - b := a + (-b)$.*

Proposition 3.2.11. *Let R be a ring and $a, b, c \in R$. Then*

(a) $-0_R = 0_R$

(b) $a - 0_R = a$.

(c) $a \cdot 0_R = 0_R = 0_R \cdot a$.

(d) $a \cdot (-b) = -(ab) = (-a) \cdot b$.

(e) $-(-a) = a$.

(f) $a - b = 0_R$ *if and only if* $a = b$.

(g) $-(a + b) = (-a) + (-b) = (-a) - b$.

(h) $-(a - b) = (-a) + b = b - a$.

(i) $(-a) \cdot (-b) = ab$.

(j) $a \cdot (b - c) = ab - ac$ *and* $(a - b) \cdot c = ac - bc$.

If R has an identity 1_R,

(k) $(-1_R) \cdot a = -a = a \cdot (-1_R)$.

Proof. (a) By Ax 4 $0_R + 0_R = 0_R$ and so by the Additive Inverse Law 3.2.6 $0_R = -0_R$.

(b) $a - 0_R \overset{\text{Def: -}}{=} a + (-0_R) \overset{\text{(a)}}{=} a + 0_R \overset{\text{Ax 4}}{=} a$.

(c) We compute

$$a \cdot 0_R \overset{\text{Ax 4}}{=} a \cdot (0_R + 0_R) \overset{\text{Ax 8}}{=} a \cdot 0_R + a \cdot 0_R,$$

and so by the Additive Identity Law 3.2.4 $a \cdot 0_R = 0_R$. Similarly $0_R \cdot a = 0_R$.

(d) We have

$$ab + a \cdot (-b) \overset{\text{Ax 8}}{=} a \cdot (b + (-b)) \overset{\text{Def} -b}{=} a \cdot 0_R \overset{\text{(c)}}{=} 0_R.$$

So by the Additive Inverse Law 3.2.6 $-(ab) = a \cdot (-b)$.

(e) By Ax 5 , $a + (-a) = 0_R$ and so by the Additive Inverse Law 3.2.6, $a = -(-a)$.

(f)

$$\begin{array}{rlr} & a-b=0_R & \\ \iff & a+(-b)=0_R & -\text{ definition of -} \\ \iff & a=-(-b) & -\text{ Additive Inverse Law 3.2.6} \\ \iff & a=b & -\text{ (e)} \end{array}$$

(g)

$$\begin{array}{rlrl} (a+b)+((-a)+(-b)) & \overset{\text{Ax 3}}{=} (b+a)+((-a)+(-b) & \overset{\text{Ax 2}}{=} & ((b+a)+(-a))+(-b) \\ & \overset{3.2.1}{=} \quad b+(-b) & \overset{\text{Ax 5}}{=} & 0_R. \end{array}$$

and so by the Additive Inverse Law 3.2.6 $-(a+b)=(-a)+(-b)$. By definition of " $-$ ", $(-a)+(-b)=(-a)-b$.

(h)

$$\begin{array}{rlrl} -(a-b) & \overset{\text{Def -}}{=} -(a+(-b)) & \overset{(g)}{=} (-a)+(-(-b)) & \overset{(e)}{=} (-a)+b \\ & & \overset{\text{Ax 3}}{=} \quad b+(-a) & \overset{\text{Def -}}{=} \quad b-a \end{array}.$$

(i) $(-a)\cdot(-b) \overset{(d)}{=} a\cdot(-(-b)) \overset{(e)}{=} a\cdot b.$

(j) $a\cdot(b-c) \overset{\text{Def -}}{=} a\cdot(b+(-c)) \overset{\text{Ax 8}}{=} a\cdot b+a\cdot(-c) \overset{(d)}{=} ab+(-(ac)) \overset{\text{Def -}}{=} ab-ac.$
Similarly $(a-b)\cdot c = ab-ac.$

(k) Suppose now that R has an additive identity. Then

$$a+((-1_R)\cdot a) \overset{(\text{Ax 10})}{=} 1_R\cdot a+(-1_R)\cdot a \overset{\text{Ax 8}}{=} (1_R+(-1_R))\cdot a \overset{\text{Ax 5}}{=} 0_R\cdot a \overset{(c)}{=} 0_R.$$

Hence by the Additive Inverse Law 3.2.6 $-a=(-1_R)\cdot a$. Similarly, $-a=a\cdot(-1_R)$. □

Lemma 3.2.12. *Let R be ring and $a,b,c\in R$. Then*

$$\begin{array}{rrcl} & c & = & b-a \\ \iff & c+a & = & b \\ \iff & a+c & = & b \end{array}$$

Proof.

$$\begin{array}{rrcll} & a+c & = & b & \\ \iff & c+a & = & b & -\text{ Ax 3} \\ \iff & (c+a)+(-a) & = & b+(-a) & -\text{ Additive Cancellation Law 3.2.2} \\ \iff & c & = & b-a & -\text{ 3.2.1 and Definition of } b-a \end{array}$$

□

Definition 3.2.13. *Let R be a ring with identity.*

(a) *Let $u \in R$. Then u is called a* unit *in R if there exists an element in R, denoted by u^{-1} and called 'u-inverse', with*

$$uu^{-1} = 1_R = u^{-1}u.$$

(b) *Let $u, v \in R$. Then v is called an (multiplicative) inverse of u if $uv = 1_R = vu$.*

(c) *Let $e \in R$. Then e is called an (multiplicative) identity of R, if $ea = a = ae$ for all $a \in R$.*

Example 3.2.14. Find the units in $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{Z}_6$.

Units in $\mathbb{Z}$: Let u be a unit in $\mathbb{Z}$. Then $uv = 1$ for some $v \in \mathbb{Z}$. So $u|1$ and so by 1.2.1 $1 \leq |u| \leq 1$. Hence $|u| = 1$ and ± 1 are the only units in $\mathbb{Z}$.

Units in $\mathbb{Q}$: Let u is a non-zero rational number. Then $u = \frac{n}{m}$ with $n, m \in \mathbb{Z}$ with $n \neq 0$ and $m \neq 0$. Thus $\frac{1}{u} = \frac{m}{n}$ is rational. So all non-zero elements in $\mathbb{Q}$ are units.

Units in $\mathbb{Z}_6$: By 2.1.2 $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ and so $\mathbb{Z}_6 = \{0, \pm 1, \pm 2, 3\}$. We compute

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

So ± 1 (that is 1 and 5) are the only units in $\mathbb{Z}_6$.

Lemma 3.2.15. (a) *Let R be a ring and e and $e' \in R$. Suppose that*

$$(*) \quad ea = a \qquad \text{and} \qquad (**) \quad ae' = a$$

for all $a \in R$. Then $e = e'$ and e is a multiplicative identity in R. In particular, a ring has at most one multiplicative identity.

(b) *Let R be a ring with identity and $x, y, u \in R$ with*

$$(+) \quad xu = 1_R \qquad \text{and} \qquad (++) \quad uy = 1_R.$$

Then $x = y$, u is a unit in R and x is an inverse of u.

Proof. (a)

$$e \overset{(*)}{=} ee' \overset{(**)}{=} e'$$

(b)

$$y \overset{(\text{Ax } 10)}{=} 1_R y \overset{(+)}{=} (xu)y \overset{\text{Ax } 7}{=} x(uy) \overset{(++)}{=} x1_R \overset{(\text{Ax } 10)}{=} x.$$

□

Theorem 3.2.16 (Multiplicative Inverse Law)**.** *Let R be a ring with identity and $u, v \in R$. Suppose u is a unit. Then*

$$\begin{array}{rrcl} & v & = & u^{-1} \\ \iff & vu & = & 1_R \\ \iff & uv & = & 1_R \end{array}$$

In particular, u^{-1} is the unique multiplicative inverse of u.

Proof. Recall first that by definition of unit:

$$(*) \quad uu^{-1} = 1_R \qquad \text{and} \qquad (**) \quad u^{-1}u = 1_R$$

First Statement $\Longrightarrow$ Second Statement': Suppose $v = u^{-1}$. Then $vu = u^{-1}u \overset{(**)}{=} 1_R$.
'Second Statement $\Longrightarrow$ Third Statement': Suppose that $vu = 1_R$. By (*) $uu^{-1} = 1_R$. Thus by 3.2.15(b) applied with $x = v$ and $y = u^{-1}$ we jave $v = u^{-1}$ and so $uv = uu^{-1} \overset{(*)}{=} 1_R$.

'Third Statement $\Longrightarrow$ First Statement': Suppose that $uv = 1_R$. By (**) $u^{-1}u = 1_R$. Thus 3.2.15 applied with $x = u^{-1}$ and $y = v$ gives $u^{-1} = v$. □

Lemma 3.2.17. *Let R be a ring with identity and a and b units in R.*

(a) *a^{-1} is a unit and $(a^{-1})^{-1} = a$.*

(b) *ab is a unit and $(ab)^{-1} = b^{-1}a^{-1}$.*

Proof. (a) By definition of a^{-1}, $aa^{-1} = 1_R = a^{-1}a$. Hence also $a^{-1}a = 1_R = aa^{-1}$. Thus a^{-1} is a unit and by the Multiplicative Inverse Law 3.2.16, $a = (a^{-1})^{-1}$.

(b) See Exercise 3.2.#7. □

Definition 3.2.18. *A ring R is called an* integral domain *provided that R is commutative, R has an identity, $1_R \neq 0_R$ and*

(Ax 11) *whenever $a, b \in R$ with $ab = 0_R$, then $a = 0_R$ or $b = 0_R$.*

Theorem 3.2.19 (Multiplicative Cancellation Law for Integral Domains)**.** *Let R be an integral domain and $a, b, c \in R$ with $a \neq 0_R$. Then*

$$\begin{array}{ccccc} & ab & = & ac \\ \iff & b & = & c \\ \iff & ba & = & ca \end{array}$$

Proof. 'First Statement $\Longrightarrow$ Second Statement:' Suppose $ab = ac$. Then

$$\begin{aligned} a(b-c) =& ab - ac && \text{3.2.11(j)} \\ =& ab - ab && \text{Principal of Substitution}, ab = ac \\ =& 0_R && \text{3.2.11(f)} \end{aligned}$$

Since R is an integral domain, (Ax 11) holds. So $a(b-c) = 0_R$ implies $a = 0_R$ or $b - c = 0_R$. By assumption $a \neq 0_R$ and so $b - c = 0_R$. Thus by 3.2.11(f), $b = c$.

'Second Statement $\Longrightarrow$ Third Statement:' If $b = c$ then $ab = ac$ by the Principal of Substitution.

'Third Statement $\Longrightarrow$ First Statement:' Since integral domains are commutative, $ba = ca$ implies $ab = ac$. □

Definition 3.2.20. *A ring R is called a* field *provided that R is commutative, R has an identity, $1_R \neq 0_R$ and*

(Ax 12) *each $a \in R$ with $a \neq 0_R$ is a unit in R.*

Example 3.2.21. Which of the following rings are fields? Which are integral domains?

(a) $\mathbb{Z}$.

(b) $\mathbb{Q}$.

(c) $\mathbb{R}$.

(d) $\mathbb{Z}_3$.

(e) $\mathbb{Z}_6$.

(f) $\mathrm{M}_2(\mathbb{R})$.

(g) $\mathbb{Z}_p$, p a prime.

All of the rings have a non-zero identity. All but $\mathrm{M}_2(\mathbb{R})$ are commutative. If a, b are non zero real numbers then $ab \neq 0$. So (Ax 11) holds for $\mathbb{R}$ and so also for $\mathbb{Z}$ and $\mathbb{Q}$. Thus $\mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$ are integral domains.

(a) 2 does not have an inverse in $\mathbb{Z}$. So $\mathbb{Z}$ is an integral domain, but not a field.

(b) The inverse of a non-zero rational numbers is rational. So $\mathbb{Q}$ is a integral domain and a field.

(c) The inverse of a non-zero real numbers is real. So $\mathbb{R}$ is a integral domain and a field.

(d) ± 1 are the only non-zero elements in $\mathbb{Z}_3$. $1 \cdot 1 = 1$ and $-1 \cdot -1 = 1$. So ± 1 are units and $\mathbb{Z}_3$ is a field. Also $\pm 1 \cdot \pm 1 = \pm 1 \neq 0$ and so $\mathbb{Z}_3$ is an integral domain.

(e) . By 3.2.14 the units in $\mathbb{Z}_6$ are ± 1 and ± 3. Thus 2 is not a unit and so $\mathbb{Z}_6$ is not a field. Note that $2 \cdot 3 = 6 = 0$ in $\mathbb{Z}_6$ and so $\mathbb{Z}_6$ is not an integral domain

(f) Note that $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. So $\mathrm{M}_2(\mathbb{R})$ is not commutative. Also $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not a unit and $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. So $\mathrm{M}_2(\mathbb{R})$ fails all conditions of a field and integral domain, except for $1_R \neq 0_R$.

(g) By 2.3.6 each non-zero element in $\mathbb{Z}_p$ has an inverse. So $\mathbb{Z}_p$ is a field. Let $A, B \in \mathbb{Z}$ with $AB = [0]_p$. Then by 2.3.2 $A = [0]_p$ or $B = [0]_p$. Thus $\mathbb{Z}_p$ is an integral domain.

Proposition 3.2.22. *Every field is an integral domain.*

Proof. Let F be a field. Then by definition, F is an commutative ring with identity and $1_F \neq 0_F$. So it remains the verify Ax 11 in 3.2.18. For this let $a, b \in F$ with

$$(*) \qquad ab = 0_F.$$

Suppose that $a \neq 0_F$. Then by the definition of a field, a is a unit. Thus a has multiplicative inverse a^{-1}. So we compute

$$0_F \overset{3.2.11(c)}{=} a^{-1} \cdot 0_F \overset{(*)}{=} a^{-1} \cdot (a \cdot b) \overset{\text{Ax 7}}{=} (a^{-1} \cdot a) \cdot b \overset{\text{Def: } a^{-1}}{=} 1_F \cdot b \overset{(\text{Ax 10})}{=} b.$$

So $b = 0_F$.

We have proven that if $a \neq 0_F$, then $b = 0_F$. So $a = 0_F$ or $b = 0_F$. Hence Ax 11 holds and F is an integral domain. □

Theorem 3.2.23. *Every finite integral domains is a field.*

Proof. Let R be a finite integral domain. Then R is a commutative ring with identity and $1_R \neq 0_R$. So it remains to show that every $a \in R$ with $a \neq 0_R$ is a unit. Set $S := \{ar \mid r \in R\}$. Define a function f by

$$f : R \to S, r \to ar.$$

Let $b, c \in R$ with $f(b) = f(c)$. Then $ab = ac$ and by the Cancellation Law 3.2.19 $b = c$. Thus f is 1-1. Also

$$\operatorname{Im} f = \{f(r) \mid r \in R\} = \{ar \mid r \in R\} = S,$$

and so f is onto. Hence f is a bijection and so $|R| = |S|$. Since $S \subseteq R$ and R is finite we conclude $R = S$. In particular, $1_R \in S$ and so there exists $b \in R$ with $1_R = ab$. Since R is commutative we also have $ba = 1_R$ and so a is a unit. □

Definition 3.2.24. *Let R be a ring and $a \in R$.*

(a) *Let $n \in \mathbb{Z}^+$. Then a^n is inductively defined by $a^1 = a$ and $a^{n+1} = a^n a$.*

(b) *If R has an identity, then $a^0 = 1_R$.*

(c) *If R has an identity and a is a unit, then $a^{-n} = (a^{-1})^n$ for all $n \in \mathbb{Z}^+$.*

Exercises 3.2:

#1. Let R be a ring and $a \in R$. Let $n, m \in \mathbb{Z}$ such that a^n and a^m are defined. (So $n, m \in \mathbb{Z}^+$, or R has an identity and $n, m \in \mathbb{N}$, or R has identity, a is a unit and $n, m \in \mathbb{Z}$.) Show that

(a) $a^n a^m = a^{n+m}$.

(b) $a^{nm} = (a^n)^m$.

#2. Prove or disprove:

(a) If R and S are integral domains, then $R \times S$ is an integral domain.

(b) If R and S are fields, then $R \times S$ is a field.

#3. Which of the following six sets are subrings of $\mathrm{M}_2(\mathbb{R})$? Which ones have an identity?

(a) All matrices of the form $\begin{bmatrix} 0 & r \\ 0 & 0 \end{bmatrix}$ with $r \in \mathbb{Q}$.

(b) All matrices of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ with $a, b, c \in \mathbb{Z}$.

(c) All matrices of the form $\begin{bmatrix} a & a \\ b & b \end{bmatrix}$ with $a, b \in \mathbb{R}$.

(d) All matrices of the form $\begin{bmatrix} a & 0 \\ a & 0 \end{bmatrix}$ with $a, b \in \mathbb{R}$.

(e) All matrices of the form $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ with $a \in \mathbb{R}$.

(f) All matrices of the form $\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$ with $a \in \mathbb{R}$.

#4. Let $\mathbb{Z}[i]$ denote the set $\{a + bi \mid a, b \in \mathbb{Z}\}$. Show that $\mathbb{Z}[i]$ is a subring of $\mathbb{C}$.

#5. An element e of a ring is said to be an **idempotent** if $e^2 = e$.

(a) Find four idempotents in $M(\mathbb{R})$.

(b) Find all idempotents in $\mathbb{Z}_{12}$.

(c) Prove that the only idempotents in an integral domain R are 0_R and 1_R.

#6. Let R be a ring and b a fixed element of R. Let $T = \{rb \mid r \in R\}$. Prove that T is a subring of R.

#7. (a) If a and b are units in a ring with identity, prove that ab is a unit with inverse $b^{-1}a^{-1}$.

(b) Give an example to show that if a and b are units, then $a^{-1}b^{-1}$ does not need to be the multiplicative inverse of ab.

#8. Let R be a ring with identity. If ab and a are units in R, prove that b is a unit.

#9. Let R be a commutative ring with identity $1_R \neq 0_R$. Prove that R is an integral domain if and only if cancellation holds in R, (that is whenever $a, b, c \in R$ with $a \neq 0_R$ and $ab = ac$ then $b = c$.)

3.3 Isomorphism and Homomorphism

Definition 3.3.1. *Let $(R, +, \cdot)$ and $(S, \oplus, \odot)$ be rings and let $f : R \to S$ be a function.*

(a) *f is called a* homomorphism *from $(R, +, \cdot)$ to $(S, \oplus, \odot)$ if*

$$f(a + b) = f(a) \oplus f(b) \qquad \text{[}f\text{ respects addition]}$$

and

$$f(a \cdot b) = f(a) \odot f(b) \qquad \text{[}f\text{ respects multiplication]}$$

for all $a, b \in R$.

(b) *f is called an* isomorphism *from $(R, +, \cdot)$ to $(S, \oplus, \odot)$, if f is a homomorphism from $(R, +, \cdot)$ to $(S, \oplus, \odot)$ and f is 1-1 and onto*

(c) *$(R, +, \cdot)$ is called* isomorphic *to $(S, \oplus, \odot)$, if there exists an isomorphism from $(R, +, \cdot)$ to $(S, \oplus, \odot)$.*

Example 3.3.2. (1) Consider $f : \mathbb{Z} \to \mathbb{R}, a \to a$.

Let $a, b \in \mathbb{Z}$. Then

$$f(a + b) = a + b = f(a) + f(b) \qquad \text{and} \qquad f(ab) = ab = f(a)f(b)$$

and so f is homomorphism. f is 1-1, but not onto. Hence f is not an isomorphism.

(2) Consider $g : \mathbb{R} \to \mathbb{R}, a \to -a$.

Let $a, b \in \mathbb{R}$. Then

$$g(a+b) = -(a+b) = -a + (-b) = g(a) + g(b).$$

and so g respects addition.

$$g(ab) = -(ab) \qquad \text{and} \qquad g(a)g(b) = (-a)(-b) = ab$$

For $a = b = 1$ we conclude that

$$g(1 \cdot 1) = -(1 \cdot 1) = -1 \qquad \text{and} \qquad g(1)g(1) = 1 \cdot 1 = 1.$$

So $g(1 \cdot 1) \neq g(1) \cdot g(1)$. Thus g does not respect multiplication, and g is not a homomorphism. But note that g is 1-1 and onto.

(3) Let R and S be rings and consider $h : R \to S, r \to 0_S$.

Let $a, b \in R$. Then

$$g(a+b) = 0_S = 0_S + 0_S = g(a) + g(b) \qquad \text{and} \qquad g(ab) = 0_S = 0_S 0_S = g(a)g(b).$$

So g is a homomorphism. g is 1-1 if and only if $R = \{0_R\}$ and g is onto if and only if $S = \{0_S\}$. Hence g is an isomorphism if and only if $R = \{0_R\}$ and $S = \{0_S\}$.

(4) Let R be a ring. Consider $\mathrm{id}_R : R \to R, r \to r$

Let $a, b \in R$. Then

$$\mathrm{id}_R(a+b) = a + b = \mathrm{id}_R(a) + \mathrm{id}_R(b) \qquad \text{and} \qquad \mathrm{id}_R(ab) = ab = \mathrm{id}_R(a)\mathrm{id}_R(b)$$

and so id_R is a homomorphism. Since id_R is 1-1 and onto, id_R is an isomorphism.

(5) Let n be a non-zero integer. Consider $h : \mathbb{Z} \to \mathbb{Z}_n, a \to [a]_n$.

Let $a, b \in \mathbb{Z}$. By definition of addition and multiplication in $\mathbb{Z}_n$

$$h(a+b) = [a+b]_n = [a]_n \oplus [b]_n = h(a) \oplus h(b) \qquad \text{and} \qquad h(ab) = [ab]_n = [a]_n \odot [b]_n = h(a) \odot h(b).$$

So h is homomorphism. Since

$$h(n) = [n]_n = [0]_n = h(0)$$

and $n \neq 0$, h is not 1-1. So h is not isomorphism.

Let $A \in \mathbb{Z}_n$. By definition of $\mathbb{Z}_n$, $A = [a]_n$ for some $a \in \mathbb{Z}$. Hence $h(a) = A$ and h is onto.

Example 3.3.3. Consider the function

$$f : \mathbb{C} \to \mathrm{M}_2(\mathbb{R}), r + si \to \begin{bmatrix} r & s \\ -s & r \end{bmatrix}$$

Let $a, b \in \mathbb{C}$. Then $a = r + si$ and $b = \tilde{r} + \tilde{s}$ for some $r, s, \tilde{r}, \tilde{s} \in \mathbb{R}$. So

$$\begin{aligned} f\Big(a + b\Big) &= f\Big((r + si) + (\tilde{r} + \tilde{s}i)\Big) \\ &= f\Big((r + \tilde{r}) + (s + \tilde{s})i\Big) \\ &= \begin{bmatrix} r + \tilde{r} & s + \tilde{s} \\ -(s + \tilde{s}) & r + \tilde{r} \end{bmatrix} \\ &= \begin{bmatrix} r & s \\ -s & r \end{bmatrix} + \begin{bmatrix} \tilde{r} & \tilde{s} \\ -\tilde{s} & \tilde{r} \end{bmatrix} \\ &= f(r + si) + f(\tilde{r} + \tilde{s}i) \\ &= f(a) + f(b) \end{aligned}$$

and

$$\begin{aligned} f\Big(ab\Big) &= f\Big((r + si)(\tilde{r} + \tilde{s}i)\Big) \\ &= f\Big((r\tilde{r} - s\tilde{s}) + (r\tilde{s} + s\tilde{r})i\Big) \\ &= \begin{bmatrix} r\tilde{r} - s\tilde{s} & r\tilde{s} + s\tilde{r} \\ -(r\tilde{s} + s\tilde{r}) & r\tilde{r} - s\tilde{s} \end{bmatrix} \\ &= \begin{bmatrix} r & s \\ -s & r \end{bmatrix} \begin{bmatrix} \tilde{r} & \tilde{s} \\ -\tilde{s} & \tilde{r} \end{bmatrix} \\ &= f(r + si)f(\tilde{r} + \tilde{s}i) \\ &= f(a)f(b). \end{aligned}$$

So f is a homomorphism. If $f(a) = f(b)$, then

$$\begin{bmatrix} r & s \\ -s & r \end{bmatrix} = \begin{bmatrix} \tilde{r} & \tilde{s} \\ -\tilde{s} & \tilde{r} \end{bmatrix}$$

and so $r = \tilde{r}$ and $s = \tilde{s}$. Hence $a = r + si = \tilde{r} + \tilde{s}i = b$ and so f is 1-1. Note that $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not of the form $\begin{bmatrix} r & s \\ -s & r \end{bmatrix}$ and so f is not onto.

Notation 3.3.4. (a) *'$f : R \to S$ is a ring homomorphism' stands for more precise '$(R, +, \cdot)$ and $(S, \oplus, \odot)$ are rings and f is a ring homomorphism from $(R, +, \cdot)$ to $(S, \oplus, \odot)$.'*

(b) *Usually we will use the symbols $+$ and $\cdot$ also for the addition and multiplication on S and so the conditions for a homomorphism become*

$$f(a+b) = f(a) + f(b) \quad \text{and} \quad f(ab) = f(a)f(b)$$

Remark 3.3.5. *Let $R = \{r_1, r_2, \ldots, r_n\}$ be a ring with n elements. Suppose that the addition and multiplication table is given by*

$$A : \begin{array}{c|cccccc} + & r_1 & \ldots & r_j & \ldots & r_n \\ \hline r_1 & a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_i & a_{i1} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_n & a_{n1} & \ldots & a_{nj} & \ldots & a_{nn} \end{array} \qquad \text{and} \qquad M : \begin{array}{c|cccccc} \cdot & r_1 & \ldots & r_j & \ldots & r_n \\ \hline r_1 & b_{11} & \ldots & b_{1j} & \ldots & b_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_i & b_{i1} & \ldots & b_{ij} & \ldots & b_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_n & b_{n1} & \ldots & b_{nj} & \ldots & b_{nn} \end{array}$$

So $r_i + r_j = a_{ij}$ and $r_i r_j = b_{ij}$ for all $1 \leq i, j \leq n$.

Let S be a ring and $f : R \to S$ a function. For $r \in R$ put $r' = f(r)$. Consider the tables A' and M' obtain from the tables A and M by replacing all entries by its image under f:

$$A' : \begin{array}{c|cccccc} & r'_1 & \ldots & r'_j & \ldots & r'_n \\ \hline r'_1 & a'_{11} & \ldots & a'_{1j} & \ldots & a'_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r'_i & a'_{i1} & \ldots & a'_{ij} & \ldots & a'_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r'_n & a'_{n1} & \ldots & a'_{nj} & \ldots & a'_{nn} \end{array} \qquad \text{and} \qquad M' : \begin{array}{c|cccccc} & r'_1 & \ldots & r'_j & \ldots & r'_n \\ \hline r'_1 & b'_{11} & \ldots & b'_{1j} & \ldots & b'_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r'_i & b'_{i1} & \ldots & b'_{ij} & \ldots & b'_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r'_n & b'_{n1} & \ldots & b'_{nj} & \ldots & b'_{nn} \end{array}$$

(a) *f is a homomorphism if and only if A' and M' are the tables for the addition and multiplication of the elements $r'_1, \ldots, r'_n$ in S, that is $r'_i + r'_j = a'_{ij}$ and $r'_i r'_j = b'_{ij}$ for all $1 \leq i, j \leq n$.*

(b) *f is 1-1 if and only if $r'_1, \ldots, r'_n$ are pairwise distinct.*

(c) *f is onto if and only if $S = \{r'_1, r'_2, \ldots, r'_n\}$.*

(d) *f is an isomorphism if and only if A' is an addition table for S and M' is a multiplication table for S.*

Proof. (a) f is a homomorphism if and only if

$$f(a+b) = a+b \quad \text{and} \quad f(ab) = f(a)f(b)$$

for all $a, b \in R$. Since $R = \{r_1, \ldots, r_n\}$, this holds if and only if

$$f(r_i + r_j) = f(r_i) + f(r_j) \quad \text{and} \quad f(r_i r_j) = f(r_i)f(r_j)$$

for all $1 \leq i, j \leq n$. Since $r_i + r_j = a_{ij}$ and $r_i r_j = b_{ij}$ this holds if and only if

$$f(a_{ij}) = f(r_i) + f(r_j) \quad \text{and} \quad f(b_{ij}) = f(r_i)f(r_j)$$

Since $f(r) = r'$, this is equivalent to

$$a'_{ij} = r'_i + r'_j \quad \text{and} \quad b'_{ij} = r'_i r'_j$$

(b) f is 1-1 if and only if for all $a, b \in R$, $f(a) = f(b)$ implies $a = b$ and so if and only if $a \neq b$ implies $f(a) \neq f(b)$. Since for each $a \in R$ there exists a unique $1 \leq i \leq n$ with $a = r_i$, f is 1-1 if and only for all $1 \leq i, j \leq n$, $i \neq j$ implies $f(r_i) \neq f(r_j)$, that is $i \neq j$ implies $r'_i \neq r'_j$.

(c) f is onto if and only if $\operatorname{Im} f = S$. Since $R = \{r_1, \ldots, r_n\}$, $\operatorname{Im} f = \{f(r_1), \ldots, f(r_n)\} = \{r'_1, \ldots, r'_n\}$. So f is onto if and only if $S = \{r'_1, \ldots, r'_n\}$.

(d) Follows from (a)-(c).

□

Example 3.3.6. Let R be the ring from example 3.1.6. Then the map

$$f : R \to \mathbb{Z}_2, 0 \to [1]_2, 1 \to [0]_2$$

is an isomorphism.

The tables for R are

$\boxplus$	0	1
0	1	0
1	0	1

and

$\boxdot$	0	1
0	0	1
1	1	1

Replacing 0 by $[1]_2$ and 1 by $[0]_2$ we obtain

$$\begin{array}{c|cc} & [1]_2 & [0]_2 \\ \hline [1]_2 & [0]_2 & [1]_2 \\ [0]_2 & [1]_2 & [0]_2 \end{array} \qquad \text{and} \qquad \begin{array}{c|cc} & [1]_2 & [0]_2 \\ \hline [1]_2 & [1]_2 & [0]_2 \\ [0]_2 & [0]_2 & [0]_2 \end{array}.$$

Note that these are addition and multiplication tables for $\mathbb{Z}_2$ and so by 3.3.5 f is an isomorphism.

Lemma 3.3.7. *Let $f : R \to S$ be a homomorphism of rings. Then*

(a) $f(0_R) = 0_S$.

(b) $f(-a) = -f(a)$ *for all* $a \in R$.

(c) $f(a - b) = f(a) - f(b)$ *for all* $a, b \in R$.

Suppose in addition that R has an identity and f is onto, then

(d) *S is a ring with identity and* $f(1_R) = 1_S$.

(e) *If u is a unit in R, then $f(u)$ is a unit in S and* $f(u^{-1}) = f(u)^{-1}$.

Proof. (a) We have

$$f(0_R) + f(0_R) \overset{\text{f hom}}{=} f(0_R + 0_R) \overset{\text{Ax 4}}{=} f(0_R).$$

So by the Additive Identity Law 3.2.4, $f(0_R) = 0_S$.

(b) We compute

$$f(a) + f(-a) \overset{\text{f hom}}{=} f(a + (-a)) \overset{\text{Ax 5}}{=} f(0_R) \overset{\text{(a)}}{=} 0_S,$$

and so by the Additive Inverse Law 3.2.6 $f(-a) = -f(a)$.

(c)

$$f(a - b) \overset{\text{Def} -}{=} f(a + (-b)) \overset{\text{f hom}}{=} f(a) + f(-b) \overset{\text{(b)}}{=} f(a) + (-f(b)) \overset{\text{def} -}{=} f(a) - f(b).$$

(d) We will first show that $f(1_R)$ is an identity in S. For this let $s \in S$. Then since f is onto, $s = f(r)$ for some $r \in R$. Thus

$$s \cdot f(1_R) = f(r)f(1_R) \overset{\text{f hom}}{=} f(r1_R) \overset{\text{(Ax 10)}}{=} f(r) = s,$$

and similarly $f(1_R) \cdot s = s$. So $f(1_R)$ is an identity in S. By 3.2.15(a) a ring has at most one identity and so $f(1_R) = 1_S$.

(e) Let u be a unit in R. We will first show that $f(u^{-1})$ is an inverse of $f(u)$:

$$f(u)f(u^{-1}) \overset{\text{f hom}}{=} f(uu^{-1}) \overset{\text{def inv}}{=} f(1_R) \overset{\text{(d)}}{=} 1_S.$$

Similarly $f(u^{-1})f(u) = 1_S$. Thus $f(u^{-1})$ is an inverse of $f(u)$ and so $f(u)$ is a unit. By 3.2.16 $f(u)^{-1}$ is the unique inverse of $f(u)$ and so $f(u^{-1}) = f(u)^{-1}$. $\square$

Example 3.3.8. Find all onto homomorphisms from $\mathbb{Z}_6$ to $\mathbb{Z}_2 \times \mathbb{Z}_3$.

Let $f : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ be an onto homomorphism. For $a, b \in \mathbb{Z}$ let

$$[a] := [a]_6, \qquad f[a] := f\big([a]_6\big), \quad \text{and} \quad [a, b] := \big([a]_2, [b]_3\big).$$

Since f is an onto homomorphism, we get from 3.3.7(d) that $f(1_{\mathbb{Z}_6}) = 1_{\mathbb{Z}_2 \times \mathbb{Z}_3}$. Since $[1]$ is the identity in $\mathbb{Z}_6$ and $[1, 1]$ is the identity in $\mathbb{Z}_2 \times \mathbb{Z}_3$ this gives $f[1] = [1, 1]$. Similarly, by 3.3.7(a), $f(0_{\mathbb{Z}_6}) = 0_{\mathbb{Z}_2 \times \mathbb{Z}_3}$ and thus $f[0] = [0, 0]$. We compute

$$\begin{aligned}
f[0] &= [0,0] \\
f[1] &= [1,1] \\
f[2] = f[1+1] = f[1] + f[1] = [1,1] + [1,1] = [2,2] &= [0,2] \\
f[3] = f[2+1] = f[2] + f[1] = [2,2] + [1,1] = [3,3] &= [1,0] \\
f[4] = f[3+1] = f[3] + f[1] = [3,3] + [1,1] = [4,4] &= [0,1] \\
f[5] = f[4+1] = f[4] + f[1] = [4,4] + [1,1] = [5,5] &= [1,2]
\end{aligned}$$

By 2.1.2 $\mathbb{Z}_6 = \{[0], [1], [2], [3], [4], [5]\}$, $\mathbb{Z}_2 = \{[0]_2, [1]_2\}$ and $\mathbb{Z}_3 = \{[0]_3, [1]_3, [2]_3\}$. Hence f is uniquely determined and

$$\mathbb{Z}_2 \times \mathbb{Z}_3 = \{(x, y) \mid x \in \mathbb{Z}_2, y \in \mathbb{Z}_3\} = \{[0,0], [0,1], [0,2], [1,0], [1,1], [1,2]\}.$$

We conclude that f is 1-1 and onto. Moreover,

$$(*) \qquad f[r] = [r, r] \text{ for all } 0 \leq r < 5.$$

We will show that the function $f : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ defined by (*) is a homomorphism. For this we first show that $f[m] = [m, m]$ for all $m \in \mathbb{Z}$. Indeed, by the Division Algorithm, $m = 6q + r$ with $q, r \in \mathbb{Z}$ and $0 \leq r < 6$. Then by 2.1.1 $[m]_6 = [r]_6$ and since $m = 2(3q) + r = 3(2q) + r$, $[m]_2 = [r]_2$ and $[m]_3 = [r]_3$. So $[m] = [r]$, $[m, m] = [r, r]$ and

$$(**) \qquad f[m] = f[r] = [r, r] = [m, m].$$

Note also that by the definition of addition and multiplication in the direct product $\mathbb{Z}_2 \times \mathbb{Z}_3$:

$$(***) \qquad [n+m, n+m] = [n, n] + [m, m] \quad \text{and} \quad [nm, nm] = [n, n][m, m]$$

Thus

$$f[n+m] \overset{(**)}{=} [n+m, n+m] \overset{(***)}{=} [n, n] + [m, m] \overset{(**)}{=} f[n] + f[m],$$

and

$$f[nm] \overset{(**)}{=} [nm,nm] \overset{(***)}{=} [n,n][m,m] \overset{(**)}{=} f[n]f[m].$$

So f is a homomorphism of rings. Since f is 1-1 and onto, f is an isomorphism and so $\mathbb{Z}_6$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_3$.

Example 3.3.9. Show that $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are not isomorphic.

Put $R := \mathbb{Z}_2 \times \mathbb{Z}_2$. Since $x + x = [0]_2$ for all $x \in \mathbb{Z}_2$ we also have

$$(x,y) + (x,y) = (x+x, y+y) = ([0]_2, [0]_2) = 0_R.$$

for all $x, y \in \mathbb{Z}_2$. Thus

$$(*) \qquad r + r = 0_R$$

for all $r \in R$. Let S be any ring isomorphic to R. We claim that $s + s = 0_S$ for all $s \in S$. Indeed, let $f : R \to S$ be an isomorphism and let $s \in S$. Since f is onto, there exists $r \in R$ with $f(r) = s$. Thus

$$s + s = f(r) + f(r) \overset{\text{f hom}}{=} f(r+r) \overset{(*)}{=} f(0_R) \overset{3.3.7(\text{a})}{=} 0_S$$

Since $[1]_4 + [1]_4 = [2]_4 \neq [0]_4$ we conclude that $\mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Corollary 3.3.10. *Let $f : R \to S$ be a homomorphism of rings. Then $\operatorname{Im} f$ is a subring of S. (Recall here that $\operatorname{Im} f = \{f(r) \mid r \in R\}$).*

Proof. It suffices to verify the four conditions in the Subring Theorem 3.2.8. Observe first that for $s \in S$,

$$(*) \qquad s \in \operatorname{Im} f \quad \Longleftrightarrow \quad s = f(r) \text{ for some } r \in R$$

Let $x, y \in \operatorname{Im} f$. Then by $(*)$:

$$(**) \qquad x = f(a) \quad \text{and} \quad y = f(b) \quad \text{for some } a, b \in R.$$

(I) By 3.3.7(a) $f(0_R) = 0_S$ and so $0_S \in \operatorname{Im} f$ by $(*)$

(II) $x + y \overset{(**)}{=} f(a) + f(b) \overset{\text{f hom}}{=} f(a+b)$. By Ax 1 $a + b \in R$. So $x + y \in \operatorname{Im} f$ by $(*)$.

(III) $xy \overset{(**)}{=} f(a)f(b) \overset{\text{f hom}}{=} f(ab)$. By Ax 6 $ab \in R$. So $xy \in \operatorname{Im} f$ by (*).

(IV) $-x \overset{(**)}{=} -f(a) \overset{3.3.7(\text{b})}{=} f(-a)$. By Ax 5 $-a \in R$. So $-x \in \operatorname{Im} f$ by (*). $\square$

Definition 3.3.11. Let R be a ring. For $n \in \mathbb{Z}$ and $a \in R$ define $na \in R$ as follows:

(i) $0a = 0_R$.

(ii) If $n \geq 0$ and na already has been defined, define $(n+1)a = na + a$.

(iii) If $n < 0$ define $na = -\Big((-n)a\Big)$.

Exercises 3.3:

#1. Let R be ring, $n, m \in \mathbb{Z}$ and $a, b \in R$. Show that

(a) $1a = a$.

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

(c) $(n+m)a = na + ma$.

(d) $(nm)a = n(ma)$.

(e) $n(a+b) = na + nb$.

(f) $n(ab) = (na)b = a(nb)$

#2. Let $f : R \to S$ be a ring homomorphism. Show that $f(na) = nf(a)$ for all $n \in \mathbb{Z}$ and $a \in R$.

#3. Let R be a ring. Show that:

(a) If $f : \mathbb{Z} \to R$ is a homomorphism, then $f(1)^2 = f(1)$.

(b) Let $a \in R$ with $a^2 = a$. Then there exists a unique homomorphism $g : \mathbb{Z} \to R$ with $g(1) = a$.

#4. Let $S = \left\{ \begin{bmatrix} a & b \\ b & a+b \end{bmatrix} \middle| \, a, b \in \mathbb{Z}_2 \right\}$. Given that S is a subring of $\mathrm{M}_2(\mathbb{Z}_2)$. Show that S is isomorphic to the ring R from Exercise 3.1.#1.

#5. (a) Give an example of a ring R and a function $f : R \to R$ such that $f(a+b) = f(a) + f(b)$ for all $a, b \in R$, but $f(ab) \neq f(a)(f(b)$ for some $a, b \in R$.

(b) Give an example of a ring R and a function $f : R \to R$ such that $f(ab) = f(a)f(b)$ for all $a, b \in R$, but $f(a+b) \neq f(a) + (f(b)$ for some $a, b \in R$.

#6. Let L be the ring of all matrices in $\mathrm{M}_2(\mathbb{Z})$ of the form $\begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$ with $a, b, c \in \mathbb{Z}$. Show that the function $f : L \to \mathbb{Z}$ given by $f\left(\begin{bmatrix} a & 0 \\ b & c \end{bmatrix}\right) = a$ is a surjective homomorphism but is not an isomorphism.

#7. Let n and m be positive integers with $n \equiv 1 \pmod{m}$. Define $f : \mathbb{Z}_m \to \mathbb{Z}_{nm}, [x]_m \to [xn]_{nm}$. Show that

(a) f is well-defined. (That is if x, y are integers with $[x]_m = [y]_m$, then $[xn]_{nm} = [yn]_{nm}$)

(b) f is a homomorphism.

(c) f is 1-1.

(d) If $n > 1$, then f is not onto.

#8. Let $f : R \to S$ be a ring homomorphism. Let B be a subring of S and define

$$A = \{r \in R \mid f(r) \in B\}.$$

Show that A is a subring of R.

3.4 Associates in commutative rings

Definition 3.4.1. *Let R be a commutative ring and $a, b \in R$. Then we say that a* divides *b in R and write $a|b$ if there exists $c \in R$ with $b = ac$* □

Lemma 3.4.2. *Let R be a commutative ring and $r \in R$. Then $0_R|r$ if and only of $r = 0_R$.*

Proof. By 3.2.11(c), $0_R = 0_R \cdot 0_R$ and so $0_R|0_R$.

Suppose now that $r \in R$ with $0_R|r$. Then there exists $s \in R$ with $r = 0_R s$ and so by 3.2.11(c), $r = 0_R$. □

Lemma 3.4.3. *Let R be a commutative ring and $a, b, c \in R$.*

(a) *$|$ is transitive, that is if $a|b$ and $b|c$, then $a|c$.*

(b) *$a|b \iff a|(-b) \iff (-a)|(-b) \iff (-a)|b$.*

(c) *If $a|b$ and $a|c$, then $a|(b + c)$ and $a|(b - c)$.*

(d) *If $a|b$ and $a|c$, then $a|(bu + cv)$ and $a|(bu - cv)$ for all $u, v \in R$*

Proof. (a) Let $a, b, c \in R$ such that $a|b$ and $b|c$. Then by definition of divide there exist r and s in R with

$$(*) \qquad b = ar \qquad \text{and} \qquad c = bs.$$

Hence

$$c \overset{(*)}{=} bs \overset{(*)}{=} (ar)s \overset{\text{Ax } 2}{=} a(rs).$$

Since R is closed under multiplication, $rs \in R$ and so $a|c$ by definition of divide.

(b) We will first show

$$(**) \qquad a|b \qquad \Longrightarrow \qquad a|(-b) \text{ and } (-a)|b$$

Suppose that a divides b. Then by definition of "divide" there exists $r \in R$ with $b = ar$. Thus

$$-b = -(ar) \overset{3.2.11(d)}{=} a(-r) \quad \text{and} \quad b = ar \overset{3.2.11(i)}{=} (-a)(-r)$$

By Ax 5, $-r \in R$ and so $a|(-b)$ and $(-a)|b$ by definition of "divide". So $(**)$ holds.

Suppose $a|b$. Then by $(**)$ $a|(-b)$.

Suppose that $a|(-b)$, then by $(**)$ applied with $-b$ in place of b, $(-a)|(-b)$.

Suppose that $(-a)|(-b)$. Then by $(**)$ applied with $-a$ and $-b$ in place of a and b, $(-a)|-(-b)$. By 3.2.11(e), $-(-b) = b$ and so $-a|b$.

Suppose that $(-a)|b$. Then by $(**)$ applied with $-a$ in place of a, $-(-a)|b$. By 3.2.11(e), $-(-a) = a$ and so $a|b$.

(c) Suppose that $a|b$ and $a|c$. Then by definition of divide there exist r and s in R with

$$(***) \qquad b = ar \quad \text{and} \quad c = as$$

Thus

$$b + c \overset{(***)}{=} ar + as \overset{\text{Ax } 8}{=} a(r + s) \qquad \text{and} \qquad b - c \overset{(***)}{=} ar - as \overset{3.2.11(j)}{=} a(r - s)$$

By Ax 1 and Ax 5, R is closed under addition and subtraction. Thus $r + s \in R$ and $r - s \in R$ and so $a|b + c$ and $a|b - c$.

(c) Suppose that $a|b$ and $a|c$ and let $u, v \in R$. By definition, $b \mid bu$ and $c \mid cv$ and so by (a) $a|bu$ and $a|cv$. Thus by (c), $a|(bu + cv)$ and $a|(bu - cv)$. □

Definition 3.4.4. *Let R be an commutative ring with identity and let $a, b \in R$. We say that a is* associated *to b, or that b is an* associate *of a and write $a \sim b$ if there exists a unit u in R with $au = b$.*

Lemma 3.4.5. *Let n be a non-zero integer and $a \in \mathbb{Z}$. Then $\gcd(a, n) = 1$ if and only if $[a]_n$ is a unit in $\mathbb{Z}_n$.*

Proof. Recall first from 2.2.4(10) that $[1]_n$ is the identity in $\mathbb{Z}_n$.

$\Longrightarrow$: Suppose that $\gcd(a, n) = 1$. Then by Exercise 8 on Homework 4, $[a_n][u]_n = [1]_n$ for some $u \in \mathbb{Z}$. Since $\mathbb{Z}_n$ is commutative this gives $[u]_n[a]_n = [1]_n$ and so $[a]_n$ is a unit.

$\Longleftarrow$: Suppose next that $[a]_n$ is a unit. Then the definition of a unit shows that there exists U in $\mathbb{Z}_n$ with $[a]_nU = [1]_n$. Then $U = [u]_n$ for some $u \in \mathbb{Z}$ and so

$$[au]_n = [a]_n[u]_n = [a]_nU = [1]_n$$

Put $d = \gcd(a, n)$. Then $d \mid a$ and $d \mid n$ and Exercise 9 on Homework 8 shows that $d \mid 1$. Thus $d = 1$ and $\gcd(a, n) = 1$.

□

Example 3.4.6. (a) Let $n \in \mathbb{Z}$. Find all associates of n in $\mathbb{Z}$.

(b) Find all associates of $0, 1, 2$ and 5 in $\mathbb{Z}_{10}$.

(a) By 3.2.14 the units in $\mathbb{Z}$ are ± 1. So the associates of n are $n \cdot \pm 1$, that is $\pm n$.

(b) By 2.1.2 $\mathbb{Z}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and so $\mathbb{Z}_{10} = \{0, \pm 1, \pm 2, \pm 3, \pm 4, 5\}$.
We compute

n	0	± 1	± 2	± 3	± 4	5
$\gcd(n, 10)$	10	1	2	1	2	5

and so by 3.4.5 the units in $\mathbb{Z}_{10}$ are ± 1 and ± 3.

So the associates of $a \in \mathbb{Z}_{10}$ are $a \cdot \pm 1$ and $a \cdot \pm 3$, that is $\pm a$ and $\pm 3a$. We compute

a	associates of a	associates of a, simplified
0	$\pm 0, \pm 3 \cdot 0$	0
1	$\pm 1, \pm 3 \cdot 1$	$\pm 1, \pm 3$
2	$\pm 2, \pm 3 \cdot 2$	$\pm 2, \pm 4$
5	$\pm 5, \pm 3 \cdot 5$	5

Lemma 3.4.7. *Let R be a commutative ring with identity. Then the relation $\sim$ ('is associated to') is an equivalence relation on R.*

Proof. **Reflexive:** Let $a \in R$. By (Ax 10), $1_R = 1_R 1_R$. Hence 1_R is a unit in R. By (Ax 10) $a 1_R = a$ and so $a \sim a$. Thus $\sim$ is reflexive.

Symmetric: Let $a, b \in R$ with $a \sim b$. Then there exists a unit $u \in R$ with $au = b$. Since u is a unit, u has an inverse u^{-1}. Hence (multiplying $au = b$ with u^{-1})

$$bu^{-1} = (au)u^{-1} \overset{\text{Ax } 2}{=} a(uu^{-1}) \overset{\text{def } u^{-1}}{=} a1_R \overset{(\text{Ax } 10)}{=} a.$$

By 3.2.17 u^{-1} is a unit in R and so $b \sim a$. Thus $\sim$ is symmetric.

Transitive: Let $a, b, c \in R$ with $a \sim b$ and $b \sim c$. Then $au = b$ and $bv = c$ for some units u and $v \in R$. Substituting the first equation in the second gives $(au)v = c$ and so by Ax 2, $a(uv) = c$. By 3.2.17 uv is a unit in R and so $a \sim c$. Thus $\sim$ is transitive.

Since $\sim$ is reflexive, symmetric and transitive, $\sim$ is an equivalence relation. $\square$

Example 3.4.8. Determine the equivalence classes of $\sim$ on $\mathbb{Z}_{10}$.

Note that for $a \in \mathbb{Z}_{10}$, $[a]_\sim = \{b \in \mathbb{Z}_{10} \mid a \sim b\}$ is the set of associates of a. So by Example 3.4.6

$$\begin{aligned}
[0]_\sim &= \{0\} \\
[1]_\sim &= \{\pm 1, \pm 3\} \\
[2]_\sim &= \{\pm 2, \pm 4\} \\
[5]_\sim &= \{5\}
\end{aligned}$$

By 2.1.2 $\mathbb{Z}_{10} = \{0, 1, \ldots, 9\} = \{0, \pm 1, \pm 2, \pm 3, \pm 4, 5\}$. So for each $x \in \mathbb{Z}_{10}$ there exists $y \in \{0, 1, 2, 5\}$ with $x \in [y]_\sim$. Thus by 0.5.8 $[x]_\sim = [y]_\sim$. So $[0]_\sim, [1]_\sim, [2]_\sim, [5]_\sim$ are all the equivalence classes of $\sim$.

Lemma 3.4.9. *Let R be a commutative ring with identity and $a, b \in R$ with $a \sim b$. Then $a|b$ and $b|a$.*

Proof. Since $a \sim b$, $au = b$ for some unit $u \in R$. So $a|b$.

By 3.4.7 the relation $\sim$ is symmetric and so $a \sim b$ implies $b \sim a$. Thus, by the result of the previous paragraph applied with a and b interchanged, $b|a$. □

Lemma 3.4.10. *Let R be a commutative ring with identity and $r \in R$. Then the following four statements are equivalent:*

(a) $1_R \sim r$.

(b) $r|1_R$

(c) *There exists s in R with $rs = 1_R$.*

(d) *r is a unit.*

Proof. (a) $\Longrightarrow$ (b): Since $1_R \sim r$, 3.4.9 gives $r|1_R$.

(b) $\Longrightarrow$ (c): Follows from the definition of 'divide'.

(c) $\Longrightarrow$ (d): Since R is commutative $rs = 1_R$ implies $sr = 1_R$. So r is a unit.

(d) $\Longrightarrow$ (a): By (Ax 10), $1_R r = r$. Since r is a unit this gives $1_R \sim r$ by definition of $\sim$. □

Lemma 3.4.11. *Let R be a commutative ring with identity and $a, b, c, d \in R$.*

(a) *If $a \sim b$ and $c \sim d$, then $a|c$ if and only if $b|d$.*

(b) *If $c \sim d$, then $a|c$ if and only if $a|d$.*

(c) *If $a \sim b$, then $a|c$ if and only if $b|c$.*

Proof. (a) Suppose that $a \sim b$ and $c \sim d$.

$\Longrightarrow$: Suppose that $a|c$. Since $a \sim b$, 3.4.9 gives $b|a$. Since $a|c$ and $|$ is transitive (3.4.3(a)) we have $b|c$. Since $c \sim d$, 3.4.9 gives $c|d$. Hence by transitivity of $|$, $b|d$.

$\Longleftarrow$: Since $\sim$ is symmetric, $b \sim a$ and $d \sim c$. So the result of previous paragraph applied with a and b interchanged and c and d interchanged shows that $b|d$ implies $a|c$.

(b) Since $\sim$ is reflexive, $a \sim a$. Hence (b) follows from (a) applied with $b = a$.

(c) Since $\sim$ is reflexive, $c \sim c$. Hence (c) follows from (a) applied with $c = d$. □

Definition 3.4.12. *Let R be a commutative ring and $a, b \in R$. We say that a and b divide each other in R and write $a \approx b$ if*

$$a|b \qquad \text{and} \qquad b|a.$$

Exercises 3.4:

#1. Let $R = \mathbb{Z}_{12}$.

(a) Find all units in R.

(b) Determine the equivalence classes of the relation $\sim$ on R.

#2. Let R be a commutative ring with identity. Prove that:

(a) $\approx$ is an equivalence relation on R.

(b) Let $a, b, c, d \in R$ with $a \approx b$ and $c \approx d$. Then $a|c$ if and only if $b|d$.

#3. Let n be a positive integer and $a, b \in \mathbb{Z}$. Put $d = \gcd(a, n)$ and $e = \gcd(b, n)$. Prove that:

(a) $[a]_n \big| [d]_n$ in $\mathbb{Z}_n$.

(b) $[a]_n \approx [d]_n$.

(c) Let $r, s \in \mathbb{Z}$ with $r|n$ in $\mathbb{Z}$. Then $[r]_n \big| [s]_n$ in $\mathbb{Z}_n$ if and only if $r|s$ in $\mathbb{Z}$.

(d) $[d]_n \big| [e]_n$ in $\mathbb{Z}_n$ if and only if $d|e$ in $\mathbb{Z}$.

(e) $[a]_n \big| [b]_n$ in $\mathbb{Z}_n$ if and only if $d|e$ in $\mathbb{Z}$.

(f) $[d]_n \approx [e]_n$ if and only if $d = e$.

(g) $[a]_n \approx [b]_n$ if and only if $d = e$.

#4. Let R be an integral domain and $a, b, c \in R$ such that $a \neq 0_F$ and $ba|ca$. Then $b|c$.

3.5 The General Associative Commutative and Distributive Laws in Rings

Definition 3.5.1. *Let R be a ring, n a positive integer and $a_1, a_2, \ldots a_n \in R$.*

(a) *For $k \in \mathbb{Z}$ with $1 \leq k \leq n$ define $\sum_{i=1}^{k} a_i$ inductively by*

(i) $\sum_{i=1}^{1} a_i = a_1$*; and*

(ii) $\sum_{i=1}^{k+1} a_i = \left(\sum_{i=1}^{k} a_i\right) + a_{k+1}$.

so $\sum_{i=1}^{n} a_i = \bigg(\Big(\ldots\big((a_1 + a_2) + a_3\big) + \ldots + a_{n-2}\Big) + a_{n-1}\bigg) + a_n.$

(b) *Inductively, we say that z is a sum of $(a_1, \ldots, a_n)$ in R provided that one of the following holds:*

(1) *$n = 1$ and $z = a_1$.*

(2) *$n > 1$ and there exist an integer k with $1 \le k < n$ and $x, y \in R$ such that x is a sum of $(a_1, \ldots, a_k)$ in R, y is a sum of $(a_{k+1}, a_{k+2}, \ldots, a_n)$ in R and $z = x + y$.*

(c) *$\prod_{i=1}^{k} a_n$ is defined similarly as in (a), just replace '$\sum$' by '$\prod$' and '$+$' by '$\cdot$'.*

(d) *A product of $(a_1, \ldots, a_n)$ in R is defined similarly as in (b), just replace 'sum' by 'product' and '$+$' by '$\cdot$'.*

(e) *Let $a \in R$. Then $a^n := \prod_{i=1}^{n} a \ \left(= \underbrace{aa \ldots a}_{n-times} \right)$.*

(f) *If R has an identity and $a \in R$, then $a^0 = 1_R$.*

We will also write $a_1 + a_2 + \ldots + a_n$ for $\sum_{i=1}^{n} a_n$ and $a_1 a_2 \ldots a_n$ for $\prod_{i=1}^{n} a_i$,

Example 3.5.2. Let R be a ring and $a, b, c, d \in R$. Find all sums of (a, b, c, d).

a is the only sum of (a).
$a + b$ is the only sum of (a, b).
$a + (b + c)$ and $(a + b) + c$ are the sums of (a, b, c).
$a + (b + (c + d)), a + ((b + c) + d), (a + b) + (c + d), (a + (b + c)) + d$ and $((a + b) + c) + d$ are the sums of (a, b, c, d).

Theorem 3.5.3 (General Associative Law, GAL)**.** *Let R be a ring and $a_1, a_2, \ldots, a_n$ elements of R. Then any sum of $(a_1, a_2, \ldots, a_n)$ in R is equal to $\sum_{i=1}^{n} a_i$ and any product of $(a_1, a_2, \ldots, a_n)$ is equal to $\prod_{i=1}^{n} a_i$*

Proof. See D.1.3 □

Theorem 3.5.4 (General Commutative Law,GCL)**.** *Let R be a ring, $a_1, a_2, \ldots, a_n \in R$ and*

$$f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$$

a 1-1 and onto function.

(a) *$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{f(i)}$.*

(b) *If R is commutative, then $\prod_{i=1}^{n} a_i = \prod_{i=1}^{n} a_{f(i)}$.*

Proof. See D.2.2 □

Theorem 3.5.5 (General Distributive Law,GDL)**.** *Let R be a ring and $a_1, \ldots, a_n, b_1, \ldots, b_m \in R$. Then*

$$\left(\sum_{i=1}^{n} a_i \right) \cdot \left(\sum_{j=1}^{m} b_j \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_i b_j \right)$$

Proof. See D.3.2. ☐

Example 3.5.6. Let R be a ring and a, b, c, d, e in R. By the General Associative Law:

$$a+b+c+d = \big(a+(b+c)\big)+d = (a+b)+(c+d) = a+\big((b+c)+d\big) = a+\big(b+(c+d)\big).$$

By the General Commutative Law:

$$a+b+c+d+e = d+c+a+b+e = b+a+c+d+e.$$

By General Distributive Law:

$$(a+b+c)(d+e) = (ad+ae)+(bd+be)+(cd+ce).$$

Chapter 4

Polynomial Rings

4.1 Addition and Multiplication

Definition 4.1.1. *Let R and P be a rings with identity and $x \in P$. Then P is called a* polynomial ring *with coefficients in R and indeterminate x provided that*

(i) *R is subring of P.*

(ii) *$ax = xa$ for all $a \in R$.*

(iii) *For each $f \in P$, there exists $n \in \mathbb{N}$ and $f_0, f_1, \ldots, f_n \in R$ such that*

$$f = \sum_{i=0}^{n} f_i x^i \ \left(= f_0 + f_1 x + \ldots + f_n x^n\right).$$

(iv) *Whenever $n, m \in \mathbb{N}$ with $n \leq m$ and $f_0, f_1, \ldots, f_n, g_0, \ldots, g_m \in R$ with*

$$\sum_{i=0}^{n} f_i x^i = \sum_{i=0}^{m} g_i x^i,$$

then $f_i = g_i$ for all $0 \leq i \leq n$ and $g_i = 0_R$ for all $n < i \leq m$.

Lemma 4.1.2. *Let R be ring with identity and $a, b \in R$.*

(a) *$a^{n+m} = a^n a^m$ for all $n, m \in \mathbb{N}$.*

(b) *If $ab = ab$, then $ab^n = b^n a$.*

Proof. (a) If $n = 0$, then $a^{n+m} = a^m = 1_R a^m = a^0 a^m$. So we may assume that $n > 0$. Similarly we may assume that $m > 0$. Then

$$a^n a^m = (\underbrace{aa\ldots a}_{n-\text{times}})(\underbrace{aa\ldots a}_{m-\text{times}}) \overset{\text{GAL}}{=} \underbrace{aa\ldots a}_{n+m-\text{times}} = a^{n+m}$$

(b) For $n = 0$ we have $ab^0 = a1_R = a = 1_R a = b^0 a$. Thus (b) holds. Suppose (b) holds for $n = k$. Then

$$ab^{k+1} = a(b^k b) = (ab^k)b = (b^k a)b = b^k(ab) = b^k(ba) = (b^k b)a = b^{k+1}a.$$

Thus (b) also holds for $n = k + 1$. So by the Principal Of Induction, (b) holds for all $n \in \mathbb{N}$. □

Lemma 4.1.3. *Let R be a ring with identity and P a polynomial ring with coefficients in R and indeterminate x. Then $1_R = 1_P$. In particular, $x = 1_R x$.*

Proof. Let $f \in P$. Then by definition of a polynomial ring there exists $n \in \mathbb{N}$ and $f_0, f_1, \ldots f_n \in R$ with

$$(*) \qquad f = \sum_{i=0}^{n} f_i x^i.$$

Let $1 \leq i \leq n$. By definition of a polynomial ring $1_R x = x 1_R$ and so by 4.1.2(b)

$$(**) \qquad 1_R x^i = x^i 1_R.$$

Thus

$$(***) \qquad (f_i x^i)1_R \overset{\text{Ax 7}}{=} f_i(x^i 1_R) \overset{(**)}{=} (f_i 1_R)x^i \overset{\text{Ax 10}}{=} f_i x^i$$

and

$$f1_R \overset{(*)}{=} \left(\sum_{i=0}^{n} f_i x^i\right) 1_R \overset{\text{GDL}}{=} \sum_{i=0}^{n} (f_i x^i)1_R \overset{(***)}{=} \sum_{i=0}^{n} f_i x^i \overset{(*)}{=} f$$

Similarly $1_R f = f$ and so 1_R is a multiplicative identity of P. Thus $1_R = 1_P$. Since $x \in P$ this gives $1_R x = 1_P x = x$. □

Theorem 4.1.4. *Let P be a ring with identity, R a subring of P, $x \in P$ and $f, g \in P$. Suppose that*

(i) *$ax = xa$ for all $a \in R$;*

(ii) *there exist $n \in \mathbb{N}$ and $f_0, \ldots, f_n \in R$ with $f = \sum_{i=0}^{n} f_i x^i$; and*

(iii) *there exist $m \in \mathbb{N}$ and $g_0, \ldots, g_m \in R$ with $g = \sum_{i=0}^{m} g_i x^i$.*

Put $f_i = 0_R$ for $i > n$ and $g_i = 0_R$ for $i > m$. Then

(a) $f + g = \displaystyle\sum_{i=0}^{\max(n,m)} (f_i + g_i)x^i$.

$$\text{(b)}\quad fg = \sum_{i=0}^{n}\left(\sum_{j=0}^{m} f_i g_j x^{i+j}\right) = \sum_{k=0}^{n+m}\left(\sum_{i=\max(0,k-m)}^{\min(n,k)} f_i g_{k-i}\right)x^k = \sum_{k=0}^{n+m}\left(\sum_{i=0}^{k} f_i g_{k-i}\right)x^k.$$

Proof. (a) Put $p=\max(n,m)$. Then $f_i = 0_R$ for all $n<i\leq p$ and $g_i = 0_R$ for all $m < i \leq p$. Hence

$$(*)\qquad f=\sum_{i=0}^{p} f_i x^i \qquad\text{and}\qquad g=\sum_{i=0}^{p} g_i x^i.$$

Thus

$$\begin{aligned} f+g &= \left(\textstyle\sum_{i=0}^{p} f_i x^i\right)+\left(\textstyle\sum_{i=0}^{p} g_i x^i\right) && -\ (*)\\ &= \textstyle\sum_{i=0}^{p}(f_i x^i + g_i x^i) && -\text{GCL and GAL}\\ &= \textstyle\sum_{i=0}^{p}(f_i+g_i)x^i && -\ \text{Ax 8} \end{aligned}$$

So (a) holds.

(b) By assumption $ax = xa$ and so by 4.1.2(b)

$$(**)\qquad ax^n = x^n a$$

for all $a\in R$ and $n\in\mathbb{N}$. We now can compute fg.

$$(+)\qquad \begin{aligned} fg &= \left(\sum_{i=0}^{n} f_i x^i\right)\cdot\left(\sum_{j=0}^{m} g_j x^j\right) && -\ \text{(ii) and (iii)}\\ &= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i x^i)(g_j x^j)\right) && -\text{GDL}\\ &= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i(x^i g_j))x^j\right) && -\text{GAL}\\ &= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i(g_j x^i))x^j\right) && -\ (**)\\ &= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i g_j)(x^i x^j)\right) && -\ \text{GAL}\\ &= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i g_j)x^{i+j}\right) && -\ 4.1.2(a) \end{aligned}$$

Let $k=i+j$ for some $0\leq i\leq n$ and $0\leq j\leq m$. Then

$$0\leq k\leq n+m,\quad i\leq k,\quad k-i=j\leq m,\quad k-m\leq i$$

and so

$$0 \le k \le n+m \qquad \text{and} \qquad \max(0, k-m) \le i \le \min(k, n).$$

Using the substitution $k = i + j$ (and so $j = k - i$) and the GCL and GAL we therefore conclude that

$$(++) \qquad \begin{aligned} \sum_{i=0}^{n}\left(\sum_{j=0}^{m} f_i g_j x^{i+j}\right) &= \sum_{k=0}^{n+m}\left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i} x^k\right) \\ &= \sum_{k=0}^{n+m}\left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i}\right) x^k \quad \text{–GDL} \end{aligned}$$

If $0 \le i < \max(0, k-m)$, then $k - i > m$ and so $g_{k-i} = 0_R$. Hence $f_i g_{k-i} = f_i 0_R = 0_R$ (by 3.2.11(c)).

If $\min(k, n) < i \le k$ for some $i \in \mathbb{N}$, then $\min(n, k) \neq k$ and so $\min(n, k) = n$ and $n < i$. Thus $f_i = 0_R$ and so $f_i g_{k-i} = 0_R g_{k-i} = 0_R$. It follows that

$$\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i} = \sum_{i=0}^{k} f_i g_{k-i}$$

and so also

$$(+++) \qquad \sum_{k=0}^{n+m}\left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i}\right) x^k = \sum_{k=0}^{n+m}\left(\sum_{i=0}^{k} f_i g_{k-i}\right) x^k.$$

Combining (+), (++) and (+++) gives (b).

□

Example 4.1.5. (1) Suppose that $R = \mathbb{Z}_2$, $f = 1 + x + x^3$ and $g = 1 + x^2 + x^3 + x^5$. Compute $f + g$.

$$\begin{aligned} f + g &= (1 + x + x^3) + (1 + x^2 + x^3 + x^5) \\ &= (1+1) + (1+0)x + (0+1)x^2 + (1+1)x^3 + (0+0)x^4 + (0+1)x^5 \\ &= 0 + 1x + 1x^2 + 0x^3 + 0x^4 + 1x^5 \\ &= x + x^2 + x^5 \end{aligned}$$

(2) Suppose that $R = \mathbb{Z}_6$, $f = 1 + x + x^2$ and $g = 1 + x + 2x^2 + 3x^3$. Compute fg.

$$\begin{aligned} fg &= (1+x+2x^2)(1+x+2x^2+3x^3) \\ &= (1\cdot 1) + (1\cdot 1 + 1\cdot 1)x + (1\cdot 2 + 1\cdot 1 + 2\cdot 1)x^2 \\ &\qquad + (1\cdot 3 + 1\cdot 2 + 2\cdot 1)x^3 + (1\cdot 3 + 2\cdot 2)x^4 + (2\cdot 3)x^5 \\ &= 1 + 2x + 5x^2 + x^3 + x^4 \end{aligned}$$

Definition 4.1.6. *Let R be a ring with identity.*

(a) *$R[x]$ denotes the polynomial ring with coefficients in R and indeterminate x constructed in F.3.1.*

(b) *Let $f \in R[x]$ and let $n \in \mathbb{N}$ and $a_0, a_1, \dots a_n \in R$ with $f = \sum_{i=0}^n a_i x^i$. Let $i \in \mathbb{N}$. If $i \leq n$ define $f_i = a_i$. If $i > n$ define $f_i = 0_R$. Then f_i is called the* coefficient *of x^i in f.(Observe that this is well defined by 4.1.1)*

(c) *$\mathbb{N}^* := \mathbb{N} \cup \{-\infty\}$. For $n \in \mathbb{N}^*$ we define $n + (-\infty) = -\infty$ and $-\infty + n = -\infty$. We extend the relation $'\leq'$ on $\mathbb{N}$ to $\mathbb{N}^*$ by declaring that $-\infty \leq n$ for all $n \in \mathbb{N}^*$.*

(d) *For $f \in R[x]$, $\deg f$ is the minimal element of $\mathbb{N}^*$ with $f_i = 0_R$ for all $i \in \mathbb{N}$ with $i > \deg f$. So $\deg 0_R = -\infty$ and if $f = \sum_{i=0}^n f_i x^i$ with $f_n \neq 0$, then $\deg f = n$.*

(e) *If $\deg f \in \mathbb{N}$ then $\operatorname{lead}(f)$ is the coefficient of $x^{\deg f}$ in f. If $\deg f = -\infty$, then $\operatorname{lead}(f) = 0_R$.*

Lemma 4.1.7. *Let R be a ring with identity and $f \in R[x]$.*

(a) *$f = 0_R$ if and only if $\deg f = -\infty$ and if and only if $\operatorname{lead}(f) = 0_R$.*

(b) *$\deg f = 0$ if and only if $f \in R$ and $f \neq 0_R$.*

(c) *$f \in R$ if and only if $\deg f \leq 0$ and if and only if $f = \operatorname{lead}(f)$.*

(d) *$f = \sum_{i=0}^{\deg f} f_i x^i$. (Here an empty sum is defined to be 0_R)*

Proof. This follows straightforward from the definition of $\deg f$ and $\operatorname{lead} f$ and we leave the details to the reader. □

Lemma 4.1.8. *Let R be a ring with identity and $f, g \in R[x]$. Then*

(a) *$\deg(f+g) \leq \max(\deg f, \deg g)$.*

(b) *$\deg(-f) = \deg f$.*

(c) *Exactly one of the following holds.*

   (1) *$\deg(fg) = \deg f + \deg g$ and $\operatorname{lead}(fg) = \operatorname{lead}(f)\operatorname{lead}(g)$.*

   (2) *$\deg(fg) < \deg f + \deg g$, $\operatorname{lead}(f)\operatorname{lead}(g) = 0_R$, $f \neq 0_R$ and $g \neq 0_R$.*

In particular, $\deg fg \leq \deg f + \deg g$.

Proof. Put $n := \deg f$ and $m := \deg g$. Then $f = \sum_{i=0}^{n} f_i x^n$ and $g = \sum_{i=0}^{m} g_i x^i$.

(a) By 4.1.4(a), $f + g = \sum_{i=0}^{\max(n,m)} (f_i + g_i)x^i$ and so $(f+g)_k = 0_R$ for $k > \max(\deg f, \deg g)$. Thus (a) holds.

(b) Note that $-f = \sum_{i=0}^{n}(-f_i)x^i$. As $f_n \neq 0_R$ we also have $-f_n \neq 0_R$ and so $\deg(-f) = \deg f$.

(c) Suppose first that $f = 0_R$. Then $fg = 0_R g = 0_R$. Hence $\deg f = -\infty$, $\deg(fg) = -\infty$, $\mathrm{lead} f = 0_R$ and $\mathrm{lead}(fg) = 0_R$. Hence

$$\deg(fg) = -\infty = -\infty + \deg g = \deg f + \deg g \text{ and } \mathrm{lead}(fg) = 0_R = 0_R \cdot \mathrm{lead}(g) = \mathrm{lead}(f)\mathrm{lead}(g)$$

So (c:1) holds in this case. Similarly, (c:1) holds if $g = 0_R$.

So suppose $f \neq 0_R \neq g$ By 4.1.4(b),

$$fg = \sum_{k=0}^{n+m} \left(\sum_{i=\min(0,k-m)}^{\max(k,n)} f_i g_{k-i} \right) x^k.$$

Thus $(fg)_k = 0_R$ for $k > n + m$ and so $\deg fg \leq n + m$. Moreover, for $k = n + m$ we have $\max(0, k-m) = \max(0, n) = n$ and $\min(n, k) = \min(n, n+m) = n$. So

$$(fg)_{n+m} = \sum_{i=n}^{n} f_i g_{n+m-i} = f_n g_m = \mathrm{lead}(f)\mathrm{lead}(g).$$

Suppose that $\mathrm{lead}(f)\mathrm{lead}(g) \neq 0_R$. Then $\deg(f+g) = n+m$ and $\mathrm{lead}(fg) = \mathrm{lead}(f)\mathrm{lead}(g)$. Thus (c:1) holds.

Suppose that $\mathrm{lead}(f)\mathrm{lead}(g) = 0_R$. Then $\deg(f+g) < n+m$ and (c:2) holds. □

Theorem 4.1.9. *Let R be a commutative ring with identity. Then $R[x]$ is commutative.*

Proof. Let $f, g \in R[x]$. Then

$$\begin{aligned}
fg &= \left(\sum_{i=0}^{n} f_i x^i\right)\left(\sum_{j=0}^{m} g_j x^j\right) && \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m} f_i g_j x^{i+j} && - \text{Theorem 4.1.4} \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m} g_j f_i x^{j+i} && - \text{R commutative} \\
&= \sum_{j=0}^{m}\sum_{i=0}^{n} g_j f_i x^{j+i} && - \text{GCL, GAL} \\
&= \left(\sum_{j=0}^{m} g_j x^j\right)\left(\sum_{i=0}^{n} f_i x^i\right) && - \text{Theorem 4.1.4} \\
&= gf &&
\end{aligned}$$

We proved that $fg = gf$ for all $f, g \in R[x]$ and so $R[x]$ is commutative. $\square$

Theorem 4.1.10. *Let R be field or an integral domain. Then*

(a) $\deg(fg) = \deg f + \deg g$ *and* $\mathrm{lead}(fg) = \mathrm{lead}(f)\mathrm{lead}(g)$ *for all* $f, g \in R[x]$.

(b) $\deg(rf) = \deg f$ *and* $\mathrm{lead}(rf) = r\,\mathrm{lead}(f)$ *for all* $r \in R$ *and* $f \in R[x]$ *with* $r \neq 0_R$.

(c) $R[x]$ *is an integral domain.*

Proof. By Theorem 3.2.22 any field is an integral domain. So in any case R is an integral domain. We will first show that

(*) If $f, g \in R$ with $\mathrm{lead}(f)\mathrm{lead}(g) = 0_R$ then $f = 0_R$ or $g = 0_R$.

Indeed since R is an integral domain, $\mathrm{lead}(f)\mathrm{lead}(g) = 0_R$ implies $\mathrm{lead}(f) = 0$ or $\mathrm{lead}(g) = 0_R$. 4.1.7 now implies $f = 0_R$ or $g = 0_R$.

(a) By 4.1.8(c)

(1) $\deg(fg) = \deg f + \deg g$ and $\mathrm{lead}(fg) = \mathrm{lead}(f)\mathrm{lead}(g)$, or

(2) $\deg(fg) < \deg f + \deg g$, $\mathrm{lead}(f)\mathrm{lead}(g) = 0_R$, $f \neq 0_R$ and $g \neq 0_R$.

In the first case (a) holds. The second case contradicts (*) and so does not occur.

(b) By 4.1.7 $\deg r = 0$ and $\mathrm{lead} r = r$. So (b) follows from (a).

(c) By 4.1.9, $R[x]$ is a commutative ring with identity 1_R. Note that $1_{R[x]} = 1_R \neq 0_R = 0_{R[x]}$. Let $fg \in R[x]$ with $fg = 0_R$. Then by (a) $\mathrm{lead}(f)\mathrm{lead}(g) = \mathrm{lead}(fg) = \mathrm{lead}(0_R) = 0_R$ and by (*), $f = 0_R$ or $g = 0_R$. Hence $R[x]$ is an integral domain. $\square$

Theorem 4.1.11 (Division Algorithm)**.** *Let F be a field and $f, g \in F[x]$ with $g \neq 0_F$. Then there exist uniquely determined $q, r \in F[x]$ with*

$$f = gq + r \quad \textit{and} \quad \deg r < \deg g.$$

Proof. Fix $g \in F[x]$ with $g \neq 0_F$. For $n \in \mathbb{N}$ let $P(n)$ be the statement:

$P(n):$ If $f \in R[x]$ with $\deg f \leq n$, then there exists $q, r \in F[x]$ with $f = gq+r$ and $\deg r < \deg g$.

Let $k \in \mathbb{N}$ such that $P(n)$ holds for all $n \in \mathbb{N}$ with $n < k$. We will show that $P(k)$ holds. So let $f \in F[x]$ with $\deg f \leq k$. Put $m = \deg g$. Note that $f = g \cdot 0_R + f$. So if $k < m$, then $P(k)$ holds for f with $q = 0_R$ and $r = f$.

So we may assume that $k \geq m$. Since $g \neq 0_R$ we have $m = \deg g \in \mathbb{N}$ and $g_m \neq 0_F$. As F is a field this implies that g_m is a unit in F. Define

$$\tilde{f} := f - g \cdot g_m^{-1} f_k x^{k-m}. \tag{1}$$

Since $-g$ has degree m and $g_m^{-1} f_k x^{k-m}$ has degree $k-m$, 4.1.8(c) shows that $-g \cdot g_m^{-1} f_k x^{k-m}$ has degree at most $m + (k-m) = k$. Since f has degree at most k we conclude from 4.1.8(a) that

$$\deg \tilde{f} = \deg(f - g \cdot g_m^{-1} f_k x^{k-m}) \leq \max\Big(\deg f, \deg(-g \cdot g_m^{-1} f_k x^{k-m})\Big) \leq k.$$

The coefficient of x^k in $\tilde{f}$ is $f_k - g_m g_m^{-1} f_k = f_k - f_k = 0_F$. Thus $\deg \tilde{f} \neq k$ and so $\deg \tilde{f} \leq k-1$. By the induction assumption, $P(k-1)$-holds and so there exist $\tilde{q}$ and $\tilde{r} \in F[x]$ with

$$\tilde{f} = g\tilde{q} + \tilde{r} \quad \text{and} \quad \deg \tilde{r} < \deg g. \tag{2}$$

We compute

$$\begin{aligned} f &= \tilde{f} + g \cdot f_k g_m^{-1} x^{k-m} && - (1) \\ &= (g\tilde{q} + \tilde{r}) + g \cdot g_m^{-1} f_k x^{k-m} && - (2) \\ &= (g\tilde{q} + g \cdot g_m^{-1} f_k x^{k-m}) + \tilde{r} && - \text{Ax } 2, \text{Ax } 3 \\ &= g \cdot (\tilde{q} + g_m^{-1} f_k x^{k-m}) + \tilde{r} && - \text{Ax } 8 \end{aligned} \tag{3}$$

Put $q = \tilde{q} + g_m^{-1} f_k x^{k-m}$ and $r = \tilde{r}$. Then by (3), $f = qg + r$ and by (2), $\deg r = \deg \tilde{r} < \deg g$. Thus $P(k)$ is proved.

By the Principal of Complete Induction 0.4.4 we conclude that $P(n)$ holds for all $n \in \mathbb{N}$. This shows the existence of q and r.

To show uniqueness suppose that for $i = 1, 2$ we have $q_i, r_i \in F[x]$ with

$$f = gq_i + r_i \quad \text{and} \quad \deg r_i < \deg g. \tag{4}$$

Then

$$gq_1 + r_1 = gq_2 + r_2$$

and so

$$g \cdot (q_1 - q_2) = r_2 - r_1. \tag{5}$$

Suppose $q_1 - q_2 \neq 0_F$ Then $\deg(q_1 - q_2) \geq 0$ and so

$$\begin{aligned} \deg g \leq{} & \deg g + \deg(q_1 - q_2) \overset{4.1.10(a)}{=} \deg(g \cdot (q_1 - q_2)) \overset{(5)}{=} \deg(r_1 - r_2) \\ \overset{4.1.8(a)}{\leq}{} & \max(\deg r_1, \deg r_2) \overset{(4)}{<} \deg g. \end{aligned}$$

(Note here that we can apply 4.1.10(a) since F is a field.)

This contradiction shows $q_1 - q_2 = 0_F$. Hence , by (5) also $r_2 - r_1 = g \cdot (q_1 - q_2) = g \cdot 0_F = 0_F$. Thus $q_1 = q_2$ and $r_1 = r_2$, see 3.2.11(f). $\square$

Definition 4.1.12. *Let F be field and $f, g \in F[x]$ with $g \neq 0_F$. Let $q, r \in F[x]$ be the unique polynomials with*

$$f = gq + r \quad \textit{and} \quad \deg r < \deg g$$

Then r is called the remainder *of f when divided by g.*

Example 4.1.13. Consider the polynomials $f = x^4 + x^3 - x + 1$ and $g = -x^2 + x - 1$ in $\mathbb{Z}_3[x]$. Compute the remainder of f when divided by g.

$$
\begin{array}{r|rrrrrrrrr}
\multicolumn{1}{r}{} & -x^2 & + & x & - & 1 & & & & \\ \cline{2-10}
-x^2+x-1 & x^4 & + & x^3 & & & - & x & + & 1 \\
 & x^4 & - & x^3 & + & x^2 & & & & \\ \cline{2-10}
 & & - & x^3 & - & x^2 & - & x & + & 1 \\
 & & - & x^3 & + & x^2 & - & x & & \\ \cline{3-10}
 & & & & & x^2 & & & + & 1 \\
 & & & & & x^2 & - & x & + & 1 \\ \cline{6-10}
 & & & & & & & x & &
\end{array}
$$

Since $\deg x = 1 < 2 = \deg(-x^2 + x - 1)$, the remainder of $x^4 + x^3 - x + 1$ when divided by $-x^2 + x + 1$ in $\mathbb{Z}_3[x]$ is x.

Exercises 4.1:

#1. Perform the indicated operation and simplify your answer:

(a) $(3x^4 + 2x^3 - 4x^2 + x + 4) + (4x^3 + x^2 + 4x + 3)$ in $\mathbb{Z}_5[x]$.

(b) $(x + 1)^3$ in $\mathbb{Z}_3[x]$.

(c) $(x - 1)^5$ in $\mathbb{Z}_5[x]$.

(d) $(x^2 - 3x + 2)(2x^3 - 4x + 1) \in \mathbb{Z}_7[x]$.

#2. Find polynomials q and r such that $f = gq + r$ and $\deg r) < \deg g$.

(a) $f = 3x^4 - 2x^3 + 6x^2 - x + 2$ and $g = x^2 + x + 1$ in $\mathbb{Q}[x]$.

(b) $f = x^4 - 7x + 1$ and $g = 2x^2 + 1$ in $\mathbb{Q}[x]$.

(c) $f = 2x^4 + x^2 - x + 1$ and $g = 2x - 1$ in $\mathbb{Z}_5[x]$.

(d) $f = 4x^4 + 2x^3 + 6x^2 + 4x + 5$ and $g = 3x^2 + 2$ in $\mathbb{Z}_7[x]$.

#3. Let R be a commutative ring. If $a_n \neq 0_R$ and $a_0 + a_1x + \ldots + a_nx^n$ is a zero-divisor in $R[x]$, then a_n is a zero divisor in R.

#4. (a) Let R be an integral domain and $f, g \in R[x]$. Assume that the leading coefficent of g is a unit in R. Verify that the Division algorithm holds for f as divident and g as divisor.

(b) Give an example in $\mathbb{Z}[x]$ to show that part (a) may be false if the leading coefficent of $g(x)$ is not a unit.[*Hint:* Exercise 4.1.5(b).]

4.2 Divisibility in $F[x]$

In a general commutative ring it may or may not be easy to decide whether a given element divides another. But for polynomial over a field it is easy, thanks to the division algorithm:

Lemma 4.2.1. *Let F be a field and $f, g \in F[x]$ with $g \neq 0_F$. Then g divides f in $F[x]$ if and only if the remainder of f when divided by g is 0_F.*

Proof. $\Longrightarrow$: Suppose that $g|f$. Then by Definition 3.4.1 $f = gq$ for some $q \in F[x]$. Thus $f = gq+0_F$. Since $\deg 0_F = -\infty < \deg g$, Definition 4.1.12 shows that 0_F is the remainder of f when divided by g.

$\Longleftarrow$: Suppose that the remainder of f when divided by g is 0_F. Then by Definition 1.1.3 $f = gq + 0_F$ for some $q \in F[x]$. Thus $f = gq$ and so Definition 3.4.1 shows that $g|f$. $\square$

Lemma 4.2.2. *Let R be a field or an integral domain and $f, g \in R[x]$. If $g \neq 0_R$ and $f|g$, then $\deg f \leq \deg g$.*

Proof. Since $f|g$, $g = fh$ for some $h \in R[x]$. If $h = 0_R$, then by 3.2.11(c), $g = fh = f0_R = 0_R$, contrary to the assumption. Thus $h \neq 0_R$ and so $\deg h \geq 0$. Thus by 4.1.10(a),

$$\deg g = \deg fh = \deg f + \deg h \geq \deg f.$$

$\square$

Lemma 4.2.3. *Let F be a field and $f \in F[x]$. Then the following statements are equivalent:*

(a) $\deg f = 0$.

(b) $f \in F$ *and* $f \neq 0_F$.

(c) $f|1_F$.

(d) $f \sim 1_F$.

(e) f *is a unit in* $F[x]$.

Proof. (a) $\Longrightarrow$ (b): See 4.1.7(b).

(b) $\Longrightarrow$ (c): Suppose that $f \in F$ and $f \neq 0_F$. Since F is a field, f has an inverse $f^{-1} \in F$. Then $f^{-1} \in F[x]$ and $ff^{-1} = 1_F$. Thus $f|1_F$ by definition of 'divide' and (c) holds.

(c) $\Longrightarrow$ (d): and (d) $\Longrightarrow$ (e): See 3.4.10.

(e) $\Longrightarrow$ (a): Since f is a unit, $1_F = fg$ for some $g \in F[x]$. So by 4.1.10(a) $\deg f + \deg g = \deg(fg) = \deg(1_F) = 0$ and so also $\deg f = \deg g = 0$. $\square$

Lemma 4.2.4. *Let F be a field and $f, g \in F[x]$. Then the following statements are equivalent:*

(a) $f \sim g$.

(b) $f|g$ *and* $g|f$.

(c) $\deg f = \deg g$ *and* $f|g$.

(d) $g \sim f$.

Proof. (a) $\Longrightarrow$ (b): See 3.4.11.

(b) $\Longrightarrow$ (c): Suppose that $f|g$ and $g|f$. We need to show that $\deg f = \deg g$. Assume first that $g = 0_F$, then since $g|f$, we get from 3.4.2 that $f = 0_F$. Hence $f = g$ and so also $\deg g = \deg f$ and thus (c) holds. Similarly, (c) holds if $f = 0_F$.

Assume that $f \neq 0_F$ and $g \neq 0_F$. Since $f|g$ and $g|f$ we conclude from 4.2.2 that $\deg f \leq \deg g$ and $\deg g \leq \deg f$. Thus $\deg g = \deg f$ and (c) holds.

(c) $\Longrightarrow$ (d): Suppose that $\deg f = \deg g$ and $f|g$. If $f = 0_F$, then $\deg g = \deg f = -\infty$ and so $g = 0_F$. Hence $f = g$ and so $f \sim g$ since $\sim$ reflexive.

Thus we may assume $f \neq 0_F$. Since $f|g$, $g = fh$ for some $h \in F[x]$. Thus by 4.1.10(a), $\deg g = \deg f + \deg h$. Since $f \neq 0_F$ we have $\deg g = \deg f \neq -\infty$ and so $\deg h = 0$. Thus by 4.2.3, h is a unit. So $g \sim f$ by definition of $\sim$.

(d) $\Longrightarrow$ (a): This holds since $\sim$ is symmetric by 3.4.7. $\square$

Definition 4.2.5. *Let F be a field and $f \in F[x]$.*

(a) *f is called* monic *if* $\operatorname{lead}(f) = 1_F$.

(b) *If $f \neq 0_F$ then $\check{f} := \operatorname{lead}(f)^{-1} f$ is called the* monic polynomial associated *to f. If $f = 0_F$ put $\check{f} = 0_F$.*

Lemma 4.2.6. *Let F be a field and $f, g \in F[x]$.*

(a) $\check{f} \sim f$.

(b) *If f and g are monic and $f \sim g$, then $f = g$.*

(c) *If $f \neq 0_F$, then $\check{f}$ is the unique monic polynomial associated to f.*

(d) $\deg \check{f} = \deg f$.

(e) *$f \sim g$ if and only if $\check{f} = \check{g}$.*

Proof. Recall from 3.4.7 that $\sim$ is an equivalence relation and so reflexive, symmetric and transitive.

(a) Suppose that $f = 0_F$. Then $\check{f} = 0_F$ and so $f \sim \check{f}$ as $\sim$ is reflexive.

Suppose that $f \neq 0_F$. Then also $\operatorname{lead}(f) \neq 0_F$ and so by 4.2.3 $\operatorname{lead}(f)$ is a unit in $F[x]$. Also $\check{f} = \operatorname{lead}(f)^{-1} f = f \operatorname{lead}(f)^{-1}$ and so $\check{f} \sim f$.

(b) By definition of $f \sim g$ we have $fu = g$ for some unit u in $F[x]$. By 4.2.3, $0_F \neq u \in F$. Hence

$$1_F \overset{g \text{ monic}}{=} \text{lead}(g) \overset{fu=g}{=} \text{lead}(fu) \overset{4.1.10(b)}{=} \text{lead}(f)u \overset{f \text{ monic}}{=} 1_F u \overset{(\text{Ax } 10)}{=} u$$

and so $u = 1_F$ and $g = fu = f1_F = f$.

(c) Suppose $f \neq 0_F$. By 4.1.10(b) Then

$$\text{lead}(\check{f}) = \text{lead}\big(\text{lead}(f)^{-1}f\big) \overset{4.1.10(b)}{=} \text{lead}(f)^{-1}\text{lead}(f) = 1_F.$$

So $\check{f}$ is monic. By (a) we have $\check{f} \sim f$ and so $\check{f}$ is a monic polynomial associated to f.

Suppose g is a monic polynomial with $g \sim f$. Since $\sim$ is symmetric we get we get $f \sim g$. By (a) $\check{f} \sim f$. As $\sim$ is transitive this gives $\check{f} \sim g$. Since both $\check{f}$ and g are monic we conclude from (b) that $g = \check{f}$.

(d) By (a) $f \sim \check{f}$ and so by 4.2.4 $\deg f = \deg \check{f}$.

(e) By (a) $f \sim \check{f}$ and $g \sim \check{g}$. Thus by 0.5.8

$$(*) \qquad [f]_\sim = [\check{f}]_\sim \qquad \text{and} \qquad [g]_\sim = [\check{g}]_\sim.$$

Using this we get

$$\begin{array}{rlll} & f & \sim g & \\ \iff & [f]_\sim & = [g]_\sim & - \ 0.5.8 \\ \iff & [\check{f}]_\sim & = [\check{g}]_\sim & - \ (*) \\ \iff & \check{f} & \sim \check{g} & - \ 0.5.8 \end{array}$$

□

Definition 4.2.7. *Let F be a field and $f, g \in F[x]$.*

(a) *$h \in F[x]$ is called a* common divisor *of f and g provided that $h|f$ and $h|g$.*

(b) *Let $d \in F[x]$. We say that d is a* greatest common divisor *of f and g and write*

$$d = \gcd(f, g)$$

provided that

(i) *d is a common divisor of f and g,*

(ii) *If c is a common divisor of f and g, then $\deg c \leq \deg d$, and*

(iii) *d is monic.*

Lemma 4.2.8. *Let F be a field and $f, g, q, d, u \in F[x]$. Suppose that*

(i) *u is a unit in $F[x]$,*

(ii) *$f = gq + ru$, and*

(iii) *$d = \gcd(g, r)$*

Then $d = \gcd(f, g)$

Proof. By definition of a greatest common divisor, $d \mid g$ and $d \mid r$. Since $f = gq + ru$ we conclude from 3.4.3(d) that $d \mid f$. Thus d is a common divisor of f and g.

Let c be any common divisor of f and g. Since $f = gq + ru$ and u is a unit we have $r = f \cdot u^{-1} - g \cdot qu^{-1}$. Thus 3.4.3(d) implies that $d \mid r$. So c is a common divisor of g and r. As d is a greatest common divisor of g and r we conclude that $\deg c \leq \deg d$. Thus d is a greatest common divisor of f and g. $\square$

Theorem 4.2.9 (Euclidean Algorithm)**.** *Let F be a field and $f, g \in F[x]$ with $g \neq 0_F$ and let E_{-1} and E_0 be the equations*

$$\begin{array}{rcrcrcl} E_{-1} & : & f & = & f \cdot 1_F & + & g \cdot 0_F \\ E_0 & : & \check{g} & = & f \cdot 0_F & + & g \cdot \operatorname{lead}(g)^{-1} \end{array},$$

Let $i \in \mathbb{N}$ and suppose inductively we defined equations $E_k, -1 \leq k \leq i$ of the form

$$E_k \quad : \quad r_k \quad = \quad f \cdot x_k \quad + \quad g \cdot y_k \quad .$$

where $r_k, x_k, y_k \in F[x]$ and r_i is monic. According to the division algorithm, let $t_{i+1}, q_{i+1} \in F[x]$ with

$$r_{i-1} = r_i q_{i+1} + t_{i+1} \text{ and } \deg t_{i+1} < \deg r_i$$

If $t_{i+1} \neq 0_F$, put $u_{i+1} = \operatorname{lead}(t_{i+1})^{-1}$. Let E_{i+1} be equation of the form $r_{i+1} = f \cdot x_{i+1} + g \cdot y_{i+1}$ obtained by first subtracting q_{i+1}-times equation E_i from E_{i-1} and then multiplying the resulting equation by u_{i+1}. Continue the algorithm with $i + 1$ in place of i.

If $t_{i+1} = 0_F$, define $d = r_i, u = x_i$ and $v = y_i$. Then

$$d = \gcd(f, g) \quad \text{and} \quad d = fu + gv$$

and the algorithm stops.

Proof. For $i \in \mathbb{N}$ let $P(i)$ be the following statement:

(1) For $-1 \leq k \leq i$ an equation E_k of the form $r_k = f \cdot x_k + g \cdot y_k$ with r_k, x_k and $y_k \in F[x]$ has been defined;

(2) for $-1 \leq k \leq i$ the equation E_k is true;

(3) r_i is monic;

(4) for all $1 \leq k \leq i$, $\deg r_k < r_{k-1}$; and

(5) If $d \in F[x]$ with $d = \gcd(r_{i-1}, r_i)$ then $d = \gcd(f, g)$.

Put $r_{-1} = f, x_{-1} = 1_F, y_{-1} = 0_F, r_0 = \check{g}, x_0 = 0_F$ and $y_0 = \mathrm{lead}(g)^{-1}$. Then for $k = -1$ and $k = 0$, E_k is the equation $r_k = f \cdot x_k + g \cdot y_k$ and so (1) holds for $i = 0$. Also E_{-1} and E_0 are true, so (2) holds for $i = 0$. $r_0 = \check{g}$ is monic and so (3) holds for $i = 0$. There is no integer k with $1 \leq k \leq 0$ and thus (4) holds for $i = 0$. Assume $d \in F[x]$ with $d = \gcd(r_{-1}, r_0)$. Then $d = \gcd(f, \check{)}k$. Note that $g = f \cdot 0_R + \check{g} \cdot \mathrm{lead}(g)$. As $\mathrm{lead}(g)$ is a unit in $F[x]$ we conclude from 4.2.8 that $d = \gcd(f, g)$.

Thus $P(0)$ holds. Suppose now that $i \in \mathbb{N}$ and that $P(i)$ holds. Then the equations

$$\begin{array}{lllllll} E_{i-1} & : & r_{i-1} & = & f \cdot x_{i-1} & + & g \cdot y_{i-1} \quad \text{and} \\ E_i & : & r_i & = & f \cdot x_i & + & g \cdot y_i. \end{array}$$

are defined and true. Also r_k, x_k and y_k are in $F[x]$ for $k = i-1$ and i,

Since r_i is monic, $r_i \neq 0_F$ and so by the Division algorithm there exist unique q_{i+1} and t_{i+1} in $F[x]$ with

$$(*) \qquad r_{i-1} = r_i q_i + t_{i+1} \text{ and } \deg t_{i+1} < \deg r_i$$

Consider the case that $t_{i+1} \neq 0_F$. Subtracting q_{i+1} times E_i from E_{i-1} we obtain the true equation

$$r_{i-1} - r_i q_{i+1} \quad = \quad f \cdot (x_{i-1} - x_i q_{i+1}) \quad + \quad g \cdot (y_{i-1} - y_i q_{i+1}).$$

Put $u_{i+1} = (\mathrm{lead} t_{i+1})^{-1}$. Multiplying the preceding equation with u_{i+1} gives the true equation

$$E_{i+1} \quad : \quad (r_{i-1} - r_i q_{i+1}) u_{i+1} \quad = \quad f \cdot (x_{i-1} - x_i q_{i+1}) u_{i+1} \quad + \quad g \cdot (y_{i-1} - y_i q_{i+1}) u_{i+1}.$$

Putting $r_{i+1} = (r_{i-1} - r_i q_{i+1}) u_{i+1}$, $x_{i+1} = (x_{i-1} - x_i q_{i+1}) u_{i+1}$ and $y_{i+1} = (y_{i-1} - y_i q_{i+1}) u_{i+1}$ we see that E_{i+1} is the equation $r_{i+1} = f \cdot x_{i+1} + g \cdot y_{i+1}$ and r_{i+1}, x_{i+1} and y_{i+1} are in $F[x]$. So (1) and (2) hold for $i+1$ in place of i.

By (*) we have $t_{i+1} = r_{i-1} - r_i q_{i+1}$ and so

$$r_{i+1} = (r_{i-1} - r_i q_{i+1}) u_{i+1} = t_{i+1} u_{i+1} = t_{i+1} \mathrm{lead}(t_{i+1})^{-1} = \check{t}_{i+1}.$$

Hence

$$r_{i+1} = \check{t}_{i+1}$$

Thus r_{i+1} is monic and (3) holds. Moreover, $t_{i+1} = r_{i+1} \mathrm{lead}(t_{i+1})$ and (*) gives

$$r_{i-1} = r_i q_i + r_{i+1} \mathrm{lead}(t_{i+1}).$$

Hence, if $d \in F[x]$ with $d = \gcd(r_i, r_{i+1})$, we conclude from 4.2.8 that $d = \gcd(r_{i-1}, r_i)$. As $P(i)(5)$ holds, this gives $d = \gcd(f, g)$ and so (5) in $P(i+1)$ holds. We proved that $P(i)$ implies $P(i+1)$ and so by the principal of induction, $P(i)$ holds for all $i \in \mathbb{N}$, which are reached before the algorithm stops. Note here that Condition (4) ensures that the algorithm stops in finitely many steps.

Suppose next that $t_{i+1} = 0_F$. Note that by 4.2.2 any common divisor of r_i and 0_F has degree at most $\deg r_i$. Since r_i is monic common divisor of r_i and 0_F we conclude that $r_i = \gcd(r_i, 0_F)$. As $t_{i+1} = 0_F$, (*) implies that $r_{i-1} = r_i q_i + 0_F$ and so 4.2.8 shows that $r_i = \gcd(r_{-i}, r_i)$. As $P(i)(5)$ holds, this shows that $r_i = \gcd(f, g)$.

By $P(i)$ the equation

$$E_i: \quad r_i = f \cdot x_i + g \cdot y_i$$

is true. So putting $d = r_i, u = x_i$ and $v = y_i$ we have

$$d = \gcd(f, g) \quad \text{and} \quad fu + gv$$

$\square$

Example 4.2.10. Let $f = 3x^4 + 4x^3 + 2x^2 + x + 1$ and $g = 2x^3 + x^2 + 2x + 3$ in $\mathbb{Z}_5[x]$. Find $u, v \in \mathbb{Z}_2[x]$ such that $fu + gv = \gcd(f, g)$.

In the following if a in integer, we just write a for $[a]_5$. We have

$$\text{lead}(g)^{-1} = 2^{-1} = 2^{-1} \cdot 1 = 2^{-1} \cdot 6 = 3$$

and so $r_0 = \breve{g} = 3g = 6x^3 + 3x^2 + 6x + 9 = x^3 + 3x^2 + x + 4$.

$$\begin{array}{rcrcccccc} E_{-1} & : & 3x^4 + x^3 + 2x^2 + x + 1 & = & f \cdot 1 & + & g \cdot 0 \\ E_0 & : & x^3 + 3x^2 + x + 4 & = & f \cdot 0 & + & g \cdot 3 \end{array} ,$$

$$\begin{array}{r|ccccccccc} & 3x & & & & & & & & \\ \hline x^3 + 3x^2 + x + 4 & 3x^4 & + & 4x^3 & + & 2x^2 & + & x & + & 1 \\ \hline & 3x^4 & + & 9x^3 & + & 3x^2 & + & 2x & & \\ \hline & & & & & -x^2 & & -x & + & 1 \end{array}$$

Subtracting $3x$ times E_0 from E_{-1} we get

$$-x^2 - x + 1 \quad = \quad f \cdot 1 \quad + \quad g \cdot -9x \qquad | \quad E_{-1} - E_0 \cdot 3x$$

and multiplying with $(-1)^{-1} = -1$ gives

$$E_1 \quad : \quad x^2+x-1 \quad = \quad f\cdot -1 \quad + \quad g\cdot 4x$$

$$\begin{array}{r|rrrrrrr}
\multicolumn{1}{r}{} & x & + & 2 & & & & \\
\cline{2-8}
x^2+x-1 & x^3 & + & 3x^2 & + & x & + & 4 \\
\multicolumn{1}{r}{} & x^3 & + & x^2 & - & x & & \\
\cline{2-8}
\multicolumn{1}{r}{} & & & 2x^2 & + & 2x & + & 4 \\
\multicolumn{1}{r}{} & & & 2x^2 & + & 2x & - & 2 \\
\cline{4-8}
\multicolumn{1}{r}{} & & & & & & & 1
\end{array}$$

Subtracting $x+2$ times E_1 from E_0 gives

$$1 \quad = \quad f\cdot\Big(0-(-1)(x+2)\Big) \quad + \quad g\cdot\Big(3-(4x)(x+2)\Big)$$

and so

$$E_2 \quad : \quad 1 \quad = \quad f\cdot(x+2) \quad + \quad g\cdot(x^2+2x+3)$$

Since $x+2$ is monic, this equation is E_2. The remainder of any polynomial when divided by 1 is zero, so the algorithm stops here. Hence

$$\gcd(f,g) = 1 = f\cdot(x+2) + g\cdot(x^2+2x+3)$$

Theorem 4.2.11. *Let F be a field and $f,g \in F[x]$ not both 0_F.*

(a) *There exists a unique greatest common divisor d of f and g.*

(b) *There exists $u,v \in F[x]$ with $d = fu+gv$.*

(c) *If c is a common divisor of f and g, then $c|d$.*

Proof. By the Euclidean algorithm 4.2.9 there exist $u,v \in F[x]$ such that $d := fu+gv$ is a greatest common divisor f and g. This proves the existence of d and (b).

To prove (c) let c be any common divisor of a and b. Since $d = fu+gv$ we conclude from 3.4.3(d) that $c|d$.

It remains to prove the uniqueness of a greatest common divisor. So let e be any greatest common divisor of f and g. Then e divides f and g and (c) shows that $e|d$. Since both d and e are greatest common divisors of f and g we have $\deg e \le \deg d$ and $\deg e \le \deg d$. Thus $\deg d = \deg e$. Since also $e|d$ we conclude from 4.2.4 that $d \sim e$. As d and e are monic this implies that $d = e$, see 4.2.6(b). Thus d is the unique greatest common divisor of f and g. $\square$

Definition 4.2.12. *Let F be a field and $f, g \in F[x]$. f and g are called* relatively prime *if f and g are not both 0_F and* $\gcd(f, g) = 1_F$.

Corollary 4.2.13. *Let F be a field and $f, g \in F[x]$. Then f and g are relatively prime if and only if there exist $u, v \in F[x]$ with $fu + gv = 1_F$.*

Proof. $\Longrightarrow$: Suppose that f and g are relatively prime. Then f and g are not both 0_F and $\gcd(f, g) = 1_F$. So by 4.2.11(c) there exist $u, v \in F[x]$ with $fu + gv = 1_F$.

$\Longleftarrow$: Suppose that there exist $u, v \in F[x]$ with $fu+gv = 1_F$. Since $1_F \neq 0_F$ this implies that f and g are not both 0_F. Note that 1_F is a monic common divisor of f and g. Let c be any common divisor of f and g. Since $1_F = fu + gv$ we conclude that $c \mid 1_F$ (see 3.4.3(d)). Hence $\deg c \leq \deg 1_F$ by 4.2.2. Thus 1_F is a greatest common divisor of f and g and so f and g are relatively prime. □

Proposition 4.2.14. *Let F be a field and $f, g, h \in F[x]$. Suppose that f and g are relatively prime and $f|gh$. Then $f|h$.*

Proof. Since f and g are relatively prime 4.2.13 shows that there exist $u, v \in F[x]$ with $fu+gv = 1_F$. Multiplication with h gives $(fu)h + (gv)h = h$ and so (using the General Commutative Law)

$$f \cdot (uh) + (gh) \cdot v = h.$$

Since f divides f and f divides gh, 3.4.3 now implies that $f|h$. □

Exercises 4.2:

#1. Let F be a field and $a, b \in F$ with $a \neq b$. Show that $x + a$ and $x + b$ are relatively prime in $F[x]$.

#2. Use the Euclidean Algorithm to find the gcd of the given polynomials in the given polynomial ring.

(a) $x^4 - x^3 - x^2 + 1$ and $x^3 - 1$ in $\mathbb{Q}[x]$.

(b) $x^5 + x^4 + 2x^3 - x^2 - x - 2$ and $x^4 + 2x^3 + 5x^2 + 4x + 4$ in $\mathbb{Q}[x]$.

(c) $x^4 + 3x^2 + 2x + 4$ and $x^2 - 1$ in $\mathbb{Z}_5[x]$.

(d) $4x^4 + 2x^3 + 6x^2 + 4x + 5$ and $3x^3 + 5x^2 + 6x$ in $\mathbb{Z}_7[x]$.

(e) $x^3 - ix^2 + 4x - 4i$ and $x^2 + 1$ in $\mathbb{C}[x]$.

(f) $x^4 + x + 1$ and $x^2 + x + 1$ in $\mathbb{Z}_2[x]$.

#3. Let F be a field and $f \in F[x]$ such that $f|g$ for every non-constant polynomial $g \in F[x]$. Show that f is a constant polynomial.

#4. Let F be a field and $f, g, h \in F[x]$ with f and g relatively prime. If $f|h$ and $g|h$, prove that $fg|h$.

#5. Let F be a field and $f, g, h \in F[x]$. Suppose that $g \neq 0_F$ and $\gcd(f, g) = 1_F$. Show that $\gcd(fh, g) = \gcd(h, g)$.

#6. Let F be a field and $f, g \in \mathbb{F}[x]$ such that h is non-zero and one of f and g is non-zero. Let $d = \gcd(f, g)$ and let $\hat{f}, \hat{g} \in F[x]$ with $f = \hat{f}d$ and $g = \hat{g}d$. Then $\gcd(\hat{f}, \hat{g}) = 1_F$.

#7. Let F be a field and $f, g, h \in F[x]$ with $f|gh$. Show that there exist $\tilde{g}, \tilde{h} \in F[x]$ with $\tilde{g}|g, \tilde{h}|h$ and $f = \tilde{g}\tilde{h}$.

4.3 Irreducible Polynomials

Definition 4.3.1. *Let F be a field and $f \in F[x]$.*

(a) *f is called constant if $f \in F$, that is if* $\deg f \leq 0$.

(b) *Then f is called* irreducible *provided that*

  (i) *f is not constant, and*

  (ii) *if $g \in F[x]$ with $g|f$, then*

$$g \sim 1_F \quad \textit{or} \quad g \sim f.$$

(c) *f is called* reducible *provided that*

  (i) *$f \neq 0_F$, and*

  (ii) *there exists $g \in F[x]$ with*

$$g|f, \quad g \nsim 1_F, \quad \textit{and} \quad g \nsim f.$$

Proposition 4.3.2. *Let F be a field and $0_F \neq f \in F[x]$. Then the following statements are equivalent:*

(a) *f is reducible.*

(b) *f is divisible by a non-constant polynomial of lower degree.*

(c) *f is the product of two polynomials of lower degree.*

(d) *f is the product of two non-constant polynomials of lower degree.*

(e) *f is the product of two non-constant polynomials.*

(f) *f is not constant and f is not irreducible.*

Proof. (a) $\Longrightarrow$ (b): Suppose f is reducible. Then by Definition 4.3.1 there exist $g \in F[x]$ with $g|f$, $g \nsim 1_F$ and $g \nsim f$. As $g|f$ and $f \neq 0_F$ we have $g \neq 0_F$ (see 3.4.2). By 4.2.3 all non-zero constant are associated to 1_F. Since $g \nsim 1_F$ we conclude that g is not constant. By 4.2.4, if $g|f$ and $\deg f = \deg g$, then $g \sim f$. As $g|f$ and $g \nsim f$ we conclude that $\deg f \neq \deg g$. Also by 4.2.2 since $g|f$ we have $\deg g \leq \deg f$ and so $\deg g < \deg f$. Thus g is a non-constant polynomials of lower degree than f which divides f. Thus (b) holds.

(b) $\Longrightarrow$ (c): Let g be a non-constant polynomial of lower degree than f with $g|f$. Then $\deg g > 0$, $\deg g < \deg f$ and $f = gh$ for some $h \in F[x]$. Since $f \neq 0_F$ we conclude $h \neq 0_F$. By 4.1.10(a) $\deg f = \deg g + \deg h$ and since $\deg g > 0$, $\deg h < \deg f$. Thus (c) holds.

(c) $\Longrightarrow$ (d): Suppose $f = gh$ with $\deg g < \deg f$ and $\deg h < \deg f$. By 4.1.10 $\deg f = \deg g + \deg h$. Since $\deg g < \deg f$ we conclude that $\deg h > 0$. So h is not constant. Similarly g is not constant. Thus (d) holds.

(d) $\Longrightarrow$ (e): Obvious.

(e) $\Longrightarrow$ (f): Suppose $f = gh$ where g and h are non-constant polynomials in $F[x]$. Then $g|f$. Since g is not constant, Lemma 4.2.3 gives $g \nsim 1_F$. Since $\deg h > 0$ and $\deg f = \deg g + \deg h$ (4.1.10(a)) we have $\deg f > \deg g$. Since g is not constant, $\deg g > 0$ and so also $\deg f > 0$ and f is not constant. Also $\deg f \neq \deg g$ and 4.2.4 gives $g \nsim f$. Thus by Definition 4.3.1 f is not irreducible. So (f) holds.

(f) $\Longrightarrow$ (a): Suppose f is not constant and f is not irreducible. Then by Definition 4.3.1 there exists $g \in F[x]$ with $g|f$, $g \nsim 1_F$ and $g \nsim f$. So by Definition 4.3.1, f is reducible and (a) holds. $\square$

Remark 4.3.3. *Let F be a field.*

(a) *A non-constant polynomial in $F[x]$ is reducible if and only if its is not irreducible.*

(b) *A constant polynomial in $F[x]$ is neither reducible nor irreducible.*

Proof. Let $f \in F[x]$ with $f \neq 0_F$. Then 4.3.2(a),(f) shows that

$(*)$ $\qquad f$ is reducible $\qquad$ if and only if $\qquad f$ non-constant and f is not irreducible.

(a): Let f be non-constant polynomial in $F[x]$. Then $f \neq 0_R$ and (*) shows that f is reducible if and only if f is not irreducible.

(b): By definition irreducible polynomials are not constant. Let $f \in F[x]$ be reducible. By definition of a reducible polynomial, $f \neq 0_R$ and so (*) shows that f is not constant. $\square$

Lemma 4.3.4. *Let F be a field and p a non-constant polynomial in $F[x]$. Then the following statement are equivalent:*

(a) *p is irreducible.*

(b) *Whenever $g, h \in F[x]$ with $p|gh$, then $p|g$ or $p|h$.*

(c) *Whenever $g, h \in F[x]$ with $p = gh$, then g or h is constant.*

Proof. (a) $\Longrightarrow$ (b): Suppose p is irreducible and let $g, h \in F[x]$ with $p|gh$. Put $d := \gcd(p, g)$. By definition of 'gcd', $d|p$ and since p is irreducible, $d \sim 1_F$ or $d \sim p$. We treat these two cases separately.

Suppose that $d \sim 1_F$. Since both d and 1_F are monic we conclude from 4.2.6 that $d = 1_F$. So p and g are relatively prime and, since $p|gh$, 4.2.14 implies $p|h$.

If $d \sim p$, then since $d|g$, 3.4.11(c) gives $p|g$.

(b) $\Longrightarrow$ (c): Suppose (b) holds and let $g, h \in F[x]$ with $p|gh$. Note that $p = p1_F$. So $p|p$ and since $p = gh$ we get $p|gh$. From (b) we conclude $p|g$ or $p|h$. Since the situation is symmetric in g and h we may assume $p|g$. Since $p \neq 0_F$ and $p = gh$ we get $g \neq 0_F$ and $h \neq 0_F$. From $p|g$ and 4.2.2 we have $\deg p \leq \deg g$. On the other hand by 4.1.10(a), $\deg p = \deg gh = \deg g + \deg h \geq \deg g$. Thus $\deg g = \deg p$ and $\deg h = 0$. So $h \in F$.

(c) $\Longrightarrow$ (a): Suppose (c) hold. Then p is not a product of two constant polynomials in $F[x]$. Hence 4.3.2(b) shows that p is reducible. Since p is not constant, this means that p is irreducible (see 4.3.3(a)). □

Lemma 4.3.5. *Let F be a field and let p be an irreducible polynomial in $F[x]$. If $a_1, \ldots, a_n \in F[x]$ and $p|a_1a_2\ldots a_n$, then $p|a_i$ for some $1 \leq i \leq n$.*

Proof. By induction on n. For $n = 1$ the statement is obviously true. So suppose the statment is true for $n = k$ and that $p|a_1\ldots a_k a_{k+1}$. By 4.3.4, $p|a_1\ldots a_k$ or $p|a_{k+1}$. In the first case the induction assumption implies that $p|a_i$ for some $1 \leq i \leq k$. So in any case $p|a_i$ for some $1 \leq i \leq k+1$. Thus the Lemma holds for $k + 1$ and so by the Principal of Mathematical Induction (0.4.2) the Lemma holds for all positive integer n. □

Lemma 4.3.6. *Let F be a field and p, q irreducible polynomials in $F[x]$. Then $p|q$ if and only if $p \sim q$.*

Proof. If $p \sim q$, then $p|q$, by 3.4.9. So suppose that $p|q$. Since q is irreducible, $p \sim 1_F$ or $p \sim q$. Since p is irreducible, $p \notin F$ and so by 4.2.3, $p \nsim 1_F$. Thus $p \sim q$. □

Lemma 4.3.7. *Let F be a field and $f, g \in F[x]$ with $f \sim g$. Then f is irreducible if and only if g is irreducible.*

Proof. $\Longrightarrow$: Suppose f is irreducible. Then $f \notin F$ and so $\deg f \geq 1$. Since $f \sim g$, 4.2.4 implies $\deg g = \deg f \geq 1$. Hence $g \notin F$. Let $h \in F[x]$ with $h|g$. Since $f \sim g$, 3.4.11 implies $h|f$. Since f is irreducible we conclude $h \sim 1_F$ or $h \sim f$. In the latter case, since $\sim$ is transitive (3.4.7) $h \sim g$. Hence $h \sim 1_F$ or $h \sim g$ and so g is irreducible.

$\Longleftarrow$: Suppose g is irreducible. Since $\sim$ is symmetric by 3.4.7, we have $g \sim f$. So we can apply the '$\Longrightarrow$'-case with f and g interchanged to conclude that f is irreducible. □

Theorem 4.3.8 (Unique Factorization Theorem)**.** *Let F be a field and f a non-constant polynomial in $F[x]$.*

(a) *f is the product of irreducible polynomials in $F[x]$.*

(b) *If n, m are positive integers and $p_1, p_2, \ldots, p_n$ and $q_1, \ldots q_m$ are irreducible polynomials in $F[x]$ with*

$$f = p_1 p_2 \ldots p_n \quad \textit{and} \quad f = q_1 q_2 \ldots q_m,$$

then $n = m$ and possibly after reordering the q_i's,

$$p_1 \sim q_1, \quad p_2 \sim q_2, \quad \ldots, \quad p_n \sim q_n.$$

In more precise terms: there exists a bijection $\pi : \{1, \ldots n\} \to \{1, \ldots m\}$ such that

$$p_1 \sim q_{\pi(1)}, \quad p_2 \sim q_{\pi(2)}, \quad \ldots, \quad p_n \sim q_{\pi(n)}.$$

Proof. (a) The proof is by complete induction on $\deg f$. So suppose that every non-constant polynomial of lower degree than f is a product of irreducible polynomials.

Suppose that f is irreducible. Then f is the product of one irreducible polynomial (namely itself).

Suppose f is not irreducible. Since $f \notin F$, 4.3.2 shows that $f = gh$ where g and h are non-constant polynomials of lower degree than f. By the induction assumption both g and h are products of irreducible polynomials. Hence also $f = gh$ is the product of irreducible polynomials.

(b) The proof of (a) is by complete induction on n. So let k be a positive integer and suppose that (b) holds whenever $n < k$. Suppose also that

$$(*) \qquad f = p_1 p_2 \ldots p_k \qquad \text{and} \qquad f = q_1 q_2 \ldots q_m,$$

where m is a positive integer and $p_1, \ldots, p_k, q_1, \ldots q_m$ are irreducible polynomials in $F[x]$.

Suppose first that f is irreducible. Then by 4.3.2 f is not the product of two non-constant polynomials in $F[x]$. Hence (*) implies $k = m = 1$. Thus $p_1 = f = q_1$. Since $\sim$ is reflexive this gives $p_1 \sim q_1$ and so (b) holds for $n = k$ in this case.

Suppose next that f is not irreducible. Then $p_1 \neq f \neq q_1$ and so $k \geq 2$ and $m \geq 2$.

Since $f = (p_1 \ldots p_{k-1})p_k$ we see that p_k divides f. By (*) $f = q_1 \ldots q_m$ and so p_k divides $q_1 \ldots q_m$. Hence by 4.3.5, $p_k | q_j$ for some $1 \leq j \leq m$. As p_k and q_j are irreducible we get from 4.3.6 that $p_k \sim q_j$. Reordering the q_i's we may assume that

$$p_k \sim q_m.$$

Then $p_k = q_m u$ for some unit $u \in F[x]$. Thus

$$\big((p_1 u)p_2 \ldots p_{k-1}\big)q_m = (p_1 \ldots p_{k-1})(q_m u) = p_1 \ldots p_{k-1} p_k = f = (q_1 \ldots q_{m-1})q_m.$$

By 4.1.10(c) $F[x]$ is an integral domain. Since $q_m \neq 0_F$, the Cancellation Law 3.2.19 gives

$$(p_1 u)p_2 \ldots p_{k-1} = q_1 \ldots q_{m-1}.$$

Since u is a unit, $p_1u \sim p_1$. Thus since p_1 is irreducible also p_1u is irreducible by 4.3.7. The induction assumption now implies that $k-1 = m-1$ and that, after reordering the q_i's,

$$p_1u \sim q_1, \quad p_2 \sim q_2, \quad \dots \quad p_{k-1} \sim q_{k-1}.$$

From $k-1 = m-1$ we get $k = m$. As $p_1 \sim p_1u$ and $p_1u \sim q_1$ we have $p_1 \sim q_1$, by transitivity of $\sim$. Thus

$$p_1 \sim q_1, \quad p_2 \sim q_2 \quad \dots \quad p_{k-1} \sim q_{k-1},$$

Moreover, as $p_k \sim q_m$ and $m = k$ we have $p_k \sim q_k$. Thus (b) holds for $n = k$. By the principal of complete induction, (b) holds for all positive integers n. $\square$

Exercises 4.3:

#1. Find all irreducible polynomials of

(a) degree two in $\mathbb{Z}_2[x]$.

(b) degree three in $\mathbb{Z}_2[x]$.

(c) degree two in $\mathbb{Z}_3[x]$.

#2. (a) Show that $x^2 + 2$ is irreducible in $\mathbb{Z}_5[x]$.

(b) Factor $x^4 - 4$ as a product of irreducibles in $\mathbb{Z}_5[x]$.

#3. Let F be a field. Prove that every non-constant polynomial f in $F[x]$ can be written in the form $f = cp_1p_2 \dots p_n$ with $c \in F$ and each p_i monic irreducible in $F[x]$. Show further that if $f = dq_1 \dots q_m$ with $d \in F$ and each q_i monic and irreducible in $F[x]$, then $m = n$, $c = d$ and after reordering and relabeling, if necessary, $p_i = q_i$ for each i.

#4. Let F be a field and $p \in F[x]$ with $p \notin F$. Show that the following two statements are equivalent:

(a) p is irreducible

(b) If $g \in F[x]$ then $p|g$ or $\gcd(p, g) = 1_F$.

#5. Let F be a field and let $p_1, p_2, \dots p_n$ be irreducible monic polynomials in $F[x]$ such that $p_i \neq p_j$ for all $1 \leq i < j \leq n$. Let $f, g \in F[x]$ and suppose that $f = p_1^{k_1} p_2^{k_2} \dots p_n^{k_n}$ and $g = p_1^{l_1} p_2^{l_2} \dots p_n^{l_n}$ for some $k_1, k_2, \dots, k_n, l_1, l_2 \dots, l_n \in \mathbb{N}$.

(a) Show that $f|g$ in $F[x]$ if and only if $k_i \leq l_i$ for all $1 \leq i \leq n$.

(b) For $1 \leq i \leq n$ define $m_i = \min(k_i, l_i)$. Show that $\gcd(f, g) = p_1^{m_1} p_2^{m_2} \dots p_n^{m_n}$.

4.4 Polynomial function

Theorem 4.4.1. *Let R and S be commutative rings with identities, $\alpha : R \to S$ a homomorphism of rings with $\alpha(1_R) = 1_S$ and let $s \in S$.*

(a) *There exists a unique ring homomorphism $\alpha_s : R[x] \to S$ such that $\alpha_s(x) = s$ and $\alpha_s(r) = \alpha(r)$ for all $r \in R$.*

(b) *For all $f = \sum_{i=0}^{\deg f} f_i x^i$ in $R[x]$, $\alpha_s(f) = \sum_{i=0}^{\deg f} \alpha(f_i)s^i$.*

Proof. Suppose first that $\beta : R[x] \to S$ is a ring homomorphism with

$$(*) \qquad \beta(x) = s \quad \text{and} \quad \beta(r) = \alpha(r)$$

for all $r \in R$. Let $f \in R[x]$.

Then

$$\begin{aligned}
\beta(f) &= \beta\left(\sum_{i=0}^{\deg f} f_i x^i\right) && -4.1.7(\mathrm{d})\\
&= \sum_{i=0}^{\deg f} \beta(f_i x^i) && -\beta \text{ is a homomorphism}\\
&= \sum_{i=0}^{\deg f} \beta(f_i)\beta(x)^i && -\beta \text{ is a homomorphism}\\
&= \sum_{i=0}^{\deg f} \alpha(f_i)s^i. && -\ (*)
\end{aligned}$$

This proves (b) and the uniqueness of α_s.

It remains to prove the existence. We use (b) to define α_s. That is we define

$$\alpha_s : R[x] \to S, \quad f \mapsto \sum_{i=0}^{\deg f} \alpha(f_i)s^i.$$

It follows that

$$\alpha_s(x) = \alpha_s(1_R x) = \alpha(1_R)s = 1_S s = s$$

and if $r \in R$, then

$$\alpha_s(r) = \alpha_s(rx^0) = \alpha(r)s^0 = \alpha(r)1_S = \alpha(r).$$

Let $f, g \in R[x]$. Put $n = \max(\deg f, \deg g)$ and $m = \deg f + \deg g$.

$$\begin{aligned}
\alpha_s(f+g) &= \alpha_s\left(\sum_{i=0}^{n}(f_i+g_i)x^i\right) && \text{– 4.1.4(a) with } R[x] \text{ in place of } P\\
&= \sum_{i=0}^{n}\alpha(f_i+g_i)s^i && \text{– definition of } \alpha_s\\
&= \sum_{i=0}^{n}\Big(\alpha(f_i)+\alpha(g_i)\Big)s^i && \text{– Since } \alpha \text{ is a homomorphism}\\
&= \left(\sum_{i=0}^{\deg f}\alpha(f_i)s^i\right)+\left(\sum_{i=0}^{\deg g}\alpha(g_i)s^i\right) && \text{– 4.1.4(a) with } (S,S,x) \text{ in place of } (R,P,x)\\
&= \alpha_s(f)+\alpha_s(g) && \text{– definition of } \alpha_s, \text{twice}
\end{aligned}$$

$$\begin{aligned}
\alpha_s(fg) &= \alpha_s\left(\sum_{k=0}^{m}\left(\sum_{i=0}^{k}f_ig_{k-i}\right)x^k\right) && \text{– 4.1.4(a) with } R[x] \text{ in place of } P\\
&= \sum_{k=0}^{m}\alpha\left(\sum_{i=0}^{k}f_ig_{k-i}\right)s^k && \text{– definition of } \alpha_s\\
&= \sum_{k=0}^{m}\left(\sum_{i=0}^{k}\alpha(f_i)\alpha(g_{k-i})\right)s^k && \text{– Since } \alpha \text{ is a homomorphism}\\
&= \left(\sum_{i=0}^{\deg f}\alpha(f_i)s^i\right)\cdot\left(\sum_{j=0}^{\deg g}\alpha(g_j)s^j\right) && \text{– 4.1.4(a) with } (S,S,x) \text{ in place of } (R,P,x)\\
&= \alpha_s(f)\cdot\alpha_s(g) && \text{– definition of } \alpha_s, \text{twice}
\end{aligned}$$

So α_s is a homomorphism and the Theorem is proved. $\square$

Example 4.4.2. Compute α_s in the following cases:

(1) R is a commutative ring with identity, $S = R$, $\alpha = \mathrm{id}_R$ and $s \in R$.

(2) R is a commutative ring with identity, $S = R[x]$, $\alpha(r) = r$ and $s = x$.

(3) $R = \mathbb{Z}$, n is an integer, $S = \mathbb{Z}_n[x]$, $\alpha(r) = [r]_n$ and $s = x$.

(1) $\alpha_s(f) = \sum_{i=0}^{\deg f}\alpha(f_i)s^i = \sum_{i=0}^{\deg f} f_is^i$.

(2) $\alpha_s(f) = \sum_{i=0}^{\deg f}\alpha(f_i)s^i = \sum_{i=0}^{\deg f} f_ix^i = f$

So α_s is identity function on $R[x]$.

(3) Note first that by Example 3.3.2 $\alpha : \mathbb{Z} \to \mathbb{Z}_n[x], r \to [r]_n$ is a homomorphism. Also

$$\alpha_s(f) = \sum_{i=0}^{\deg f} \alpha(f_i)s^i = \sum_{i=0}^{\deg f} [f_i]_n x^i$$

So $\alpha_s(f)$ is obtain from f by viewing each coefficient as congruence class modulo n rather than an integer.

Definition 4.4.3. *Let I be a set and R a ring.*

(a) $\mathrm{Fun}(I, R)$ *is the set of all functions from I to R.*

(b) *For $\alpha, \beta \in \mathrm{Fun}(I, R)$ define $\alpha + \beta$ in $\mathrm{Fun}(I, R)$ by*

$$(\alpha + \beta)(i) = \alpha(i) + \beta(i)$$

for all $i \in I$.

(c) *For $\alpha, \beta \in \mathrm{Fun}(I, R)$ define $\alpha\beta$ in $\mathrm{Fun}(I, R)$ by*

$$(\alpha\beta)(i) = \alpha(i)\beta(i)$$

for all $i \in I$.

(d) *For $r \in R$ define $r^* \in \mathrm{Fun}(I, R)$ by*

$$r^*(i) = r$$

for all $i \in I$.

(e) $\mathrm{Fun}(R) = \mathrm{Fun}(R, R)$.

Lemma 4.4.4. *Let I be a set and R a ring.*

(a) $\mathrm{Fun}(I, R)$ *together with the above addition and multiplication is a ring.*

(b) 0_R^* *is the additive identity in* $\mathrm{Fun}(I, R)$.

(c) *If R has a multiplicative identity 1_R, then 1_R^* is a multiplicative identity in* $\mathrm{Fun}(I, R)$.

(d) $(-\alpha)(i) = -\alpha(i)$ *for all $\alpha \in \mathrm{Fun}(I, R)$, $i \in I$.*

(e) *The function $\tau : R \to \mathrm{Fun}(I, R), r \to r^*$ is a homomorphism. If $I \neq \emptyset$, than τ is 1-1.*

Proof. Note that $\mathrm{Fun}(I, R) = \bigtimes_{i \in I} R$ and so (a)-(d) follow from F.1.2.

(e) Let $a, b \in R$ and $i \in I$. Then

$$\begin{aligned}
(a+b)^*(i) &= a + b && - \text{ definition of } (a+b)^* \\
&= a^*(i) + b^*(i) && - \text{ definition of } a^* \text{ and } b^* \\
&= (a^* + b^*)(i) && - \text{ definition of addition of functions}
\end{aligned}$$

Thus $(a+b)^* = a^* + b^*$ by 0.3.11 and so $\tau(a+b) = \tau(a) + \tau(b)$ by definition of τ.

Similarly,

$$\begin{aligned}
(ab)^*(i) &= ab && - \text{ definition of } (ab)^* \\
&= a^*(i)b^*(i) && - \text{ definition of } a^* \text{ and } b^* \\
&= (a^*b^*)(i) && - \text{ definition of multiplication of function}
\end{aligned}$$

Hence $(ab)^* = a^*b^*$ by 0.3.11 and so $\tau(ab) = \tau(a)\tau(b)$ by definition of τ.

Thus τ is a homomorphism .

Suppose that $I \neq \emptyset$ and $\tau(a) = \tau(b)$. Then $a^* = b^*$ and there exists $i \in I$. So $a = a^*(i) = b^*(i) = b$ and τ is 1-1. □

Notation 4.4.5. *Let R be a commutative ring with identity and $f \in R[x]$. For $f = \sum_{i=0}^{\deg f} f_i x^i \in F[x]$ define the function*

$$f^* : R \to R$$

by

$$f^*(r) = \sum_{i=0}^{\deg f} f_i r^i$$

for all $r \in R$.

f^* is called the polynomial function *on R induced by f.*

Remark 4.4.6. *Let R be a commutative ring with identity.*

(a) *Let* $\mathrm{id} : R \to R, r \to r$ *be the identity function on R and for $r \in R$ let* $\mathrm{id}_r : R[x] \to R$ *be the homomorphism from 4.4.1. Then*

$$f^*(r) = \mathrm{id}_r(f)$$

for all $f \in F[x]$ and $r \in R$.

(b) *Let $f \in R[x]$ be constant polynomial. Then the definitions of $f^* \in \mathrm{Fun}(R)$ in 4.4.5 and in 4.4.3 coincide.*

Proof. (a): By Example 4.4.2(1) $\mathrm{id}_r(f) = \sum_{i=0} \deg f f_i r^i$ and so $\mathrm{id}_r(f) = f^*(r)$.

(b) Since $f \in F$, $f = fx^0$ and so $f^*(r) = f$ for all $r \in R$. □

The following example shows that it is very important to distinguish between a polynomial f and its induced polynomial function f^*.

Example 4.4.7. Determine the functions induced by the polynomials of degree at most two in $\mathbb{Z}_2[x]$.

f	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
$f^*(0)$	0	1	0	1	0	1	0	1
$f^*(1)$	0	1	1	0	1	0	0	1

We conclude that $x^* = (x^2)^*$. So two distinct polynomials can lead to the same polynomial function. Also $(x^2+x)^*$ is the zero function but x^2+x is not the zero polynomial.

Theorem 4.4.8. *Let R be commutative ring with identity.*

(a) *$f^* \in \mathrm{Fun}(R)$ for all $f \in R[x]$.*

(b) *$(f+g)^*(r) = f^*(r) + g^*(r)$ and $(fg)^*(r) = f^*(r)g^*(r)$ for all $f, g \in R[x]$ and $r \in R$.*

(c) *$(f+g)^* = f^* + g^*$ and $f^*g^* = f^*g^*$ for all $f, g \in R[x]$.*

(d) *The function $R[x] \to \mathrm{Fun}(R)$, $f \to f^*$ is a ring homomorphism.*

Proof. (a) By definition f^* is a function from R to R. Hence $f^* \in \mathrm{Fun}(R)$.

(b)

$$\begin{aligned}
(f+g)^*(r) &= \mathrm{id}_r(f+g) && -4.4.6(a) \\
&= \mathrm{id}_r(f) + \mathrm{id}_r(g) && - \mathrm{id}_r \text{ is a homomorphism} \\
&= f^*(r) + g^*(r) && -4.4.6(a), \text{twice}
\end{aligned}$$

and similarly

$$\begin{aligned}
(fg)^*(r) &= \mathrm{id}_r(fg) && -4.4.6(a) \\
&= \mathrm{id}_r(f)\mathrm{id}_r(g) && - \mathrm{id}_r \text{ is a homomorphism} \\
&= f^*(r)g^*(r) && -4.4.6(a), \text{twice}
\end{aligned}$$

(c) Let $r \in R$. Then

$$\begin{aligned}
(f+g)^*(r) &= f^*(r) + g^*(r) && - \text{(b)} \\
&= (f^* + g^*)(r) && - \text{Definition of addition in } \mathrm{Fun}(R)
\end{aligned}$$

So $(f+g)^* = f^* + g^*$. Similarly

$$\begin{aligned}
(fg)^*(r) &= f^*(r)g^*(r) && - \text{(b)} \\
&= (f^*g^*)(r) && - \text{Definition of multiplication in } \mathrm{Fun}(R)
\end{aligned}$$

and so $(fg)^* = f^*g^*$.

(d) Follows from (c). $\square$

Lemma 4.4.9. *Let F be a field, $f \in F[x]$ and $a \in F$. Then the remainder of f when divided by $x - a$ is $f^*(a)$.*

Proof. Let r be the remainder of f when divided by $x - a$. So $r \in F[x]$, $\deg r < \deg(x - a)$ and there exists $q \in F[x]$ with

$$(*) \qquad f = q \cdot (x - a) + r.$$

Since $\deg(x - a) = 1$ we have $\deg r \leq 0$ and so $r \in F$. Thus

$$(**) \qquad r^*(t) = r$$

for all $t \in R$.

$$\begin{array}{ccccc} f^*(a) & \overset{(*)}{=} & (q \cdot (x - a) + r)^*(a) & \overset{4.4.8(b)}{=} & (q \cdot (x - a))^*(a) + r^*(a) \\ & \overset{4.4.8(b)}{=} & q^*(a) \cdot (x - a)^*(a) + r^*(a) & \overset{\text{Def } (x-a)^*}{=} & q^*(a)(a - a) + r^*(a) \\ & \overset{(**)}{=} & q^*(a)(a - a) + r & \overset{3.2.11(f)}{=} & q^*(a) \cdot 0_F + r \\ & \overset{3.2.11(c)}{=} & 0_F + r & \overset{\text{Ax } 4}{=} & r \end{array}$$

□

Definition 4.4.10. *Let R be a commutative ring with identity, $f \in R[x]$ and $a \in R$. Then a is called a root of f if $f^*(a) = 0_R$.*

Theorem 4.4.11 (Factor Theorem)**.** *Let F a field, $f \in F[x]$ and $a \in F$. Then a is a root of f if and only if $x - a | f$.*

Proof. Let r be the remainder of f when divided by $x - a$. Then

$$\begin{array}{ccl} & x - a | f & \\ \iff & r = 0_F & - \ 4.2.1 \\ \iff & f^*(a) = 0_F & - f^*(a) = r \text{ by } 4.4.9 \\ \iff & a \text{ is a root of } f & - \text{ Definition of root} \end{array}$$

□

Lemma 4.4.12. *Let R be commutative ring with identity and $f \in R[x]$.*

(a) *Let $g \in R[x]$ with $g | f$. Then any root of g in R is also a root of f in R.*

(b) *Let $a \in R$ and $g, h \in R[x]$ with $f = gh$. Suppose that R is field or an integral domain. Then a is a root of f if and only if a is a root of g or a is a root of h.*

Proof. (a): Let a be a root of g. Then $g^*(a) = 0_R$. Since $g|f$, there exists $h \in R[x]$ with $f = gh$. Then

$$f^*(a) = (gh)^*(a) \overset{4.4.8(c)}{=} g^*(a)h^*(a) = 0_R \cdot h^*(a) = 0_R.$$

Thus a is a root of f. So (a) holds.

(b) : Suppose that R is field or an integral domain. By 3.2.22 all fields are integral domains. Thus R is an integral domain and so (Ax 11) holds. Hence

$$\begin{array}{lccl}
 & & a \text{ is a root of } f & \\
\iff & & f^*(a) = 0_R & -\text{ definition of root} \\
\iff & & (gh)^*(a) = 0_R & -f = gh \\
\iff & & g^*(a)h^*(a) = 0_R & -4.4.8(c) \\
\iff & g^*(a) = 0_R & \text{or} \quad h^*(a) = 0_R & -(\text{Ax } 11) \\
\iff & a \text{ is a root of } g & \text{or} \quad a \text{ is a root of } h & -\text{definition of root, twice}
\end{array}$$

□

Example 4.4.13. (1) Let R be a commutative ring with identity and $a \in R$. Find the roots of $x - a$ in R.

Let $b \in R$. Then $(x-a)^*(b) = b - a$. So b is a root of $x - a$ if and only if $b - a = 0_R$ and if and only if $b = a$. Hence a is the only root of $x - a$.

(2) Find the roots of $x^2 - 1$ in $\mathbb{Z}$. Note that

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

Since $\mathbb{Z}$ is an integral domain, 4.4.12 show that the roots of x^2-1 are the roots of $x-1$ together with the roots of $x-(-1)$. So by (1) the roots of $x^2 - 1$ are 1 and -1.

(3) Find the roots of $x^2 - 1$ in $\mathbb{Z}_8$.

Since $\mathbb{Z}_8$ is not an integral domain, the argument in (2) does not work. We compute in $\mathbb{Z}_8$

$$0^2-1 = -1, (\pm 1)^2-1 = 1-1 = 0, (\pm 2)^2-1 = 4-1 = 3, (\pm 3)^2 = 9-1 = 8 = 0, 4^2-1 = 15 = -1.$$

So the roots of $x^2 - 1$ are ± 1 and ± 3. Note here that $(3-1)(3+1) = 2 \cdot 4 = 8 = 0$. So the extra root 3 comes from the fact that $2 \cdot 4 = 0$ in $\mathbb{Z}_8$ but neither 2 nor 4 is zero.

Theorem 4.4.14 (Root Theorem)**.** *Let F be a field and $f \in F[x]$ a non-zero polynomial. Then there exist a non-negative integer m, elements $a_1, \ldots, a_m \in F$ and $q \in F[x]$ such that*

(a) $m \leq \deg f$.

(b) $f = q \cdot (x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_m)$.

(c) *q has no roots in F.*

(d) *$\{a_1, a_2, \ldots, a_m\}$ is the set of roots of f in F.*

In particular, the number of roots of f is at most $\deg f$.

Proof. The proof is by complete induction on $\deg f$. So let $k \in \mathbb{N}$ and suppose that theorem holds for polynomials of degree less than k. Let f be a polynomial of degree k.

Suppose that f has no roots. Then the theorem holds with $q = f$ and $m = 0$.

Suppose next that f has a root a. Then by the Factor Theorem 4.4.11, $x - a | f$ and so

$$(*) \qquad f = g \cdot (x - a)$$

for some $g \in F[x]$. By 4.1.10 $\deg f = \deg g + \deg(x - a) = \deg g + 1$ and so $\deg g = k - 1$. Hence by the induction assumption there exist a non-negative integer n, elements $a_1, \ldots, a_n \in F$ and $q \in F[x]$ such that

(i) $n \leq \deg g$.

(ii) $g = q \cdot (x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_n)$

(iii) q has no roots in F.

(iv) $\{a_1, a_2, \ldots, a_n\}$ is the set of roots of g.

Put $m = n + 1$ and $a_m = a$. Then $m = n + 1 \overset{\text{(i)}}{\leq} \deg g + 1 = (k - 1) + 1 = k = \deg f$ and so (a) holds. From $f = g \cdot (x - a) = g \cdot (x - a_m)$ and (ii) we conclude that (b) holds. By (iii), (c) holds.

Let $b \in F$. Since $f = g \cdot (x - a_m)$, 4.4.12 shows that b is a root of f if and only if b is a root of g or g is a root of $x - a_m$. Using (iv) we conclude that b root of f if and only if $b \in \{a_1, a_2, \ldots a_n\}$ or $b - a_m = 0_F$ and so if and only if $b \in \{a_1, a_2 \ldots, a_n, a_m\} = \{a_1, \ldots, a_m\}$. Thus also (d) holds. $\square$

Remark 4.4.15. *$x^2 - 1$ has four roots in $\mathbb{Z}_8$, namely ± 1 and ± 3, see Example 4.4.13(3). So in rings without* (Ax 11) *a polynomial can have more roots than its degree.*

Lemma 4.4.16. *Let F be a field and $f \in F[x]$,*

(a) *If $\deg f = 1$, then f has a root in F.*

(b) *If $\deg f \geq 2$ and f is irreducible, then f has no root in F.*

(c) *If $\deg f = 2$ or 3, then f is irreducible if and only if f has no roots in F.*

Proof. See Exercise #1

$\square$

Exercises 4.4:

#1. Let F be a field and $f \in F[x]$. Show that

(a) If $\deg f = 1$, then f has a root in F.

(b) If $\deg f \geq 2$ and f is irreducible, then f has no root in F.

(c) If $\deg f = 2$ or 3, then f is irreducible if and only if f has no roots in F.

#2. Let F be an infinite field. Then the map $F[x] \to \mathrm{Fun}(F), f \to f^*$ is 1-1 homomorphism. In particular, if f and g in $F[x]$ induce the same function from F to F, then $f = g$.

#3. Show that $x - 1_F$ divides $a_n x^n + \ldots a_1 x + a_0$ in $F[x]$ if and only if $a_0 + a_1 + \ldots + a_n = 0$.

#4. (a) Show that $x^7 - x$ induces the zero function on $\mathbb{Z}_7$.

(b) Use (a) and Theorem 4.4.14 to write $x^7 - x$ is a product of irreducible monic polynomials in $\mathbb{Z}_7$.

#5. Let R be an integral domain and $n \in \mathbb{N}$ Let $f, g \in R[x]$. Put $n = \deg f$. If $f = 0_R$ define $f^\bullet = 0_R$ and $m_f = 0$. If $f \neq 0_R$ define

$$f^\bullet = \sum_{i=0}^{n} f_{n-i} x^i$$

and let $m_f \in \mathbb{N}$ be minimal with $f_{m_f} \neq 0_F$. Prove that

(a) $\deg f = m_f + \deg f^\bullet$.

(b) $f = x^{f_m} \cdot (f^\bullet)^\bullet$

(c) $(fg)^\bullet = f^\bullet g^\bullet$.

(d) Let $k, l \in \mathbb{N}$ and suppose that $f_0 \neq 0_R$. Then f is the product of polynomials of degree k and l in $R[x]$ if and only if $f^\bullet$ is the product of polynomials of degree k and l in $R[x]$.

(e) Suppose in addition that R is a field and let $a \in R$. Show that a is a root of $f^\bullet$ if and only if $a \neq 0_R$ and a is a root of f.

#6. Let p be a prime. Let $f, g \in \mathbb{Z}_p[x]$ and let $f^*, g^* : \mathbb{Z}_p \to \mathbb{Z}_p$ be the corresponding polynomial functions. Show that:

(a) If $\deg f < p$ and f^* is the zero function, then $f = 0_F$.

(b) If $\deg f < p, \deg g < p$ and $f \neq g$, then $f^* \neq g^*$.

(c) There are exactly p^p polynomials of degree less than p in $\mathbb{Z}_p[x]$.

(d) There exist at least p^p polynomial functions from $\mathbb{Z}_p$ to $\mathbb{Z}_p$.

(e) There are exactly p^p functions from $\mathbb{Z}_p$ to $\mathbb{Z}_p$.

(f) All functions from $\mathbb{Z}_p$ to $\mathbb{Z}_p$ are polynomial functions.

Chapter 5

Congruence Classes in F[x]

5.1 The Congruence Relation

Definition 5.1.1. *Let F be a field and $p \in F[x]$. Then the relation $\equiv \pmod{p}$ on $F[x]$ is defined by*

$$f \equiv g \pmod{p} \quad \text{if} \quad p|f-g \text{ in } F[x]$$

If $f \equiv g \pmod{p}$ we say that f and g are congruent *modulo p.*

Example 5.1.2. Let $f = x^3 + x^2 + 1, g = x^2 + x$ and $p = x^2 + x + 1$ in $\mathbb{Z}_2[x]$. Is $f \equiv g \pmod{p}$?

f and g are congruent modulo p if and only if p divides $f - g$ and so by 3.4.2 f if and only if the remainder of $f - g$ when divided by p is 0_F. So we can use the division algorithm to check whether f and g are congruent modulo p.

We have $f - g = x^3 + x + 1$ and

$$\begin{array}{r|rrrrrrr}
\multicolumn{1}{r}{} & x & + & 1 & & & & \\ \cline{2-8}
x^2+x+1 & x^3 & & & + & x & + & 1 \\
 & x^3 & + & x^2 & + & x & & \\ \cline{2-8}
 & & & x^2 & & & + & 1 \\
 & & & x^2 & + & x & + & 1 \\ \cline{3-8}
 & & & & & x & &
\end{array}$$

So the remainder of $f - g$ when divided by p is not zero and therefore

$$x^3 + x^2 + 1 \not\equiv x^2 + x \pmod{x^2 + x + 1}$$

in $\mathbb{Z}_2[x]$.

Theorem 5.1.3. *Let F be a field and $p \in F[x]$. Then the relation '$\equiv \pmod{p}$' is an equivalence relation on $F[x]$.*

Proof. We need to verify that '$\equiv \pmod p$' is reflexive, symmetric and transitive.

Reflexive: Let $f \in F[x]$. Then $f - f = 0_F = p \cdot 0_F$. So $p|f - f$ and $f \equiv f \pmod p$.

Symmetric: Let $f, g \in F[x]$ with $f \equiv g \pmod p$. Then $p|f - g$. Since $g - f = -(f - g)$, 3.4.3(b) implies that $p|g - f$. Thus $g \equiv f \pmod p$.

Transitive: Let $f, g, h \in F[x]$ with $f \equiv g \pmod p$ and $g \equiv h \pmod p$. By definition of $\equiv \pmod p$ we have $p|f - g$ and $p|g - h$. Observe that $f - h = (f - g) + (g - h)$ and so by 3.4.3(c), $p|f - h$. Thus $f \equiv h \pmod p$. □

Notation 5.1.4. *Let F be a field and $f, p \in F[x]$.*

(a) *$[f]_p$ denotes the equivalence class of '$\equiv \pmod p$' containing f. So*

$$[f]_p = \{g \in F[x] \mid f \equiv g \pmod p\}$$

$[f]_p$ is called the congruence class *of f modulo p.*

(b) *$F[x]/(p)$ is the set of congruence classes modulo p in $F[x]$. So*

$$F[x]/(p) = \left\{ [f]_p \;\middle|\; f \in F[x] \right\}$$

Theorem 5.1.5. *Let F be a field and $f, g, p \in F[x]$ with $p \neq 0_F$. Then the following statements are equivalent:*

(a) *$f = g + pk$ for some $k \in F[x]$.*

(b) *$f - g = pk$ for some $k \in F[x]$.*

(c) *$p|f - g$.*

(d) *$f \equiv g \pmod p$.*

(e) *$g \in [f]_p$.*

(f) *$[f]_p \cap [g]_p \neq \emptyset$.*

(g) *$[f]_p = [g]_p$.*

(h) *$f \in [g]_p$.*

(i) *$g \equiv f \pmod p$.*

(j) *$p|g - f$.*

(k) *$g - f = pl$ for some $l \in F[x]$.*

(l) *$g = f + pl$ for some $l \in F[x]$.*

(m) *f and g have the same remainder when divided by p.*

Proof. (a) $\Longleftrightarrow$ (b): and (k) $\Longleftrightarrow$ (l): This holds by 3.2.12.

(b) $\Longleftrightarrow$ (c): and (k) $\Longleftrightarrow$ (j): Follows from the definition of 'divide'.

(c) $\Longleftrightarrow$ (d): and (i) $\Longleftrightarrow$ (j): Follows from the definition of '$\equiv \pmod p$'.

By 5.1.3 '$\equiv \pmod p$' is an equivalence relation. We we can apply Theorem 0.5.8 and conclude that statements (d)- (i) are equivalent.

It follows that statements (a)-(l) are equivalent.

Let r_1 and r_2 be the remainders of f and g, respectively, when divided by p. Then there exist $q_1, q_2 \in F[x]$ with

$$\begin{aligned} f &= pq_1 + r_1 \quad \text{and} \quad \deg r_1 < \deg p \\ g &= pq_2 + r_2 \quad \text{and} \quad \deg r_2 < \deg p \end{aligned}$$

(m) $\Longrightarrow$ (b): Suppose (m) holds. Then $r_1 = r_2$ and

$$g - f = (pq_2 + r_2) - (pq_1 + r_1) = p \cdot (q_2 - q_1) + (r_2 - r_1) = p \cdot (q_2 - q_1).$$

So (b) holds with $k = q_2 - q_1$.

(a) $\Longrightarrow$ (m): Suppose $f = g + pk$ for some $k \in F[x]$. Then $f = (pq_2 + r_2) + pk = p(q_2 + k) + r_2$. Note that $q_2 + k \in F[x]$, $r_2 \in F[x]$ and $\deg r_2 < \deg p$. So r_2 is the remainder of f when divided by p and (m) holds. □

Theorem 5.1.6. *Let F be a field and $f, p \in F$ with $p \neq 0_F$. Then there exists a unique $r \in F[x]$ with $\deg r < \deg p$ and $[f]_p = [r]_p$, namely r is the remainder of f when divided by p.*

Proof. Let s be the remainder of f when divided by p and let $r \in F[x]$ with $\deg r < \deg p$. Since $r = p0_F + r$ and $\deg r < \deg p$, r is the remainder of r when divided by p. By 5.1.5, $[f]_p = [r]_p$ if and only f and s have the same remainder when divided by n, and so if and only if $s = r$. □

Lemma 5.1.7. *Let F be a field and $p \in F[x]$ with $p \neq 0_F$. Then*

$$F[x]/(p) = \left\{ \ [r]_p \ \middle| \ r \in F[x], \deg r < \deg p \ \right\}.$$

Proof. By definition

$$F[x]/(p) = \left\{ \ [f]_p \ \middle| \ f \in F[x] \ \right\}.$$

By 5.1.6 for each $f \in \mathbb{F}[x]$, there exists $r \in F[x]$ with $[f]_p = [r]_p$ and $\deg r < \deg p$. Thus

$$\left\{ \ [f]_p \ \middle| \ f \in F[x] \ \right\} \subseteq \left\{ \ [r]_p \ \middle| \ r \in F[x], \deg r < \deg p \ \right\}$$

The reversed inclusion is obvious. □

Example 5.1.8. Determine

(a) $\mathbb{Z}_3[x]/(x^2 + 1)$, and

(b) $\mathbb{Q}[x]/(x^3 - x + 1)$.

(a) Put $p = x^2 + 1$ in $\mathbb{Z}_3[x]$. Then $\deg p = 2$. Since $\mathbb{Z}_3 = \{0, 1, 2\}$, the polynomials of degree less than 2 in $\mathbb{Z}_3[x]$ are

$$0,\ 1,\ 2,\ x,\ x+1,\ x+2,\ 2x,\ 2x+1,\ 2x+2.$$

Thus 5.1.7 shows that

$$\begin{aligned}\mathbb{Z}_3[x]/(x^2+1) &= \left\{\ [f]_p \ \middle|\ p \in \mathbb{Z}_2[x], \deg f < 2\ \right\} \\ &= \left\{[0]_p,\ [1]_p,\ [2]_p,\ [x]_p,\ [x+1]_p,\ [x+2]_p,\ [2x]_p,\ [2x+1]_p,\ [2x+2]_p\right\}.\end{aligned}$$

(b) Any polynomial of degree less than 3 can be written as $a + bx + cx^2$, where $a, b, c \in \mathbb{Q}$. Thus

$$\mathbb{Q}[x]/(x^3 - x + 1) = \{[a + bx + cx^2]_{x^3-x+1} \mid a, b, c \in \mathbb{Q}\}.$$

Exercises 5.1:

#1. Let $f, g, p \in \mathbb{Q}[x]$. Determine whether $f \equiv g \pmod{p}$.

(a) $f = x^5 - 2x^4 + 4x^3 - 3x + 1$, $g = 3x^4 + 2x^3 - 5x^2 + 2$, $p = x^2 + 1$;

(b) $f = x^4 + 2x^3 - 3x^2 + x - 5$, $g = x^4 + x^3 - 5x^2 + 12x - 25$, $p = x^2 + 1$;

(c) $f = 3x^5 + 4x^4 + 5x^3 - 6x^2 + 5x - 7$, $g = 2x^5 + 6x^4 + x^3 + 2x^2 + 2x - 5$, $p = x^3 - x^2 + x - 1$.

#2. Show that, under congruence modulo $x^3 + 2x + 1$ in $\mathbb{Z}_3[x]$ there are exactly 27 congruence classes.

#3. Prove or disprove: Let F be a field and $f, g, k, p \in F[x]$. If p is nonzero, p is relatively prime to k and $fk \equiv gk \pmod{p}$, then $f \equiv g \pmod{p}$.

#4. Prove or disprove: Let F be a field and $f, g, p \in F[x]$. If p is irreducible and $fg \equiv 0_F \pmod{p}$, then $f \equiv 0_F \pmod{p}$ or $g \equiv 0_F \pmod{p}$.

5.2 Congruence Class Arithmetic

Theorem 5.2.1. *Let F be a field and $f, g, \tilde{f}, \tilde{g}, p$ in $F[x]$ with $p \neq 0_F$. Suppose that*

$$[f]_p = [\tilde{f}]_p \qquad \text{and} \qquad [g]_p = [\tilde{g}]_p.$$

Then

$$[f+g]_p = [\tilde{f} + \tilde{g}]_p \qquad \text{and} \qquad [fg]_p = [\tilde{f}\tilde{g}]_p$$

Proof. Since $[f]_p = [\tilde{f}]_p$ and $[g]_p = [\tilde{g}]_p$ we conclude from 5.1.5 that $\tilde{f} = f + pk$ and $\tilde{g} = g + pl$ for some $k, l \in F[x]$. Hence

$$\tilde{f} + \tilde{g} = (f + pk) + (g + pl) = (f + g) + p \cdot (k + l).$$

Since $k + l \in F[x]$, 5.1.5 gives

$$[f + g]_p = [\tilde{f} + \tilde{g}]_p.$$

Also

$$\tilde{f} \cdot \tilde{g} = (f + pk)(g + pl) = fg + p \cdot (kg + fl + kpl),$$

and since $kg + fl + kpl \in F[x]$, 5.1.5 implies

$$[fg]_p = [\tilde{f}\tilde{g}]_p.$$

□

Definition 5.2.2. *Let F be a field and $p \in F[x]$ with $p \neq 0_F$. We define an addition and multiplication on $F[x]/(p)$ by*

$$[f]_p + [g]_p = [f + g]_p \quad \textit{and} \quad [f]_p \cdot [g]_p = [f \cdot g]_p$$

for all $f, g \in F[x]$. By 5.2.1 this is well defined.

Example 5.2.3. Compute the addition and multiplication table for $\mathbb{Z}_2[x]/(x^2 + x)$.

We write $[f]$ for $[f]_{x^2+x}$. Since $\mathbb{Z}_2 = \{0, 1\}$, the polynomial of degree less than 2 in $\mathbb{Z}_2[x]$ are $0, 1, x, x + 1$. Thus 5.1.7 gives

$$\mathbb{Z}_2[x]/(x^2 + x) = \big\{[0], [1], [x], [x + 1]\big\}.$$

We compute

$+$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[0]$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[1]$	$[1]$	$[0]$	$[x+1]$	$[x]$
$[x]$	$[x]$	$[x+1]$	$[0]$	$[1]$
$[x+1]$	$[x+1]$	$[x]$	$[1]$	$[0]$

$\cdot$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[x]$	$[0]$	$[x]$	$[x]$	$[0]$
$[x+1]$	$[0]$	$[x+1]$	$[0]$	$[x+1]$

Note here that

$$[x][x + 1] = [x(x + 1)] = [x^2 + x] = [0]$$

and

$$[x+1][x+1] = [(x+1)(x+1)] = [x^2+1] = [(x^2+1)+(x^2+x)] = [x+1]$$

Observe from the above tables that $\mathbb{Z}_2[x]/(x^2+x)$ contains the subring $\{[0],[1]\}$ isomorphic to $\mathbb{Z}_2$. The next theorem shows that a similar statement holds in general.

Theorem 5.2.4. *Let F be a field and $p \in F[x]$ with $p \neq 0_F$.*

(a) *The function*

$$\sigma : \quad F[x] \to F[x]/(p), \quad f \mapsto [f]_p.$$

is an onto homomorphism of rings.

(b) *$F[x]/(p)$ is a commutative ring with identity $[1_F]_p$.*

(c) *Put $\hat{F} = \{[a]_p | a \in F\}$. Then $\hat{F}$ is a subring of $F[x]/(p)$.*

(d) *Suppose $p \notin F$. Then the function*

$$\tau : \quad F \to \hat{F}, \quad a \mapsto [a]_p.$$

is an isomorphism of rings. In particular, $\hat{F}$ is a subring of $F[x]/(p)$ isomorphic to F.

Proof. (a) Let $f, g \in F[x]$. Then

$$\sigma(f+g) = [f+g]_p = [f]_p + [g]_p = \sigma(f) + \sigma(g)$$

and

$$\sigma(fg) = [fg]_p = [f]_p[g]_p = \sigma(f)\sigma(g)$$

So σ is a homomorphism. If $a \in F[x]/(p)$, then $a = [f]_p$ for some $a \in f \in F[x]$. So $\sigma(f) = a$ and σ is onto.

(b) See E.0.3.

(c) $\hat{F} = \{[a]_p \mid a \in F\} = \{\sigma(a) \mid a \in F\}$. Since F is a subring of $F[x]$ and σ is a homomorphism we conclude from Exercise 6 on the Review for Exam 2 that $\hat{F}$ is a subring of $F[x]/(p)$.

(d) We need to show that τ is a 1-1 and onto homomorphism. By (a), σ is a homomorphism. Observe that $\tau(a) = \sigma(a)$ for all $a \in F$. Hence also τ is a homomorphism.

Let $d \in \hat{F}$. Then $d = [a]_p$ for some $a \in F$ and so $d = \tau(a)$. Thus τ is onto.

Let $a, b \in F$ with $\tau(a) = \tau(b)$. Then $[a]_p = [b]_p$. Since $p \notin F$, $\deg p \geq 1$ and since $a, b \in F$, $\deg a \leq 0$ and $\deg b \leq 0$. Thus $\deg a < \deg p$ and $\deg b < \deg p$. Since $[a]_p = [b]_p$ we conclude from 5.1.6 that $a = b$. So τ is 1-1 and (d) holds. □

The preceding theorem shows that $F[x]/(p)$ contains a subring isomorphic to F. This suggest that there exists a ring isomorphic to $F[x]/(p)$ containg F has a subring. The next theorem shows that this is indeed true.

Theorem 5.2.5. *Let F be a field and $p \in F[x]$ with $p \notin F$. Then there exist a ring R and $\alpha \in R$ such that*

(a) *F is a subring of R,*

(b) *there exists an isomorphism $\Phi : R \to F[x]/(p)$ with $\Phi(\alpha) = [x]_p$ and $\Phi(a) = [a]_p$ for all $a \in F$.*

(c) *R is a commutative ring with identity $1_R = 1_F$.*

Proof. Let $S = F[x]/(p) \setminus \hat{F}$ and $R = S \cup F$. (So for $a \in F$ we removed $[a]_p$ from $F[x]/(p)$ and replaced it by a.) Define $\Phi : R \to F[x]/(p)$ by

$$\Phi(r) = [r]_p \text{ if } r \in F \text{ and } \Phi(r) = r \text{ if } r \in S$$

Then its is easy to check that Φ is a bijection. Next we define an addition $\oplus$ and a multiplication $\odot$ on R by

$$r \oplus s = \Phi^{-1}(\Phi(r) + \Phi(s)) \quad \text{and} \quad r \odot s := \Phi^{-1}(\Phi(r)\Phi(s)) \tag{1}$$

Observe that $\Phi(\Phi^{-1}(u)) = u$ for all $u \in F[x]/(p)$. So applying Φ to both sides of (1) gives

$$\Phi(r \oplus s) = \Phi(r) + \Phi(s) \quad \text{and} \quad \Phi(r \odot s) = \Phi(r)\Phi(s)$$

for all $r, s \in R$. E.0.3 implies that R is ring and Φ is an isomorphism. Put $\alpha = [x]_p$. Then $\alpha \in S$ and so $\alpha \in R$. Moreover $\Phi(\alpha) = \Phi([x]_p) = [x]_p$. Let $a \in F$. Then $a \in R$ and $\Phi(a) = [a]_p$. Thus (b) holds.

For $a, b \in F$ we have

$$a \oplus b = \Phi^{-1}(\Phi(a) + \Phi(b)) = \Phi^{-1}([a]_p + [b]_p) = \Phi^{-1}([a + b]_p) = a + b \in F$$

and

$$a \odot b = \Phi^{-1}(\Phi(a)\Phi(b)) = \Phi^{-1}([a]_p[b]_p) = \Phi^{-1}([ab]_p) = ab \in F$$

So F is a subring of R. Thus also (a) is proved.

By 5.2.4 $F[x]/(p)$ is a commutative ring with identity $[1_F]_p$. Since Φ is an isomorphism we conclude that R is a commutative ring with identity 1_F. So (c) holds. □

Remark 5.2.6. *Let R and S be commutative rings with identities. Suppose that S is a subring of R and $1_S = 1_R$. Then we identify the polynomial*

$$f = \sum_{i=0}^{n} f_i x^i \quad \textit{in } S[x]$$

with the polynomial

$$g = \sum_{i=0}^{n} f_i x^i \quad \textit{in } R[x]$$

Note that with this identification, $S[x]$ becomes a subring of $R[x]$. But also note that the functions

$$f^* : \quad S \to S, \quad a \ \mapsto \ \sum_{i=0}^{n} f_i a^i$$

and

$$g^* : \quad R \to R, \quad a \ \mapsto \ \sum_{i=0}^{n} f_i a^i$$

are not the same unless $S \neq R$, since they have different domains. Nevertheless, we use the notation

$$f^*(a) := \sum_{i=0}^{n} f_i a^i.$$

even for $a \in R$.

For example consider

$$f = x^2 + 1 \in \mathbb{Q}[x] \qquad \text{and} \qquad g = x^2 + 1 \in \mathbb{R}[x].$$

Then $f = g$. But the functions

$$f^* : \mathbb{Q} \to \mathbb{Q}, a \to a^2 + 1 \qquad \text{and} \qquad g^* : \mathbb{R} \to \mathbb{R}, a \to a^2 + 1$$

are not the same. But abusing notations we write

$$f^*(\sqrt{2}) = (\sqrt{2})^2 + 1 = 3.$$

Notation 5.2.7. *Let F be a field and $p \in F[x]$ with $p \notin F$. Let R and α be as in 5.2.5. We denote the ring R by $F_p[\alpha]$. (If $F = \mathbb{Z}_q$ for some prime integer q, we will use the notation $\mathbb{Z}_{q,p}[\alpha]$)*

Theorem 5.2.8. *Let F be a field and $p \in F[x]$ with $p \notin F$ and let α and Φ be as in 5.2.5.*

(a) *For all $f \in F[x]$, $\Phi\big(f^*(\alpha)\big) = [f]_p$.*

(b) *Let $f, g \in F[x]$. Then $f^*(\alpha) = g^*(\alpha)$ if and only if $[f]_p = [g]_p$.*

(c) *For each $\beta \in F_p[\alpha]$ there exists a unique $f \in F[x]$ with $\deg f < \deg p$ and $f^*(\alpha) = \beta$.*

(d) *Let $n = \deg p$. Then for each $\beta \in F_p[\alpha]$ there exist unique $b_0, b_1, \ldots, b_{n-1} \in F$ with*

$$\beta = b_0 + b_1\alpha + \ldots + b_{n-1}\alpha^{n-1}.$$

(e) *Let $f \in F[x]$, then $f^*(\alpha) = 0_F$ if and only if $p \mid f$ in $F[x]$.*

(f) *α is a root of p in $F_p[\alpha]$.*

Proof. (a)

$$\Phi\big(f^*(\alpha)\big) = \Phi\left(\sum_{i=0}^{\deg f} f_i\alpha^i\right) = \sum_{i=0}^{\deg f} \Phi(f_i)\Phi(\alpha)^i \overset{5.2.5}{=} \sum_{i=0}^{\deg f} [f_i]_p[x]_p^i = \left[\sum_{i=0}^{\deg f} f_i x^i\right]_p = [f]_p.$$

(b)

$$\begin{array}{ccc} & f^*(\alpha) = g^*(\alpha) & \\ \iff & \Phi(f^*(\alpha)) = \Phi(g^*(\alpha)) & -\ \Phi \text{ is 1-1} \\ \iff & [f]_p = [g]_p & -\ (a) \end{array}$$

(c) Let $\beta \in F_p[\alpha]$ and $f \in F[x]$. Then

$$\begin{array}{ccc} & f^*(\alpha) = \beta & \\ \iff & \Phi(f^*(\alpha)) = \Phi(\beta) & -\ \Phi \text{ is 1-1} \\ \iff & [f]_p = \Phi(\beta) & -\ (a) \end{array}$$

Since $\Phi(\beta) \in F[x]/(p)$, 5.1.6 shows that there exists a unique $f \in F[x]$ with $\deg f < \deg p$ and $[f]_p = \Phi(\beta)$. It follows that f is also the unique $f \in F[x]$ with $\deg f < \deg p$ and $f^*(\alpha) = \beta$. Thus (c) holds.

(d) Let $b_0, \dots b_{n-1} \in \mathbb{F}$ and put $f = b_0 + b_1 + \dots + b_{n-1}x^{n-1}$. Then f is a polynomial with $\deg f < \deg p$ and $b_0, \dots, b_{n-1}$ are uniquely determined by f. Also

$$f^*(\alpha) = b_0 + b_1\alpha + \dots + b_{n-1}\alpha^{n-1}$$

and so (d) follows from (c).

(e)

$$\begin{array}{ccc} & f^*(\alpha) = 0_F & \\ \iff & f^*(\alpha) = 0_F^*(\alpha) & -\ \text{definition of } 0_F^* \\ \iff & [f]_p = [0_F] & -(b) \\ \iff & p \mid f - 0_F & -\ 5.1.5 \\ \iff & p \mid f & -\ 3.2.11(b) \end{array}$$

(f) Since $p \mid p$ this follows from (e). □

Example 5.2.9. Let $p = x^2 + x \in \mathbb{Z}_2[x]$. Determine the addition and multiplication table of $\mathbb{Z}_{2,p}[\alpha]$.

$+$	0	1	α	$\alpha+1$
0	0	1	α	$\alpha+1$
1	1	0	$\alpha+1$	α
α	α	$\alpha+1$	0	1
$\alpha+1$	$\alpha+1$	α	1	0

$\cdot$	0	1	α	$\alpha+1$
0	0	0	0	0
1	0	1	α	$\alpha+1$
α	0	α	α	0
$\alpha+1$	0	$\alpha+1$	0	$\alpha+1$

This can be read of from Example 5.2.3. But it also can be computed from the preceeding theorem: By 5.2.8(d) any element of $F[\alpha]$ can be uniquely written as $b_0 + b_1\alpha$ with $b_0, b_1 \in \mathbb{Z}_2$. By 2.1.2 $\mathbb{Z}_2 = \{0, 1\}$ and so

$$\mathbb{Z}_{2,p}[\alpha] = \{0 + 0\alpha,\ 0 + 1\alpha,\ 1 + 0\alpha,\ 1 + 1\alpha\} = \{0,\ 1,\ \alpha,\ \alpha + 1\}.$$

By 5.2.8(f) $p^*(\alpha) = 0$. So

$$\alpha^2 + \alpha = 0 \qquad \text{and} \qquad \alpha^2 = -\alpha = \alpha.$$

(Note here that $\alpha+\alpha = 2\alpha = 0$ and so $-\alpha = \alpha$.) This allows us to compute the multiplication table, for example

$$(\alpha + 1)(\alpha + 1) = \alpha^2 + \alpha + \alpha + 1 = \alpha^2 + 1 = \alpha + 1.$$

and

$$\alpha(\alpha + 1) = \alpha^2 + \alpha = 0$$

Exercises 5.2:

#1. Write out the addition and multiplication table of $\mathbb{Z}_2[x]/(x^3 + x + 1)$. Is $\mathbb{Z}_2[x]/(x^3 + x + 1)$ a field?

#2. Each element of $\mathbb{Q}[x]/(x^2-3)$ is can be uniquely written in the form $[ax+b]$ (Why?). Determine the rules of addition and multiplication of congruence classes.(In other words, if the product of $[ax + b][cx + d]$ is the class $[rx + c]$ describe how to find r and s from a, b, c, d, and similarly for addition.)

#3. In each part explain why $t \in F[x]/(p)$ is a unit and find its inverse.

$$\begin{array}{lcccl} (a) & t = & [2x-3] & \in & \mathbb{Q}[x]/(x^2-2) \\ (b) & t = & [x^2+x+1] & \in & \mathbb{Z}_3[x]/(x^2+1) \\ (c) & t = & [x^2+x+1] & \in & \mathbb{Z}_2[x]/(x^3+x+1) \end{array}$$

5.3 $F_p[\alpha]$ when p is irreducible

In this section we determine when $F_p[\alpha]$ is a field.

Lemma 5.3.1. *Let F be a field, $p \in F[x]$ with $p \notin F$ and $f \in F[x]$.*

(a) *$f^*(\alpha)$ is a unit in $F_p[\alpha]$ if and only if* $\gcd(f,p) = 1_F$.

(b) *If $1_F = fg + ph$ for some $g, h \in \mathbb{F}[x]$, then $g^*(\alpha)$ is an inverse of $f^*(\alpha)$.*

Proof. (a) We have

$$
\begin{array}{lcl}
 & f^*(\alpha) \text{ is a unit in } F_p[\alpha] & \\
\Longleftrightarrow & f^*(\alpha)\beta = 1_F \text{ for some } \beta \in F_p[\alpha] & -F_p[\alpha] \text{ is commutative, 3.4.10} \\
\Longleftrightarrow & f^*(\alpha)g^*(\alpha) = 1_F \text{ for some } g \in F[x] & -\ 5.2.8(c) \\
\Longleftrightarrow & (fg)^*(\alpha) = 1_F^*(\alpha) \text{ for some } g \in F[x] & -\ 4.4.8 \\
\Longleftrightarrow & [fg]_p = [1_F]_p \text{ for some } g \in F[x] & -\ 5.2.8(b) \\
\Longleftrightarrow & 1_F = fg + ph \text{ for some } g, h \in F[x] & -\ 5.1.5(a)(i) \\
\Longleftrightarrow & \gcd(f,p) = 1_F & -\ 4.2.13
\end{array}
$$

(b) From the above list of equivalent statement, $1_F = fg + ph$ implies $f^*(\alpha)g^*(\alpha) = 1_F$. Since $F_p[\alpha]$ is commutative we also have $g^*(\alpha)f^*(\alpha) = 1_F$ and so $g^*(\alpha)$ is an inverse of $f^*(\alpha)$. □

Proposition 5.3.2. *Let F be a field and $p \in F[x]$ with $p \notin F$. Then the following statements are equivalent:*

(a) *p is irreducible in $F[x]$.*

(b) *$F_p[\alpha]$ is a field.*

(c) *$F_p[\alpha]$ is an integral domain.*

Proof. (a) $\Longrightarrow$ (b): By 5.2.5(c) $F_p[\alpha]$ is a commutative ring with identity 1_F. Suppose p is irreducible and let $\beta \in F_p[\alpha]$ with $\beta \neq 0_F$. By 5.2.8(c), $\beta = f^*(\alpha)$ for some $f \in F[x]$. Then $f^*(\alpha) \neq 0_F$ and 5.2.8(e), gives $p \nmid f$. Since p is irreducible, Exercise 4.3#4 shows that $\gcd(f,p) = 1_F$. Hence by Lemma 5.3.1 $\beta = f^*(\alpha)$ is a unit in $F_p[\alpha]$. Also since F is a field, $1_F \neq 0_F$ and since (by 5.2.5(c)) $1_F = 1_{F_p[\alpha]}$ and $0_F = 0_{F_p[\alpha]}$, all the conditions of a field (see Definition 3.2.20) hold for $F_p[\alpha]$.

(b) $\Longrightarrow$ (c): If $F_p[\alpha]$ is a field, then by Corollary 3.2.22 $F_p[\alpha]$ is an integral domain.

(c) $\Longrightarrow$ (a): Suppose $F_p[\alpha]$ is an integral domain and (for a contradiction) that p is not irreducible. Since $p \notin F$, 4.3.2 shows that $p = gh$ where g and h are non constant polynomials of degree less than p. Since $g \neq 0_F$ and both g and 0_F have degree less than p, 5.2.8(c) shows that $g^*(\alpha) \neq 0^*_F(\alpha)$. As $0^*_F(\alpha) = 0_F$ this gives $g^*(\alpha) \neq 0_F$. Similarly, $h^*(\alpha) \neq 0_F$. But

$$g^*(\alpha)h^*(\alpha) = (gh)^*(\alpha) = p^*(\alpha) = 0_F$$

a contradiction since by definition (Ax 11) holds in integral domains (see 3.2.18). □

Corollary 5.3.3. *Let F be a field and p an irreducible polynomial in $F[x]$. Then $F_p[\alpha]$ is a field containing F as subring, and α is a root of p in $F_p[\alpha]$.*

Proof. By 5.2.5 F is a subring of $F_p[\alpha]$. Since p is irreducible, 5.3.2 implies that $F_p[\alpha]$ is field. By 5.2.8 α is a root of p in $F_p[\alpha]$. □

Example 5.3.4. Put $K := \mathbb{R}_{x^2+1}[\alpha]$. Determine the addition and multiplication in K and show that K is a field.

By 5.2.8(f) we know that α is a root of x^2+1 in K. Hence $\alpha^2 + 1 = 0$ and so

$$\alpha^2 = -1.$$

By 5.2.8, every element of K can be uniquely written as $a + b\alpha$ with $a, b \in \mathbb{R}$. We have

$$(a + b\alpha) + (c + d\alpha) = (a + c) + (b + d)\alpha$$

and

$$(a + b\alpha)(c + d\alpha) = ac + (bc + ad)\alpha + bd\alpha^2 = ac + (bc + ad)\alpha + bd(-1) = (ac - bd) + (ad + bc)\alpha.$$

Suppose that $a + \beta\alpha \neq 0$. Then $\alpha \neq 0$ or $b \neq 0$ and so $a^2 + b^2 > 0$. Also

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

Hence $a + b\alpha$ is a unit in K and so K is a field.

We remark that is now straight forward to verify that

$$\phi : \mathbb{R}_{x^2+1}[\alpha] \to \mathbb{C}, \quad a + b\alpha \mapsto a + bi$$

is an isomorphism from $\mathbb{R}_{x^2+1}[\alpha]$ to the complex numbers $\mathbb{C}$.

Corollary 5.3.5. *Let F be a field and $f \in F[x]$.*

(a) *Suppose $f \notin F$. Then there exists a field K with F as a subring such that f has a root in K.*

(b) *There exist a field L with F as a subring, $n \in \mathbb{N}$, and elements $c, a_1, a_2 \ldots, a_n$ in L such that*

$$f = c \cdot (x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_n)$$

Proof. (a) By 4.3.8(a), f is a product of irreducible polynomials. In particular, there exists an irreducible polynomial p in $F[x]$ dividing f. By 5.3.3 $K = F_p[\alpha]$ is a field containing F and α is a root of p in K. Since $p|f$, 4.4.12 shows that α is a root of f in K.

(b) We will prove (b) by induction on $\deg f$. If $\deg f \leq 0$, then $f \in F$. So (b) holds with $n = 0, c = f$ and $L = F$. Suppose that $k \in \mathbb{N}$ and (b) holds for any field F and any polynomial of degree k in $F[x]$. Let f be a polynomial of degree $k+1$ in $F[x]$. Then $\deg f \geq 1$ and so by (a) there exists a field K with F as a subring and a root α of f in K. By the Factor Theorem 4.4.11 $x - \alpha$ divides f in $K[x]$ and so $f = g \cdot (x - \alpha)$ for some $g \in K[x]$. Thus $\deg g = k$ and so by the induction assumption, there exists a field L with K as a subring and elements $c, a_1, \ldots a_k$ in L with

$$g = c \cdot (x - a_1) \cdot \ldots \cdot (x - a_k).$$

Put $a_{k+1} = \alpha$. Then

$$f = g \cdot (x - \alpha) = c \cdot (x - a_1) \cdot \ldots \cdot (x - a_k) \cdot (x - a_{k+1}).$$

Since F is a subring of K and K is subring of L, F is subring of L. So (b) holds for polynomials of degree $k+1$. Hence, by the Principal of Mathematical Induction, (b) holds for polynomials of arbitrary degree. □

Exercises 5.3:

#1. Determine which of the following congruence-class rings are fields.

(a) $\mathbb{Z}_3[x]/(x^3 + 2x^2 + x + 1)$.

(b) $\mathbb{Z}_5[x]/(2x^3 - 4x^2 + 2x + 1)$.

(c) $\mathbb{Z}_2[x]/(x^4 + x^2 + 1)$.

#2. (a) Verify that $\mathbb{Q}(\sqrt{3}) := \{r + s\sqrt{3} | r, s \in \mathbb{Q}\}$ is a subfield of $\mathbb{R}$.

(b) Show that $\mathbb{Q}(\sqrt{3})$ is isomorphic to $\mathbb{Q}[x]/(x^2 - 3)$.

#3. (a) Show that $\mathbb{Z}_2[x]/(x^3 + x + 1)$ is a field.

(b) Show that $x^3 + x + 1$ has three distinct roots in $\mathbb{Z}_2[x]/(x^3 + x + 1)$.

Chapter 6

Ideals and Quotients

6.1 Ideals

Definition 6.1.1. *Let I be a subset of the ring R.*

(a) *We say that I* absorbs *R if*

$$ra \in I \quad \textit{and} \quad ar \in I \qquad \textit{for all } a \in I, r \in R$$

(b) *We say that I is an* ideal *of R if I is a subring of R and I absorbs R.*

Theorem 6.1.2 (Ideal Theorem)**.** *Let I be a subset of the ring R. Then I is an ideal in R if and only if the following four conditions holds:*

(i) $0_R \in I$.

(ii) $a + b \in I$ *for all* $a, b \in I$.

(iii) $ra \in I$ *and* $ar \in I$ *for all* $a \in I$ *and* $r \in R$.

(iv) $-a \in I$ *for all* $a \in I$.

Proof. $\Longrightarrow$: Suppose first that I is an ideal in R. By Definition 6.1.1 S absorbs R and S is a subring. Thus (iii) holds and by the Subring Theorem 3.2.8 also (i), (ii) and (iv) hold.

$\Longleftarrow$: Suppose that (ii)-(iv) hold. (iii) implies $ab \in I$ for all $a, b \in I$. So the Subring Theorem 3.2.8 shows that I is a subring of R. By (iii), I absorbs R and so I is an ideal in R. $\square$

Example 6.1.3. (1) $\{3n \mid n \in \mathbb{Z}^+\}$ is an ideal in $\mathbb{Z}$.

(2) Let F be a field and $a \in F$. Then $\{f \in F[x] \mid f^*(a) = 0_F\}$ is an ideal in $F[x]$.

(3) Let R be a ring, I an ideal in R. Then $\{f \in R[x] \mid f_i \in I \text{ for all } i \in \mathbb{N}\}$ is an ideal in $R[x]$.

(4) Let R and S be rings. Then $R \times \{0_S\}$ is an ideal in $R \times S$.

Proof. See Exercise #1 $\square$

Definition 6.1.4. *Let R be a ring.*

(a) *Let $a \in R$. Then $aR = \{ar \mid a \in R\}$.*

(b) *Suppose R is commutative and $I \subseteq R$. Then I is called a principal ideal in R if $I = aR$ for some $a \in R$.*

Lemma 6.1.5. *Let R be a commutative ring with identity and $a \in R$. Then aR is the smallest ideal in R containing a, that is*

(a) *$a \in aR$,*

(b) *aR is an ideal in R, and*

(c) *$aR \subseteq I$, whenever I is an ideal in R with $a \in I$.*

Proof. (a): Note that $a = a \cdot 1_R$ and so $a \in aR$.

(b) Let $b, c \in aR$ and $r \in R$. Then

$$b = as \qquad \text{and} \qquad c = at.$$

for some $s, t \in R$. Thus

$$\begin{gathered}
0_R = a0_R \in aR, \\
b + c = as + at = a(s+t) \in aR, \\
rb = br = (as)r = a(sr) \in aR \\
-b = -(as) = a(-s) \in aR.
\end{gathered}$$

So by 6.1.2 aR is an ideal in R.

(c): Let I be any ideal of R containing a. Since $a \in I$ and I absorbs R, $ar \in I$ for all $r \in R$ and so $aR \subseteq I$. $\square$

Definition 6.1.6. *Let I be an ideal in the ring R. The relation '$\equiv \pmod{I}$' on R is defined by*

$$a \equiv b \pmod{I} \qquad \textit{if} \qquad a - b \in I$$

Remark 6.1.7. (a) *Let $a, b, n \in \mathbb{Z}$. Then*

$$a \equiv b \pmod{n} \quad \Longleftrightarrow \quad a \equiv b \pmod{n\mathbb{Z}}$$

(b) *Let F be a field and $f, g, p \in F[x]$ with $p \neq 0_F$. Then*

$$f \equiv g \pmod{p} \iff f \equiv g \pmod{pF[x]}$$

Proof. We will prove(b). The proof for (a) is virtually the same.

$$\begin{array}{rcl} & f \equiv g \pmod{p} & \\ \iff & f - g = pk \text{ for some } k \in F[x] & -\ 5.1.5 \\ \iff & f - g \in pF[x] & -\text{Definition of } pF[x] \\ \iff & f \equiv g \pmod{pF[x]} & -\text{Definition of } \equiv \pmod{I}\ \ 6.1.10 \end{array}$$

□

Proposition 6.1.8. *Let I be an ideal in R. Then '$\equiv \pmod{I}$' is an equivalence relation on R.*

Proof. We need to show that '$\equiv \pmod{I}$' is reflexive, symmetric and transitive. Let $a, b, c \in R$.

Reflexive By 3.2.11 $a - a = 0_R$ and by the Ideal Theorem $0_R \in I$. Thus $a - a \in I$ and so $a \equiv a \pmod{I}$ by definition of '$\equiv \pmod{I}$'.

Symmetric Suppose $a \equiv b \pmod{I}$. Then $a - b \in I$ and so by Ideal Theorem $-(a - b) \in I$. By 3.2.11 $b - a = -(a - b)$. Hence $b - a \in I$ and so $b \equiv a \pmod{I}$ by definition of '$\equiv \pmod{I}$'.

Transitive Suppose $a \equiv b \pmod{I}$ and $b \equiv c \pmod{I}$, then $a - b \in I$ and $b - c \in I$. Hence by the Ideal Theorem $(a - b) + (b - c) \in I$. As $a - c = (a - b) + (b - c)$ this gives $a - c \in I$. Thus $a \equiv c \pmod{I}$. □

Definition 6.1.9. *Let R be a ring and I an ideal in R.*

(a) *Let $a \in I$. Then $a + I$ denotes the equivalence class of '$\equiv \pmod{I}$' containing a. So*

$$a + I = \{b \in R \mid a \equiv b \pmod{I}\} = \{b \in R \mid a - b \in I\}$$

$a + I$ is called the coset *of I in R containing a.*

(b) *R/I is the set of cosets of I in R/I. So*

$$R/I = \{a + I \mid a \in R\}$$

and R/I is the set of equivalence classes of '$\equiv \pmod{I}$'

Theorem 6.1.10. *Let R be ring and I an ideal in R. Let $a, b \in R$. Then the following statements are equivalent*

(a) *$a = b + i$ for some $i \in I$.*

(b) *$a - b = i$ for some $i \in I$*

(c) *$a - b \in I$.*

(d) *$a \equiv b \pmod I$.*

(e) *$b \in a + I$.*

(f) *$(a + I) \cap (b + I) \neq \emptyset$.*

(g) *$a + I = b + I$.*

(h) *$a \in b + I$.*

(i) *$b \equiv a \pmod I$.*

(j) *$b - a \in I$.*

(k) *$b - a = j$ for some $j \in I$.*

(l) *$b = a + j$ for some $j \in I$.*

Proof. (a) $\iff$ (b): and (k) $\iff$ (l): This holds by 3.2.12.

(b) $\iff$ (c): and (j) $\iff$ (k): Obvious.

(c) $\iff$ (d): and (i) $\iff$ (j): This holds by definition of '$\equiv \pmod I$'.

By 6.1.8 we know that '$\equiv \pmod I$) is an equivalence relation. Also $a + I$ is the equivalence class of a and so Theorem 0.5.8 implies that (d)-(i) are equivalent. □

Corollary 6.1.11. *Let I be an ideal in the ring R.*

(a) *Let $a \in R$. Then $a + I = \{a + i \mid i \in I\}$.*

(b) *$0_R + I = I$. In particular, I is a coset of I in R.*

(c) *Any two cosets of I are either disjoint or equal.*

Proof. Let $a, b \in R$.

(a) By 6.1.10(a),(h) we have $b \in a + I$ if and only if $b = a + i$ for some $i \in I$ and so if and only if $b \in \{a + i \mid i \in I\}$.

(b) By (a) $0_R + I = \{0_r + i \mid i \in I\} = \{i \mid i \in I\} = I$.

(c) Suppose $a + I$ and $b + I$ are not disjoint. Then $(a + I) \cap (b + I) \neq \emptyset$ and 6.1.10(f),(g) shows that $a + I = b + I$. So two cosets of I in R are either disjoint or equal. □

Exercises 6.1:

#1. Show that:

(a) $\{3n \mid n \in \mathbb{Z}^+\}$ is an ideal in $\mathbb{Z}$.

(b) Let F be a field and $a \in F$. Then $\{f \in F[x] \mid f^*(a) = 0_F\}$ is an ideal in $F[x]$.

(c) Let R be a ring, I an ideal in R. Then $\{f \in R[x] \mid f_i \in I \text{ for all } i \in \mathbb{N}\}$ is an ideal in R.

(d) Let R and S be rings. Then $R \times \{0_S\}$ is an ideal in $R \times S$.

#2. Let $I_1, I_2, \dots I_n$ be ideals in the ring R. Show that $I_1 + I_2 + \dots + I_n$ is the smallest ideal in R containing $I_1, I_2, \dots, I_n$ and I_n.

#3. Is the set $J = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & r \end{bmatrix} \middle| r \in \mathbb{R} \right\}$ an ideal in the ring $\mathrm{M}_2(\mathbb{R})$ of 2×2 matrices over $\mathbb{R}$?

#4. If I is an ideal in the ring R and J is an ideal in the ring S, prove that $I \times J$ is an ideal in the ring $R \times S$.

#5. Let F be a field and I an ideal in $F[x]$. Show that I is a principal ideal. *Hint:* If $I \neq \{0_F\}$ choose $d \in I$ with $d \neq 0_F$ and $\deg(d)$ minimal. Show that $I = F[x]d$.

#6. Let $\Phi : R \to S$ be a homomorphism of rings and let J be an ideal in S. Put $I = \{a \in R \mid \Phi(a) \in J\}$. Show that I is an ideal in R.

6.2 Quotient Rings

Proposition 6.2.1. *Let I be an ideal in R and $a, b, \tilde{a}, \tilde{b} \in R$ with*

$$a + I = \tilde{a} + I \qquad \textit{and} \qquad b + I = \tilde{b} + I.$$

Then

$$(a + b) + I = (\tilde{a} + \tilde{b}) + I \qquad \textit{and} \qquad ab + I = \tilde{a}\tilde{b} + I.$$

Proof. Since $a + I = \tilde{a} + I$ 6.1.10 implies that $\tilde{a} = a + i$ for some $i \in I$. Similarly $\tilde{b} = b + j$ for some $j \in I$.

Thus

$$\tilde{a} + \tilde{b} = (a + i) + (b + j) = (a + b) + (i + j).$$

Since $i, j \in I$ and I is closed under addition, $i + j \in I$ and so by 6.1.10 $(a + b) + I = (\tilde{a} + \tilde{b}) + I$.

Also

$$\tilde{a}\tilde{b} = (a + i)(b + j) = ab + (aj + ib + ij)$$

Since $i, j \in I$ and I absorbs R we conclude that aj, ib and ij all are in I. Since I is closed under addition this implies that $aj + ib + ij \in I$ and so $ab + I = \tilde{a}\tilde{b} + I$ by 6.1.10. □

Definition 6.2.2. *Let I be an ideal in the ring R. Then we define an addition $+$ and multiplication $\cdot$ on R by*

$$(a + I) + (b + I) = (a + b) + I \quad \textit{and} \quad (a + I) \cdot (b + I) = ab + I$$

for all $a, b \in R$.

Note that by the preceding proposition the addition and multiplication on R/I are well defined.

Remark 6.2.3. (a) *Let $n \in \mathbb{Z}$. Then $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.*

(b) *Let F be a field and $p \in F[x]$. Then $F[x]/(p) = F[x]/pF[x]$.*

Proof. This follows from Remark 6.1.7 $\square$

Theorem 6.2.4. *Let R be a ring and I an ideal in R*

(a) *The function $\pi : R \to R/I$, $a \to a + I$ is an onto homomorphism.*

(b) *$(R/I, +, \cdot)$ is a ring.*

(c) *$0_{R/I} = 0_R + I = I$.*

(d) *If R is commutative, then R/I is commutative.*

(e) *If R has an identity, then R/I has an identity and $1_{R/I} = 1_R + I$.*

Proof. (a) Let $a, b \in R$. Then

$$\pi(a+b) \overset{\text{Def } \pi}{=} (a+b) + I \overset{\text{Def } +}{=} (a+I) + (b+I) \overset{\text{Def } \pi}{=} \pi(a) + \pi(b)$$

and

$$\pi(ab) \overset{\text{Def } \pi}{=} ab + I \overset{\text{Def } \cdot}{=} (a+I)(b+I) \overset{\text{Def } \pi}{=} \pi(a)\pi(b)$$

So π is a homomorphism. Let $u \in R/I$. By definition, $R/I = \{a + I \mid a \in R\}$ and so there exists $a \in R$ with $u = a + I$. Thus $\pi(a) = a + I = u$ and so π is onto.

(b), (c) and (d): By (a) π is an onto homomorphism. Thus we can apply E.0.3 and conclude that (b),(c) and (d) hold.

(e): By (a) π is an onto homomorphism. Thus (e) follows from 3.3.7(d). $\square$

Lemma 6.2.5. *Let R be a ring and I an ideal in R. Let $r \in R$. Then the following statements are equivalent:*

(a) *$r \in I$.*

(b) *$r + I = I$.*

(c) *$r + I = 0_{R/I}$.*

Proof. (a) $\iff$ (b): By 6.1.10 $r \in 0_R + I$ if and only of $r + I = 0_R + I$. By 6.2.4(c) $0_R + I = I$ and so (a) and (b) are equivalent.

(b) $\iff$ (c): By 6.2.4(c) $0_{R/I} = I$ and so (b) and (c) are equivalent. $\square$

Definition 6.2.6. (a) *Let $f : R \to S$ be a homomorphism of rings. Then*

$$\ker f = \{a \in R \mid f(a) = 0_R\}.$$

$\ker f$ *is called the* kernel *of f.*

(b) *Let I be an ideal in the ring R. The function*

$$\pi : \quad R \to R/I, \quad r \to r+I$$

is called the natural homomorphism *from R to R/I.*

Lemma 6.2.7. *Let $f : R \to S$ be homomorphism of rings. Then $\ker f$ is an ideal in R.*

Proof. By definition, $\ker f$ is a subset of R. We will now verify the four conditions of the Ideal Theorem 6.1.2. So let $a, b \in \ker f$ and $r \in R$. By definition of $\ker f$,

$$(*) \qquad f(a) = 0_S \quad \text{and} \quad f(b) = 0_S.$$

(i) $f(a+b) \overset{\text{f hom}}{=} f(a)+f(b) \overset{(*)}{=} 0_S + 0_S \overset{\text{Ax 4}}{=} 0_S$ and so $a+b \in \ker f$ by definition of $\ker f$.

(ii) $f(ra) \overset{\text{f hom}}{=} f(r)f(a) \overset{(*)}{=} f(r)0_S \overset{3.2.11(c)}{=} 0_S$ and so $ra \in \ker f$ by definition of $\ker f$. Similarly, $ar \in \ker f$.

(iii) $f(0_R) \overset{3.3.7(a)}{=} 0_S$ and so $0_R \in \ker f$ by definition of $\ker f$.

(iv) $f(-a) \overset{3.3.7(b)}{=} -f(a) \overset{(*)}{=} -0_S \overset{3.2.11(a)}{=} 0_S$ and so $-a \in \ker f$ by definition of $\ker f$. □

Example 6.2.8. Define

$$\Phi : \quad \mathbb{R}[x] \to \mathbb{C}, \qquad f \mapsto f^*(i)$$

Verify that Φ is a homomorphism and compute $\ker \Phi$.

Define $\rho : \mathbb{R} \to \mathbb{C}, r \to r$. Then ρ is a homomorphism and Φ is the function ρ_i from Lemma 4.4.1. So Φ is a homomorphism. s

Let $f \in F[x]$. We need to determine when $f^*(i) = 0$. According to the Division algorithm, $f = (x^2+1) \cdot q + r$, where $q, r \in \mathbb{R}[x]$ with $\deg(r) < \deg(x^2+1) = 2$. Then $r = a + bx$ for some $a, b \in \mathbb{R}$ and so

$$(*) \qquad f^*(i) = \Big((x^2+1) \cdot q + r\Big)^*(i) = (i^2+1) \cdot q^*(i) + r^*(i) = 0 \cdot q^*(i) + (a+bi) = a+bi$$

It follows that

$$\begin{array}{lll}
 & f \in \ker \Phi & \\
\iff & \Phi(f) = 0 & -\text{ definition of } \ker \Phi \\
\iff & f^*(i) = 0 & -\text{ definition of } \Phi \\
\iff & a + bi = 0 & -\ (*) \\
\iff & a = 0 \text{ and } b = 0 & -\text{ Property of } \mathbb{C} \\
\iff & a + bx = 0 & -\text{ definition of polynomial ring} \\
\iff & r = 0 & -\ r = a + bx \\
\iff & f = (x^2+1) \cdot q \text{ for some } q \in \mathbb{R}[x] & -\text{ Division algorithm} \\
\iff & f \in (x^2+1)\mathbb{R}[x] & -\text{ Definition of } (x^2+1)\mathbb{R}[x]
\end{array}$$

Thus $\ker \Phi = (x^2+1)\mathbb{R}[x]$.

Lemma 6.2.9. *Let R be a ring, I an ideal in R and $\pi : R \to R/I, a \to a + I$ the* natural homomorphism *from R to I. Then* $\ker \pi = I$. *In particular, a subset of I is an ideal in R if and only if it is the kernel of a ring homomorphism with domain R.*

Proof. Let $r \in R$. Then

$$\begin{array}{ccl} & r \in \ker \pi & \\ \iff & \pi(r) = 0_{R/I} & - \text{definition of } \ker \pi \\ \iff & r + I = 0_{R/I} & - \text{definition of } \pi \\ \iff & r \in I & - 6.2.5 \end{array}$$

Thus $\ker \pi = I$. □

Lemma 6.2.10. *Let $f : R \to S$ be a ring homomorphism.*

(a) *Let $a, b \in R$. Then*

$$\begin{array}{cc} & f(a) = f(b) \\ \iff & a - b \in \ker f \\ \iff & a + \ker f = b + \ker f \end{array}$$

(b) *f is 1-1 if and only if* $\ker f = \{0_R\}$.

Proof. (a)

$$\begin{array}{ccccl} & f(a) & = & f(b) & \\ \iff & f(a) - f(b) & = & 0_S & - 3.2.11(f) \\ \iff & f(a - b) & = & 0_S & - 3.3.7(c) \\ \iff & & a - b \in \ker f & & - \text{Definition of } \ker f \\ \iff & a + \ker f & = & b + \ker f & - 6.1.10 \end{array}$$

(b) $\Longrightarrow$: Suppose f is 1-1 and let $a \in R$. Then

$$\begin{array}{ccl} & a \in \ker f & \\ \iff & f(a) = 0_S & - \text{ Definition of } \ker f \\ \iff & f(a) = f(0_R) & - \text{ } 3.3.7(a) \\ \iff & a = 0_R & - f \text{ is 1-1} \end{array}$$

Thus $\ker f = \{0_R\}$.

$\Longleftarrow$: Suppose $\ker f = \{0_R\}$ and let $a, b \in R$ with $f(a) = f(b)$. Then by (b) $a - b \in \ker f$. As $\ker f = \{0_R\}$ this gives $a - b = 0_R$, so $a = b$ by 3.2.11(f). Hence f is 1-1. □

Theorem 6.2.11 (First Isomorphism Theorem)**.** *Let $f : R \to S$ be a ring homomorphism. Recall that* $\operatorname{Im} f = \{f(a) \mid a \in R\}$*. The function*

$$\overline{f}: \quad R/\ker f \to \operatorname{Im} f, \quad (a + \ker f) \mapsto f(a)$$

is a well-defined ring isomorphism. In particular $R/\ker f$ *and* $\operatorname{Im} f$ *are isomorphic rings*

Proof. By 6.2.10 $f(a) = f(b)$ if and only if $a + \ker f = b + \ker f$. The forward direction shows that $\overline{f}$ is 1-1 and backwards direction shows that $\overline{f}$ is well-defined. If $s \in \operatorname{Im} f$, then $s = f(a)$ for some $a \in R$ and so $\overline{f}(a + \ker f) = f(a) = s$. Hence $\overline{f}$ is onto. It remains to verify that $\overline{f}$ is a ring homomorphism. We compute

$$\begin{aligned} \overline{f}\Big((a + \ker f) + (b + \ker f)\Big) &\overset{\text{Def } +}{=} \overline{f}\Big((a+b) + \ker f\Big) \overset{\text{Def } \overline{f}}{=} f(a+b) \\ &\overset{f \text{ hom}}{=} f(a) + f(b) \overset{\text{Def } \overline{f}}{=} \overline{f}(a + \ker f) + \overline{f}(b + \ker f) \end{aligned}$$

and

$$\begin{aligned} \overline{f}\Big((a + \ker f) \cdot (b + \ker f)\Big) &\overset{\text{Def } \cdot}{=} \overline{f}\Big(ab + \ker f\Big) \overset{\text{Def } \overline{f}}{=} f(ab) \\ &\overset{f \text{ hom}}{=} f(a) \cdot f(b) \overset{\text{Def } \overline{f}}{=} \overline{f}(a + \ker f) \cdot \overline{f}(b + \ker f) \end{aligned}$$

and so $\overline{f}$ is a homomorphism. $\square$

Example 6.2.12. Let n and m be non-zero integers with $\gcd(n, m) = 1$. Apply the isomorphism theorem to the homomorphism

$$f : \mathbb{Z} \to \mathbb{Z}_n \times \mathbb{Z}_m, \quad a \to \big([a]_n, [b]_m)\big)$$

We first compute $\ker f$

$$\begin{array}{lcl} & a \in \ker f & \\ \iff & f(a) = 0_{\mathbb{Z}_n \times \mathbb{Z}_m} & -\text{ definition of } \ker \pi \\ \iff & \big([a]_n, [b]_m\big) = ([0]_n, [0]_m) & -\text{ definition of } f \\ \iff & [a]_n = [0]_n \quad \text{and} \quad [b]_m = [0]_m & -\ 0.3.2 \\ \iff & n|a \quad \text{and} \quad m|a & -\ 2.3.1 \\ \iff & nm|a & -\ \gcd(n,m) = 1, \text{Exercise 1.2\#2} \\ \iff & a = nmk \quad \text{for some} \quad k \in \mathbb{Z} & -\text{ definition of 'divide'} \\ \iff & a \in nm\mathbb{Z} & -\text{ definition of } nm\mathbb{Z} \end{array}$$

Thus $\ker f = nm\mathbb{Z}$ and so

$$\mathbb{Z}/\ker f = \mathbb{Z}/nm\mathbb{Z} = \mathbb{Z}_{nm}$$

where the last equality holds by 6.2.3(a).

By the First Isomorphism Theorem $\mathbb{Z}/\ker f$ is isomorphic to $\operatorname{Im} f$ and so

$$\mathbb{Z}_{nm} \quad \text{is isomorphic to} \quad \operatorname{Im} f.$$

Thus

$$|\operatorname{Im} f| = |\mathbb{Z}_{nm}| = nm.$$

Also

$$|\mathbb{Z}_n \times \mathbb{Z}_m| = |\mathbb{Z}_n| \cdot |\mathbb{Z}_m| = nm.$$

Hence $|\operatorname{Im} f| = |\mathbb{Z}_n \times \mathbb{Z}_m|$. Since $\operatorname{Im} f \subseteq \mathbb{Z}_n \times \mathbb{Z}_m$ this gives $\operatorname{Im} f = \mathbb{Z}_n \times \mathbb{Z}_m$. This gives

$$\mathbb{Z}_{nm} \quad \text{is isomorphic to} \quad \mathbb{Z}_n \times \mathbb{Z}_m.$$

Appendix A

Logic

A.1 Rules of Logic

In the following we collect a few statements which are always true.

Lemma A.1.1. *Let P, Q and R be statements, let T be a true statement and F a false statement. Then each of the following statements holds.*

(LR 1) $F \Longrightarrow P$.

(LR 2) $P \Longrightarrow T$.

(LR 3) not -(not -P) $\Longleftrightarrow P$.

(LR 4) (not -$P \Longrightarrow F$) $\Longrightarrow P$.

(LR 5) P or T.

(LR 6) not -(P and F).

(LR 7) (P and T) $\Longleftrightarrow P$.

(LR 8) (P or F) $\Longleftrightarrow P$.

(LR 9) (P and P) $\Longleftrightarrow P$.

(LR 10) (P or P) $\Longleftrightarrow P$.

(LR 11) P or not -P.

(LR 12) not -(P and not -P).

(LR 13) (P and Q) $\Longleftrightarrow$ (Q and P).

(LR 14) (P or Q) $\Longleftrightarrow$ (Q or P).

(LR 15) $(P \iff Q) \iff \Big((P \text{ and } Q) \text{ or} \,(\text{not -}P \text{ and not -}Q)\Big)$

(LR 16) $(P \implies Q) \iff (\text{not -}P \text{ or } Q)$.

(LR 17) $\text{not -}(P \implies Q) \iff (P \text{ and not -}Q)$.

(LR 18) $\Big(P \text{ and} \,(P \implies Q)\Big) \implies Q$.

(LR 19) $\Big((P \implies Q) \text{ and} \,(Q \implies P)\Big) \iff (P \iff Q)$.

(LR 20) $(P \implies Q) \iff (\text{not -}Q \implies \text{not -}P)$

(LR 21) $(P \iff Q) \iff (\text{not -}P \iff \text{not -}Q)$.

(LR 22) $\text{not -}(P \text{ and } Q) \iff (\text{not -}P \text{ or not -}Q)$

(LR 23) $\text{not -}(P \text{ or } Q) \iff (\text{not -}P \text{ and not -}Q)$

(LR 24) $\Big((P \text{ and } Q) \text{ and } R\Big) \iff \Big(P \text{ and} \,(Q \text{ and } R)\Big)$.

(LR 25) $\Big((P \text{ or } Q) \text{ or } R\Big) \iff \Big(P \text{ or} \,(Q \text{ or } R)\Big)$.

(LR 26) $\Big((P \text{ and } Q) \text{ or } R\Big) \iff \Big((P \text{ or } R) \text{ and} \,(Q \text{ or } R)\Big)$.

(LR 27) $\Big((P \text{ or } Q) \text{ and } R\Big) \iff \Big((P \text{ and } R) \text{ or} \,(Q \text{ and } R)\Big)$.

(LR 28) $\Big((P \implies Q) \text{ and} \,(Q \implies R)\Big) \implies (P \implies R)$

(LR 29) $\Big((P \iff Q) \text{ and} \,(Q \iff R)\Big) \implies (P \iff R)$

Proof. If any of these statements are not evident to you, you should use a truth table to verify it. □

Appendix B

Relations, Functions and Partitions

B.1 The inverse of a function

Definition B.1.1. *Let $f : A \to B$ and $g : B \to A$ be functions.*

(a) *g is called a left inverse of f if $g \circ f = \mathrm{id}_A$.*

(b) *g is called a right inverse of g if $f \circ g = \mathrm{id}_B$.*

(c) *g is a called an inverse of f if $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.*

Lemma B.1.2. *Let $f : A \to B$ and $h : B \to A$ be functions. Then the following statements are equivalent.*

(a) *g is a left inverse of f.*

(b) *f is a right inverse of g.*

(c) *$g(f(a)) = a$ for all $a \in A$.*

(d) *For all $a \in A$ and $b \in B$:*

$$f(a) = b \quad \Longrightarrow \quad a = g(b)$$

Proof. (a) $\Longrightarrow$ (b): Suppose that g is a left inverse of f. Then $g \circ f = \mathrm{id}_A$ and so f is a right inverse of g.

(b) $\Longrightarrow$ (c): Suppose that f is a right inverse of g. Then by definition of 'right inverse'

(1) $$g \circ f = \mathrm{id}_A$$

Let $a \in A$. Then

$$\begin{aligned} g\big(f(a)\big) &= (g \circ f)(a) && - \text{definition of composition} \\ &= \mathrm{id}_A(a) && -(1) \\ &= a && - \text{definition of } \mathrm{id}_A \end{aligned}$$

(c) $\Longrightarrow$ (d): Suppose that $g(f(a)) = a$ for all $a \in A$. Let $a \in A$ and $b \in B$ with $f(a) = b$. Then by the principal of substitution $g(f(a)) = g(b)$, and since $g(f(a)) = a$, we get $a = g(b)$.

(d) $\Longrightarrow$ (a): Suppose that for all $a \in A, b \in B$:

$$f(a) = b \Longrightarrow a = g(b) \tag{2)}$$

Let $a \in A$ and put

$$b = f(a) \tag{3}$$

Then by (2)

$$a = g(b) \tag{4}$$

and so

$$\begin{aligned} (g \circ f)(a) &= g\big(f(a)\big) && - \text{definition of composition} \\ &= g(b) && (3) \\ &= a && (4) \\ &= \mathrm{id}_A(a) && - \text{definition of } \mathrm{id}_A \end{aligned}$$

Thus by 0.3.11 $g \circ f = \mathrm{id}_A$. Hence g is a left inverse of f. □

Lemma B.1.3. *Let $f : A \to B$ and $h : B \to A$ be functions. Then the following statements are equivalent.*

(a) *g is an inverse of f.*

(b) *f is a inverse of g.*

(c) *$g(fa) = a$ for all $a \in A$ and $f(gb) = b$ for all $b \in A$.*

(d) *For all $a \in A$ and $b \in B$:*

$$fa = b \quad \Longleftrightarrow \quad a = gb$$

Proof. Note that g is an inverse of f if and only if g is a left and a right inverse of f. Thus the lemma follows from B.1.2 □

Theorem B.1.4. *Let $f : A \to B$ be a function and suppose $A \neq \emptyset$.*

(a) *f is 1-1 if and only if f has a right inverse.*

(b) *f is onto if and only if f has left inverse.*

(c) *f is a 1-1 correspondence if and only f has inverse.*

Proof. $\Longrightarrow$: Since A is not empty we can fix an element $a_0 \in A$. Let $b \in B$. If $b \in \operatorname{Im} f$ choose $a_b \in A$ with $fa_b = b$. If $b \notin \operatorname{Im} f$, put $a_b = a_0$. Define

$$g : B \to A, \quad b \to a_b$$

(a) Suppose f is 1-1. Let $a \in A$ and $b \in B$ with $b = fa$. Then $b \in \operatorname{Im} f$ and $fa_b = b = fa$. Since f is 1-1, we conclude that $a_b = b$ and so $ga = a_b = b$. Thus by B.1.2, g is right inverse of f.

(b) Suppose f is onto. Let $a \in A$ and $b \in B$ with $gb = a$. Then $a = a_b$. Since f is onto, $B = \operatorname{Im} f$ and so $a \in \operatorname{Im} f$ and $f(a_b) = b$. Hence $fa = b$ and so by B.1.2 (with the roles of f and f interchanged), g is left inverse of f.

(c) Suppose f is a 1-1 correspondence. Then f is 1-1 and onto and so by the proof of (a) and (b), g is left and right inverse of f. So g is an inverse of f.

$\Longleftarrow$:

(a) Suppose g is a left inverse of f and let $a, c \in A$ with $fa = fc$. Then by the principal of substitution, $g(fa) = g(fc)$. By B.1.2 $g(fa) = a$ and $g(fb) = b$. So $a = b$ and f -s 1-1.

(b) Suppose g is a right inverse of f and let $b \in B$. Then by B.1.2, $f(gb) = b$ and so f is onto.

(c) Suppose f has an inverse. Then f has a left and a right inverse and so by (a) and (b), f is 1-1 and onto. So f is a 1-1 correspondence. $\square$

B.2 Partitions

Definition B.2.1. *Let A be a set and Δ set of non-empty subsets of A.*

(a) *Δ is called a* partition *of A if for each $a \in A$ there exists a unique $D \in \Delta$ with $a \in D$.*

(b) $\sim_\Delta = \Big(A, A, \Big\{(a, b) \in A \times A \mid \{a, b\} \subseteq D \textit{ for some } D \in \Delta\Big\}\Big).$

Example B.2.2. The relation corresponding to a partition $\Delta = \Big\{\{1, 3\}, \{2\}\Big\}$ of $A = \{1, 2, 3\}$

$\{1, 3\}$ is the only member of Δ containing 1, $\{2\}$ is the only member of Δ containing 2 and $\{1, 3\}$ is the only member of Δ containing 3. So Δ is a partition of A.

Note that $\{1, 2\}$ is not contained in an element of Δ and so $1 \nsim_\Delta 2$. $\{1, 3\}$ is contained in $\{1, 3\}$ and so $1 \sim_\Delta 3$. Altogether the relation $\sim_\Delta$ can be described by the following table

$\sim_\Delta$	1	2	3
1	x	$-$	x
2	$-$	x	$-$
3	x	$-$	x

where we placed an x in row a and column b of the table iff $a \sim_\Delta b$.

We now computed the classes of $\sim_\Delta$. We have

$$[1] = \{b \in A \mid 1 \sim_\Delta b\} = \{1, 3\}$$
$$[2] = \{b \in A \mid 2 \sim_\Delta b\} = \{2\}$$

and

$$[3] = \{b \in A \mid 3 \sim_\Delta b\} = \{1, 3\}$$

Thus $A/ \sim_\Delta = \{\{1, 3\}, \{2\}\} = \Delta$.

So the set of classes of relation $\sim_\Delta$ is just the original partition Δ. The next theorem shows that this is true for any partition.

Proposition B.2.3. *Let A be set.*

(a) *If $\sim$ is an equivalence relation, then $A/\sim$ is a partition of A and $\sim = \sim_{A/\sim}$.*

(b) *If Δ is partition of A, then $\sim_\Delta$ is an equivalence relation and $\Delta = A/ \sim_\Delta$.*

Proof. (a) Let $a \in A$. Since $\sim$ is reflexive we have $a \sim a$ and so $a \in [a]$ by definition of $[a]$. Let $D \in A/ \sim$ with $a \in D$. Then $D = [b]$ for some $b \in A$ and so $a \in [b]$. 0.5.8 implies $[a] = [b] = D$. So $[a]$ is the unique member of $A/ \sim$ containing a. Thus $A/ \sim$ is a partition of A. Put $\approx = \sim_{A/\sim}$. Then $a \approx b$ if and only if $\{a, b\} \subseteq D$ for some $D \in A/ \sim$. We need to show that $a \approx b$ if and only if $a \sim b$.

So let $a, b \in A$ with $a \approx b$. Then $\{a, b\} \subseteq D$ for some $D \in A/ \sim$. By the previous paragraph, $[a]$ is the only member of $A/ \sim$ containing a. Thus $D = [a]$ and similarly $D = [b]$. Thus $[a] = [b]$ and 0.5.8 implies $a \sim b$.

Now let $a, b \in A$ with $a \sim b$. Then both a and b are contained in $[b]$ and so $a \approx b$.

We proved that $a \approx b$ if and only if $a \sim b$ and so (a) is proved.

(b) Let $a \in A$. Since Δ is a partition, there exists $D \in \Delta$ with $a \in \Delta$. Thus $\{a, a\} \subseteq D$ and hence $a \sim_\Delta a$. So $\sim_\Delta$ is reflexive. If $a \sim_\Delta b$ then $\{a, \beta\} \subseteq D$ for some $D \in \Delta$. Then also $\{b, a\} \subseteq D$ and hence $b \sim_\Delta$. There $\sim$ is symmetric. Now suppose that $a, b, c \in A$ with $a \sim_\Delta b$ and $b \sim_\Delta c$. Then there exists $D, E \in \Delta$ with $a, b \in D$ and $b, c \in E$. Since b is contained in a unique member of Δ, $D = E$ and so $a \sim_\Delta c$. Thus $\sim_\Delta$ is an equivalence relation.

It remains to show that $\Delta = A/ \sim_\Delta$. For $a \in A$ let $[a] = [a]_{\sim\Delta}$. We will prove:

$(*)$ *Let $D \in \Delta$ and $a \in D$. Then $D = [a]$.*

Let $b \in D$. Then $\{a, b\} \in D$ and so $a \sim_\Delta b$ by definition of $\sim_\Delta$. Thus $b \in [a]$ by definition of $[a]$. It follows that $D \subseteq [a]$.

Let $b \in [a]$. Then $a \sim_\Delta b$ by definition of $[a]$ and thus $\{a, b\} \in E$ for some $E \in \Delta$. Since Δ is a partition, a is contained in a unique member of Δ and so $E = D$. Thus $b \in D$ and so $[a] \subseteq D$. We proved $D \subseteq [a]$ and $[a] \subseteq D$ and so $(*)$ holds.

Let $D \in \Delta$. Since Δ is a partition of A, D is non-empty subset of A. So we can pick $a \in D$ and $(*)$ implies $D = [a]$. Thus $D \in A/ \sim_\Delta$ and so $\Delta \subseteq A/ \sim_\Delta$

Let $E \in A/ \sim_\Delta$. Then $E = [a]$ for some $a \in A$. Since Δ is a partition, $a \in D$ for some $D \in \Delta$. $(*)$ gives $D = [a] = E$ and so $E \in \Delta$. This shows $A/ \sim_\Delta \subseteq \Delta$.

Together with $\Delta \subseteq A/ \sim_\Delta$ this gives $\Delta = A/ \sim_\Delta$ and (b) is proved. $\square$

Appendix C

Real numbers, integers and natural numbers

In this part of the appendix we list properties of the real numbers, integers and natural numbers we assume to be true.

C.1 Definition of the real numbers

Definition C.1.1. *The real numbers are a quadtruple* $(\mathbb{R}, +, \cdot, \leq)$ *such that*

($\mathbb{R}$ i) $\mathbb{R}$ *is a set (whose elements are called* real numbers*)*

($\mathbb{R}$ ii) $+$ *is a function (called* addition*)* , $\mathbb{R} \times \mathbb{R}$ *is a subset of the domain of* $+$ *and*

$$a + b \in \mathbb{R} \qquad \text{(Closure of addition)}$$

for all $a, b \in \mathbb{R}$*, where* $a \oplus b$ *denotes the image of* (a, b) *under* $+$*;*

($\mathbb{R}$ iii) $\cdot$ *is a function (*called *multiplication),* $\mathbb{R} \times \mathbb{R}$ *is a subset of the domain of* $\cdot$ *and*

$$a \cdot b \in \mathbb{R} \qquad \text{(Closure of multiplication)}$$

for all $a, b \in \mathbb{R}$ *where* $a \cdot b$ *denotes the image of* (a, b) *under* $\cdot$*. We will also use the notion* ab *for* $a \cdot b$*.*

($\mathbb{R}$ iv) $\leq$ *is a relation from* $\mathbb{R}$ *and* $\mathbb{R}$*;*

and such that the following statements hold:

($\mathbb{R}$ Ax 1) $a + b = b + a$ *for all* $a, b \in \mathbb{R}$. *(*Commutativity of Addition*)*

($\mathbb{R}$ Ax 2) $a + (b + c) = (a + b) + c$ *for all* $a, b, c \in \mathbb{R}$*;* *(*Associativity of Addition*)*

($\mathbb{R}$ Ax 3) *There exists an element in $\mathbb{R}$, denoted by 0 (and called* zero*), such that $a+0=a$ and $0+a=a$ for all $a \in \mathbb{R}$;* (Existence of Additive Identity)

($\mathbb{R}$ Ax 4) *For each $a \in \mathbb{R}$ there exists an element in $\mathbb{R}$, denoted by $-a$ (and called negative a) such that $a+(-a)=0$ and $(-a)+a=0$;* (Existence of Additive Inverse)

($\mathbb{R}$ Ax 5) *$a(b+c)=ab+ac$ for all $a,b,c \in \mathbb{R}$.* (Right Distributivity)

($\mathbb{R}$ Ax 6) *$(a+b)c=ac+bc$ for all $a,b,c \in \mathbb{R}$* (Left Distributivity)

($\mathbb{R}$ Ax 7) *$(ab)c=a(bc)$ for all $a,b,c \in \mathbb{R}$* (Associativity of Multiplication)

($\mathbb{R}$ Ax 8) *There exists an element in $\mathbb{R}$, denoted by 1 (and called* one*), such that $1a=a$ for all $a \in R$.* (Multiplicative Identity)

($\mathbb{R}$ Ax 9) *For each $a \in \mathbb{R}$ with $a \neq 0$ there exists an element in $\mathbb{R}$, denoted by $\frac{1}{a}$ (and called 'a* inverse*') such that $aa^{-1}=1$ and $a^{-1}a=1$;*

(Existence of Multiplicative Inverse)

($\mathbb{R}$ Ax 10) *For all $a,b \in \mathbb{R}$,*

$$(a \leq b \text{ and } b \leq a) \iff (a=b)$$

($\mathbb{R}$ Ax 11) *For all $a,b,c \in \mathbb{R}$,*

$$(a \leq b \text{ and } b \leq c) \implies (a \leq c)$$

($\mathbb{R}$ Ax 12) *For all $a,b,c \in \mathbb{R}$,*

$$(a \leq b \text{ and } 0 \leq c) \implies (ac \leq bc)$$

($\mathbb{R}$ Ax 13) *For all $a,b,c \in \mathbb{R}$,*

$$(a \leq b) \implies (a+c \leq b+c)$$

($\mathbb{R}$ Ax 14) *Each bounded, non-empty subset of $\mathbb{R}$ has a least upper bound. That is, if S is a non-empty subset of $\mathbb{R}$ and there exists $u \in \mathbb{R}$ with $s \leq u$ for all $s \in S$, then there exists $m \in R$ such that for all $r \in \mathbb{R}$,*

$$\Big(s \leq r \text{ for all } s \in S\Big) \iff \Big(m \leq r\Big)$$

($\mathbb{R}$ Ax 15) *For all $a,b \in \mathbb{R}$ such that $b \neq 0$ and $0 \leq b$ there exists a positive integer n such that $a \leq nb$. (Here na is inductively defined by $1a=a$ and $(n+1)a=na+a$).*

Definition C.1.2. *The relations $<$, $\geq$ and $>$ on $\mathbb{R}$ are defined as follows: Let $a,b \in \mathbb{R}$, then*

(a) *$a<b$ if $a \leq b$ and $a \neq b$.*

(b) *$a \geq b$ if $b \leq a$.*

(c) *$a>b$ if $b \leq a$ and $a \neq b$*

C.2 Algebraic properties of the integers

Lemma C.2.1. *Let $a, b, c \in \mathbb{Z}$. Then*

(1) $a + b \in \mathbb{Z}$.

(2) $a + (b + c) = (a + b) + c$.

(3) $a + b = b + a$.

(4) $a + 0 = a = 0 + a$.

(5) *There exists* $x \in \mathbb{Z}$ *with* $a + x = 0$.

(6) $ab \in \mathbb{Z}$.

(7) $a(bc) = (ab)c$.

(8) $a(b + c) = ab + ac$ *and* $(a + b)c = ac + bc$.

(9) $ab = ba$.

(10) $a1 = a = 1a$.

(11) *If* $ab = 0$ *then* $a = 0$ *or* $b = 0$.

C.3 Properties of the order on the integers

Lemma C.3.1. *Let a, b, c be integers.*

(a) *Exactly one of* $a < b, a = b$ *and* $b < a$ *holds.*

(b) *If* $a < b$ *and* $b < c$, *then* $a < c$.

(c) *If* $c > 0$, *then* $a < b$ *if and only if* $ac < bc$.

(d) *If* $c < 0$, *then* $a < b$ *if and only if* $bc < ac$.

(e) *If* $a < b$, *then* $a + c < b + c$.

(f) 1 *is the smallest positive integer.*

C.4 Properties of the natural numbers

Lemma C.4.1. *Let $a, b \in \mathbb{N}$. Then*

(a) $a + b \in \mathbb{N}$.

(b) $ab \in \mathbb{N}$.

Theorem C.4.2 (Well-Ordering Axiom)**.** *Let S be a non-empty subset of $\mathbb{N}$. Then S has a minimal element* □

Appendix D

The Associative, Commutative and Distributive Laws

D.1 The General Associative Law

Definition D.1.1. *Let G be a set.*

(a) *A* binary *operation on G is a function $+$ such that $G \times G$ is a subset of the domain of $+$ and $+(a, b) \in G$ for all $a, b \in G$.*

(b) *If $+$ is a binary operation on G and $a, b \in G$, then we write $a + b$ for $+(a, b)$.*

(c) *A binary operation $+$ on G is called associative if $a + (b + c) = (a + b) + c$ for all $a, b, c \in G$.*

Definition D.1.2. *Let G be a set and $+ : G \times G \to G, (a, b) \to a + b$ a function. Let n be a positive integer and $a_1, a_2, \ldots a_n \in G$. Define $\sum_{i=1}^{1} a_i = a_1$ and inductively for $n > 1$*

$$\sum_{i=1}^{n} a_i = \left(\sum_{i=1}^{n-1} a_i\right) + a_n.$$

so $\sum_{i=1}^{n} a_i = \Big(\Big(\ldots((a_1 + a_2) + a_3) + \ldots + a_{n-2}\Big) + a_{n-1}\Big) + a_n.$

Inductively, we say that z is a sum of $(a_1, \ldots, a_n)$ provided that one of the following holds:

(1) *$n = 1$ and $z = a_1$.*

(2) *$n > 1$ and there exists an integer k with $1 \leq k < n$ and $x, y \in G$ such that x is a sum of $(a_1, \ldots, a_k)$, y is a sum of $(a_{k+1}, a_{k+2}, \ldots, a_n)$ and $z = x + y$.*

For example a is the only sum of (a), $a + b$ is the only sum of (a, b), $a + (b + c)$ and $(a + b) + c$ are the sums of (a, b, c), and $a + (b + (c + d)), a + ((b + c) + d), (a + b) + (c + d), (a + (b + c)) + d$ and $((a + b) + c) + d$ are the sums of (a, b, c, d).

Theorem D.1.3 (General Associative Law)**.** *Let + be an associative binary operation on the set G. Then any sum of $(a_1, a_2, \ldots, a_n)$ is equal to $\sum_{i=1}^n a_i$.*

Proof. The proof is by complete induction. For a positive integer n let $P(n)$ be the statement:

If $a_1, a_2, \ldots a_n$ are elements of G and z is a sum of $(a_1, a_2, \ldots, a_n)$, then $z = \sum_{i=1}^n a_i$.

Suppose now that n is a positive integer with n and $P(k)$ is true all integer $1 \leq k < n$. Let $a_1, a_2, \ldots a_n$ be elements of G and z is a sum of $(a_1, a_2, \ldots, a_n)$. We need to show that $z = \sum_{i=1}^n a_i$.

Assume that $n = 1$. By definition a_1 is the only sum of (a_1) and $\sum_{i=1}^1 a_1 = a_1$. So $z = a_1 = \sum_{i=1}^n a_i$

Assume next that $n > 1$. We will first show that

(*) If u is any sum of $(a_1, \ldots, a_{n-1})$, then $u + a_n = \sum_{i=1}^n a_i$.

Indeed by the induction assumption, $P(n-1)$ is true and so $u = \sum_{i=1}^{n-1} a_i$. Thus $u + a_n = \sum_{i=1}^{n-1} a_i + a_n$ and the definition of $\sum_{i=1}^n a_i$ implies $u + a_n = \sum_{i=1}^n a_i$. So (*) is true.

By the definition of 'sum' there exists $1 \leq k < n$, a sum x of $(a_1, \ldots, a_k)$ and a sum y of $(a_{k+1}, \ldots, a_n)$ such that $z = x + y$.

Case 1: $k = n - 1$.

In this case x is a sum of $(a_1, \ldots, a_{n-1})$ and y a sum of (a_n). So $y = a_n$ and by (**) applied with $x = u$ we have $z = x + y = x + a_n = \sum_{i=1}^n a_i$.

Case 2: $1 \leq k < n - 1$.

Observe that $n - k \leq n - 1 < n$ and so by the induction assumption $P(n-k)$ holds. Since y is a sum of $a_{k+1}, \ldots, a_n)$ we conclude that $y = \sum_{i=1}^{n-k} a_{k+i}$. Since $k < n - 1$, $1 < n - k$ and so by definition of Σ, $y = \sum_{i=1}^{n-k-1} a_{k+i} + a_n$. Since + is associative we compute

$$z = x + y = x + (\sum_{i=1}^{n-k} a_{k+i} + a_n) = (x + \sum_{i=1}^{n-k-1} a_{k+i}) + a_n$$

Put $u = x + \sum_{i=1}^{n-k-1} a_{k+i}$. Then $z = u + a_n$. Also x is a sum of $(a_1, \ldots, a_k)$ and $\sum_{i=1}^{n-k-1} a_{k+i}$ is a sum of $(a_k, \ldots, a_{n-1})$. So by definition of a sum, u is a sum of $(a_1, \ldots, a_{n-1})$. Thus by (**), $z = u + a_n = \sum_{i=1}^n a_i$.

We proved that in both cases $z = \sum_{i=1}^n a_i$. Thus $P(n)$ holds. By the principal of complete induction, $P(n)$ holds for all positive integers n. □

D.2 The general commutative law

Definition D.2.1. *A binary operation + on a set G is called* commutative *if $a + b = b + a$ for all $a, b \in G$.*

Theorem D.2.2 (General Commutative Law I). *Let + be an associative and commutative binary operation on a set G. Let $a_1, a_2, \ldots, a_n \in G$ and $f : [1 \ldots n] \to [1 \ldots n]$ a bijection. Then*

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{f(i)}$$

Proof. Obsere that the theorem clearly holds for $n = 1$. Suppose inductively its true for $n - 1$.

Since f is onto there exists a unique integer k with $f(k) = n$.

Define $g : \{1, \ldots n-1\} \to \{1, \ldots, n-1\}$ by $g(i) = f(i)$ if $i < k$ and $g(i) = f(i+1)$ if $i \geq k$. We claim that g is a bijection. For this let $1 \leq l \leq n-1$ be an integer. Then $l = f(m)$ for some $1 \leq m \leq n$. Since $l \neq n$ and f is 1-1, $m \neq k$. If $m < k$, then $g(m) = f(m) = l$ and if $m > k$, then $g(m-1) = f(m) = l$. Thus g is onto and by G.1.7(b) g is also 1-1. By assumption the theorem is true for $n - 1$ and so

$$(*) \qquad \sum_{i=1}^{n-1} a_i = \sum_{i=1}^{n-1} a_{g(i)}$$

Using the general associative law (GAL, Theorem D.1.3) we have

$$\begin{array}{rcl}
 & & \sum_{i=1}^{n} a_{f(i)} \\
\text{(GAL)} & = & (\sum_{i=1}^{k-1} a_{f(i)}) + (a_{f(k)} + \sum_{i=k+1}^{n} a_{f(i)}) \\
(n = f(k)) & = & (\sum_{i=1}^{k-1} a_{f(i)}) + (a_n + \sum_{i=k+1}^{n} a_{f(i)}) \\
(`+' \text{ commutative }) & = & (\sum_{i=1}^{k-1} a_{f(i)}) + (\sum_{i=k+1}^{n} a_{f(i)} + a_n) \\
('+'\text{associative }) & = & ((\sum_{i=1}^{k-1} a_{f(i)}) + (\sum_{i=k+1}^{n} a_{f(i)})) + a_n \\
(\text{Substitution } j = i+1) & = & ((\sum_{i=1}^{k-1} a_{f(i)}) + (\sum_{j=k}^{n-1} a_{f(j+1)})) + a_n \\
(\text{definition of } g) & = & ((\sum_{i=1}^{k-1} a_{g(i)}) + (\sum_{j=k}^{n-1} a_{g(j)})) + a_n \\
\text{(GAL)} & = & (\sum_{i=1}^{n-1} a_{g(i)}) + a_n \\
(*) & = & (\sum_{i=1}^{n-1} a_i) + a_n \\
(\text{definition of } \sum) & = & \sum_{i=1}^{n} a_i
\end{array}$$

So the Theorem holds for n and thus by the Principal of Mathematical induction for all positive integers. □

Corollary D.2.3. *Let + be an associative and commutative binary operation on a set G. I a non-empty finite set and for $i \in I$ let $b_i \in G$. Let $g, h : \{1, \ldots, n\} \to I$ be bijections, then*

$$\sum_{i=1}^{n} b_{g(i)} = \sum_{i=1}^{n} b_{h(i)}$$

Proof. For $1 \leq i \leq n$, define $a_i = b_{g(i)}$. Let $f = g^{-1} \circ h$. Then f is a bijection. Moreover, $g \circ f = h$ and $a_{f(i)} = b_{g(f(i))} = b_{h(i))}$. Thus

$$\sum_{i=1}^{n} b_{h(i)} = \sum_{i=1}^{n} a_{f(i)} \overset{\text{D.2.2}}{=} \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_{g(i)}$$

□

Definition D.2.4. *Let $+$ be an associative and commutative binary operation on a set G. I a finite set and for $i \in I$ let $b_i \in G$. Then $\sum_{i\in I} a_i := \sum_{i=1}^{n} b_{f(i)}$, where $n = |I|$ and $f := \{1, \ldots, n\}$ is bijection. (Observe here that by D.2.3 this does not depend on the choice of f.)*

Theorem D.2.5 (General Commutative Law II)**.** *Let $+$ be an associative and commutative binary operation on a set G. I a finite set, $(I_j, | \ j \in J)$ a partition of I and for $i \in I$ let $a_i \in G$. Then*

$$\sum_{i\in I} a_i = \sum_{j\in J} \left(\sum_{i \in I_J} a_i \right)$$

Proof. The proof is by induction on $|J|$. If $|J| = 1$, the result is clearly true. Suppose next that $|J| = 2$ and say $J = \{j_1, j_2\}$. Let $f_i : \{1, \ldots, n_i\} \to I_{j_i}$ be a bijection and define $f : \{1 \ldots, n_1 + n_2\} \to I$ by $f(i) = f_1(i)$ if $1 \leq i \leq n_1$ and $f(i) = f_2(i - n_1)$ if $n_1 + 1 \leq i \leq n_1 + n_2$. Then clearly f is a onto and so by G.1.7(b), f is 1-1. We compute

$$\begin{array}{rcc}
\sum_{i\in I} a_i & = & \sum_{i=1}^{n_1+n_2} a_{f(i)} \\
& \overset{\text{GAL}}{=} & \left(\sum_{i=1}^{n_1} a_{f(i)}\right) + \left(\sum_{i=n_1+1}^{n_1+n_2} a_{f(i)}\right) \\
& = & \left(\sum_{i=1}^{n_1} a_{f_1(i)}\right) + \left(\sum_{i=1}^{n_2} a_{f_2(i)}\right) \\
& = & \left(\sum_{i\in I_{j_1}} a_i\right) + \left(\sum_{i\in I_{j_2}} a_i\right) \\
& = & \sum_{j\in J} \left(\sum_{i\in I_j} a_i\right)
\end{array}$$

Thus the theorem holds if $|J| = 2$. Suppose now that the theorem is true whenever $|J| = k$. We need to show it is also true if $|J| = k + 1$. Let $j \in J$ and put $Y = I \setminus J_j$. Then $(I_k \mid j \neq k \in J)$ is a partition of Y and (I_j, Y) is partition of I. By the induction assumption, $\sum_{i\in Y} a_i = \sum_{j\neq k\in J} \left(\sum_{i\in I_k} a_i\right)$ and so by the $|J| = 2$-case

$$\begin{array}{rcc}
\sum_{i\in I} a_i & = & \left(\sum_{i\in I_j} a_i\right) + \left(\sum_{i\in Y} a_i\right) \\
& = & \left(\sum_{i\in I_j} a_i\right) + \left(\sum_{j\neq k\in J} \left(\sum_{i\in I_k} a_i\right)\right) \\
& = & \sum_{j\in J} \left(\sum_{i\in I_J} a_i\right)
\end{array}$$

The theorem now follows from the Principal of Mathematical Induction. □

D.3 The General Distributive Law

Definition D.3.1. *Let $(+, \cdot)$ be a pair of binary operation on the set G. We say that*

(a) $(+, \cdot)$ *is* left-distributive *if* $a(b + c) = (ab) + (ac)$ *for all* $a, b, c \in G$.

(b) $(+, \cdot)$ *is* right-distributive *if* $(b + c)a = (ba) + (ca)$ *for all* $a, b, c \in G$.

(c) $(+, \cdot)$ *is* distributive *if its is right- and left-distributive.*

Theorem D.3.2 (General Distributive Law)**.** *Let $(+, \cdot)$ be a pair of binary operations on the set G.*

(a) *Suppose $(+, \cdot)$ is left-distributive and let $a, b_1, \dots b_m \in G$. Then*

$$a \cdot (\sum_{j=1}^{m} b_j) = \sum_{j=1}^{m} ab_j$$

(b) *Suppose $(+, \cdot)$ is right-distributive and let $a_1, \dots a_n, b \in G$. Then*

$$(\sum_{i=1}^{m} a_i) \cdot b = \sum_{i=1}^{n} a_i b$$

(c) *Suppose $(+, \cdot)$ is distributive and let $a_1, \dots a_n, b_1, \dots b_m \in G$. Then*

$$(\sum_{i=1}^{n} a_i) \cdot (\sum_{j=1}^{m} b_j) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_i b_j \right)$$

Proof. (a) Clearly (a) is true for $m = 1$. Suppose now (a) is true for k and let $a, b_1, \dots b_{k+1} \in G$. Then

$$\begin{array}{rcl} & & a \cdot \left(\sum_{i=1}^{k+1} b_i \right) \\ \text{(definition of } \sum) & = & a \cdot \left((\sum_{i=1}^{k} b_i) + b_{k+1} \right) \\ \text{(left-distributive)} & = & a \cdot \left(\sum_{i=1}^{k} b_i \right) + a \cdot b_{k+1} \\ \text{(induction assumption)} & = & \left(\sum_{i=1}^{k} ab_i \right) + ab_{k+1} \\ \text{(definition of } \sum) & = & \sum_{i=1}^{k+1} ab_i \end{array}$$

Thus (a) holds for $k + 1$ and so by induction for all positive integers n.

The proof of (b) is virtually the same as the proof of (a) and we leave the details to the reader.

(c)

$$\left(\sum_{i=1}^{m} a_i \right) \cdot \left(\sum_{i=1}^{k} b_i \right) \overset{(b)}{=} \sum_{i=1}^{n} \left(a_i \sum_{j=1}^{m} b_j \right) \overset{(a)}{=} \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_i b_j \right)$$

$\square$

Appendix E

Verifying Ring Axioms

Proposition E.0.3. *Let $(R, +, \cdot)$ be ring and $(S, \oplus, \odot)$ a set with binary operations $\oplus$ and $\odot$. Suppose there exists an onto homomorphism $\Phi : R \to S$ (that is an onto function $\Phi : R \to S$ with $\Phi(a + b) = \Phi(a) \oplus \Phi(b)$ and $\Phi(ab) = \Phi(a) \odot \Phi(b)$ for all $a, b \in R$. Then*

(a) *$(S, \oplus, \odot)$ is a ring and Φ is ring homomorphism.*

(b) *If R is commutative, so is S.*

Proof. (a) Clearly if S is a ring, then Φ is a ring homomorphism. So we only need to verify the eight ring axioms. For this let $a, b, c \in S$. Since Φ is onto ther exist $x, y, z \in R$ with $\Phi(x) = a, \Phi(y) = b$ and $\Phi(z) = c$.

Ax 1 By assumption $\oplus$ is binary operation. So Ax 1 holds for S.

Ax 2

$$\begin{array}{cccccccc} & a \oplus (b \oplus c) & = & \Phi(x) \oplus (\Phi(y) \oplus \Phi(z)) & = & \Phi(x) \oplus \Phi(y + z) & = & \Phi(x + (y + z)) \\ = & \Phi((x + y) + z)) & = & \Phi(x + y) \oplus \Phi(z) & = & (\Phi(x) \oplus \Phi(y)) \oplus \Phi(z) & = & (a \oplus b) \oplus c \end{array}$$

Ax 3 $a \oplus b = \Phi(x) \oplus \Phi(y) = \Phi(x + y) = \Phi(y + x) = \Phi(y) \oplus \Phi(x) = b \oplus a$

Ax 4 Put $0_S = \Phi(0_R)$. Then

$$a \oplus 0_S = \Phi(x) \oplus \Phi(0_R) = \Phi(x + 0_R) = \Phi(x) = a$$

$$0_S + a = \Phi(0_R) \oplus \Phi(x) = \Phi(0_R + x) = \Phi(x) = a.$$

Ax 5 Put $d = \Phi(-x)$. Then

$$a \oplus d = \Phi(x) \oplus \Phi(-x) = \Phi(x + (-x)) = \Phi(0_R) = 0_S$$

Ax 6 By assumption $\odot$ is binary operation . So Ax 6 holds for S.

Ax 7

$$\begin{array}{rcccccc} a \odot (b \odot c) & = & \Phi(x) \odot (\Phi(y) \odot \Phi(z)) & = & \Phi(x) \odot \Phi(yz) & = & \Phi(x(yz)) \\ = \quad \Phi((xy)z)) & = & \Phi(xy) \odot \Phi(z) & = & (\Phi(x) \odot \Phi(y)) \odot \Phi(z) & = & (a \odot b) \odot c \end{array}$$

Ax 8

$$\begin{array}{rcccccc} a \odot (b \oplus c) & = & \Phi(x) \odot (\Phi(y) \oplus \Phi(z)) & = & \Phi(x) \odot \Phi(y+z) & = & \Phi(x(y+z)) \\ = \quad \Phi(xy+xz) & = & \Phi(xy) + \Phi(xz) & = & (\Phi(x) \odot \Phi(y)) + (\Phi(x) \odot \Phi(z)) & = & (a \odot b) \oplus (a \odot c) \end{array}$$

Similarly $(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c)$.
(b) Suppose R is commutative then
3.1.2 $\quad a \odot b = \Phi(x) \odot \Phi(y) = \Phi(xy) = \Phi(yx) = \Phi(y) \odot \Phi(x) = b \odot a$ $\square$

Appendix F

Constructing rings from given rings

F.1 Direct products of rings

Definition F.1.1. *Let $(R_i)_{i\in I}$ be a family of rings (that is I is a set and for each $i \in I$, R_i is a ring).*

(a) *$\bigtimes_{i\in I} R_i$ is the set of all functions $r : I \to \bigcup_{i\in I} R_i, i \to r_i$ such that $r_i \in R_i$ for all $i \in I$.*

(b) *$\bigtimes_{i\in I} R_i$ is called the direct product of $(R_i)_{\in I}$.*

(c) *We denote $r \in \bigtimes_{i\in I} R_i$ by $(r_i)_{i\in I}$, $(r_i)_i$ or (r_i).*

(d) *For $r = (r_i)$ and $s = (s_i)$ in R define $r + s = (r_i + s_i)$ and $rs = (r_i s_i)$.*

Lemma F.1.2. *Let $(R_i)_{i\in I}$ be a family of rings.*

(a) *$R := \bigtimes_{i\in I} R_i$ is a ring.*

(b) $0_R = (0_{R_i})_{i\in I}$.

(c) $-(r_i) = (-r_i)$.

(d) *If each R_i is a ring with identity, then also $\bigtimes_{i\in I} R_i$ is a ring with identity and $1_R = (1_{R_i})$.*

(e) *If each R_i is commutative, then $\bigtimes_{i\in I} R_i$ is commutative.*

Proof. Left as an exercise. □

F.2 Matrix rings

Definition F.2.1. *Let R be a ring and m, n positive integers.*

(a) *An $m \times n$-*matrix *with coefficients in R is a function*

$$A : \{1, \dots, m\} \times \{1, \dots, n\} \to R, \quad (i, j) \mapsto a_{ij}.$$

(b) *We denote an $m \times n$-matrix A by $[a_{ij}]_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$, $[a_{ij}]_{ij}$, $[a_{ij}]$ or*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

(c) *Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ matrices with coefficients in R. Then $A + B$ is the $m \times n$-matrix $A + B := [a_{ij} + b_{ij}]$.*

(d) *Let $A = [a_{ij}]_{ij}$ be an $m \times n$-matrix and $B = [b_{jk}]_{jk}$ an $n \times p$ matrix with coefficients in R. Then AB is the $m \times p$ matrix $AB = [\sum_{j=1}^{n} a_{ij}b_{jk}]_{ik}$.*

(e) $\mathrm{M}_{mn}(R)$ *denotes the set of all $m \times n$ matrices with coefficients in R.* $\mathrm{M}_n(R) = \mathrm{M}_{nn}(R)$.

It might be useful to write out the above definitions of $A + B$ and AB in longhand notation:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1}+b_{m2} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}$$

and

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{mp} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+\ldots+a_{1n}b_{n1} & a_{11}b_{12}+a_{12}b_{22}+\ldots+a_{1n}b_{n2} & \dots & a_{11}b_{1p}+a_{12}b_{2p}+\ldots+a_{1n}b_{np} \\ a_{21}b_{11}+a_{22}b_{21}+\ldots+a_{2n}b_{n1} & a_{21}b_{12}+a_{22}b_{22}+\ldots+a_{2n}b_{n2} & \dots & a_{21}b_{1p}+a_{22}b_{2p}+\ldots+a_{2n}b_{np} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1}b_{11}+a_{m2}b_{21}+\ldots+a_{mn}b_{n1} & a_{m1}b_{12}+a_{m2}b_{22}+\ldots+a_{mn}b_{n2} & \dots & a_{m1}b_{1p}+a_{m2}b_{2p}+\ldots+a_{mn}b_{np} \end{bmatrix}$$

Lemma F.2.2. *Let n be an integer and R an ring. Then*

(a) $(\mathrm{M}_n(R), +, \cdot)$ *is a ring.*

(b) $0_{\mathrm{M}_n(R)} = (0_R)_{ij}$.

(c) $-[a_{ij}] = [-a_{ij}]$ *for any* $[a_{ij}] \in \mathrm{M}_n(R)$.

(d) *If R has an identity, then $\mathrm{M}_n(R)$ has an identity and $1_{\mathrm{M}_n(R)} = (\delta_{ij})$, where*

$$\delta_{ij} = \begin{cases} 1_R & \textit{if } i = j \\ 0_R & \textit{if } i \neq j \end{cases}$$

Proof. Put $J = \{1, \ldots, n\} \times \{1, \ldots, m\}$ and observe that $(\mathrm{M}_n(R), +) = (\bigtimes_{j\in J} R, +)$. So F.1.2 implies that Ax 1-Ax 5, (b) and (c) hold.

Clearly Ax 6 holds. To verify Ax 7 let $A = [a_{ij}], B = [b_{jk}]$ and $C = [c_{kl}]$ be in $\mathrm{M}_n(R)$. Put $D = AB$ and $E = BC$. Then

$$(AB)C = DC = \left[\sum_{k=1}^{n} d_{ik}c_{kl}\right]_{il} = \left[\sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij}b_{jk}\right) c_{kl}\right]_{il} = \left[\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ij}b_{jk}c_{kl}\right]_{il}$$

and

$$A(BC) = AE = \left[\sum_{j=1}^{n} a_{ij}e_{jl}\right]_{il} = \left[\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk}c_{kl}\right)\right]_{il} = \left[\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ij}b_{jk}c_{kl}\right]_{il}$$

Thus $A(BC) = (AB)C$.

$$\begin{aligned}(A+B)C = [a_{ij} + b_{ij}]_{ij} \cdot [c_{jk}]_{ik} &= \left[\sum_{j=1}^{n}(a_{ij} + b_{ij})c_{jk}\right]_{ik} \\ = \left[\sum_{j=1}^{n} a_{ij}c_{jk}\right]_{ik} &+ \left[\sum_{j=1}^{n} b_{ij}c_{jk}\right]_{ik} = AC + BC.\end{aligned}$$

So $(A+B)C = AC + BC$ and similarly $A(B+C) = AB + AC$. Thus $\mathrm{M}_n(R)$ is a ring.

Suppose now that R has an identity 1_R. Put $I = [\delta_{ij}]_{ij}$, where

$$\delta_{ij} = \begin{cases} 1_R & \text{if } i = j \\ 0_R & \text{if } i = j \end{cases}$$

If $i \neq j$, then $\delta_{ij}a_{jk} = 0_R a_{jk} = 0_R$ and if $i = j$ then $\delta_{ij}a_{jk} = 1_F a_{ik} = a_{ik}$. Thus

$$IA = \left[\sum_{j=1} \delta_{ij}a_{jk}\right]_{ik} = [a_{ik}]_{ik} = A$$

and similarly $AI = A$. Thus A is an identity in R and so (d) holds. $\square$

F.3 Polynomial Rings

In this section we show that if R is ring with identity then existence of a polynomial ring with coefficients in R.

Theorem F.3.1. *Let R be a ring. Let P be the set of all functions $f : \mathbb{N} \to R$ such that there exists $m \in \mathbb{N}^*$ with*

$$f(i) = 0_R \text{ for all } i > m \tag{1}$$

We define an addition and multiplication on P by

$$(f+g)(i) = f(i) + g(i) \quad \text{and} \quad (fg)(i) = \sum_{k=0}^{i} f(i)g(k-i) \tag{2}$$

(a) *P is a ring.*

(b) *For $r \in R$ define $r^\circ \in P$ by*

$$r^\circ(i) := \begin{cases} r & \text{if } i = 0 \\ 0_R & \text{if } i \neq 0 \end{cases} \tag{3}$$

Then the map $R \to P, r \to r^\circ$ is a 1-1 homomorphism.

(c) *Suppose R has an identity and define $x \in P$ by*

$$x(i) := \begin{cases} 1_R & \text{if } i = 1 \\ 0_R & \text{if } i \neq 1 \end{cases}$$

Then (after identifying $r \in R$ with r° in P), P is a polynomial ring with coefficients in R and indeterminate x.

Proof. Let $f, g \in P$. Let $\deg f$ be the minimal $m \in \mathbb{N}^*$ for which (1) holds. Observe that (2) defines functions $f+g$ and fg from $\mathbb{N}$ to R. So to show that $f+g$ and fg are in P we need to verify that (1) holds for $f+g$ and fg as well. Let $m = \max \deg f, \deg g$ and $n = \deg f + \deg g$. Then for $i > m$, $f(i) = 0_R$ and $g(i) = 0_R$ and so also $(f+g)(i) = 0_R$. Also if $i > n$ and $0 \leq k \leq i$, then either $k < \deg f$ or $i - k > \deg g$. In either case $f(k)g(i-k) = 0_R$ and so $(fg)(i) = 0_R$. So we indeed have $f + g \in P$ and $fg \in P$. Thus axiom Ax 1 and Ax 6 hold. We now verify the remaining axioms one by one. Observe that f and g in P are equal if and only if $f(i) = g(i)$ for all $i \in \mathbb{N}$. Let $f, g, h \in P$ and $i \in \mathbb{N}$.

(**Ax 2**

$$\begin{aligned}((f+g)+h)(i) &= (f+g)(i)+h(i) &&= (f(i)+g(i))+h(i) &&= f(i)+(g(i)+h(i))\\ &= f(i)+(g(i)+h(i)) &&= f(i)+(g+h)(i) &&= (f+(g+h))(i)\end{aligned}$$

(**Ax 3** $(f+g)(i) = f(i)+g(i) = g(i)+f(i) = (g+f)(i)$

(**Ax 4** Define $0_P \in P$ by $0_P(i) = 0_R$ for all $i \in \mathbb{N}$. Then

$$\begin{aligned}(f+0_P)(i) = f(i)+0_P(i) = f(i)+0_R = f(i)\\ (0_P+f)(i) = 0_P(i)+f(i) = 0_R+f(i) = f(i)\end{aligned}$$

Ax 5 Define $-f \in P$ by $(-f)(i) = -f(i)$ for all $i \in \mathbb{N}$. Then

$$(f+(-f))(i) = f(i)+(-f)(i) = f(i)+(-f(i)) = 0_R = 0_P(i)$$

Ax 7 Any triple of non-negative integers (k,l,p) with $k+l+p=i$ be uniquely written as $(k, j-k, i-j)$ where $0 \le j \le i$ and $0 \le k \le j-k)$ and uniquely as $(k,l,i-k-l)$ where $0 \le i \le k$ and $0 \le l \le i-k$. This is used in the fourth equality sign in the following computation:

$$\begin{aligned}((fg)h)(i) &= \sum_{j=0}^{i}(fg)(j)\cdot h(i-j) &&= \sum_{j=0}^{i}\left(\left(\sum_{k=0}^{j} f(k)g(j-k)\right)h(i-j)\right)\\ &= \sum_{j=0}^{i}\left(\sum_{k=0}^{j} f(k)g(j-k))h(i-j)\right) &&= \sum_{k=0}^{i}\left(\sum_{l=0}^{i-k} f(k)g(l)h(i-k-l)\right))\\ &= \sum_{k=0}^{i}\left(f(k)\left(\sum_{l=0}^{i-k} g(l)h(i-k-l)\right)\right) &&= \sum_{k=0}^{i} f(k)\cdot(gh)(i-k)\\ & &&= (f(gh))(i)\end{aligned}$$

Ax 8

$$\begin{aligned}(f\cdot(g+h))(i) &= \sum_{j=0}^{i} f(j)\cdot(g+h)(i-j) &&= \sum_{j=0}^{i} f(j)\cdot(g(i-j)+h(i-j))\\ &= \sum_{j=0}^{i} f(j)g(i-j)+f(j)h(i-j) &&= \sum_{j=0}^{i} f(j)g(i-j)+\sum_{j=0}^{i} f(j)h(i-j)\\ &= (fg)(i)+(fh)(i) &&= (fg+fh)(i)\end{aligned}$$

$$\begin{aligned}((f+g)\cdot h)(i) &= \sum_{j=0}^{i}(f+g)(j)\cdot h(i-j) &&= \sum_{j=0}^{i}(f(j)+g(j))\cdot h(i-j)\\ &= \sum_{j=0}^{i} f(j)h(i-j)+g(j)h(i-j) &&= \sum_{j=0}^{i} f(j)h(i-j)+\sum_{j=0}^{i} g(j)h(i-j)\\ &= (fh)(i)+(gh)(i) &&= (fh+gh)(i)\end{aligned}$$

Since Ax 1 through Ax 8 hold we conclude that P is a ring and (a) is proved. Let $r, s \in R$ and $k, l \in \mathbb{N}$. We compute

$$(r+s)^\circ(i) \quad = \quad \begin{cases} r+s & \text{if } i=0 \\ 0_R & \text{if } i \neq 0 \end{cases} \quad = \quad r^\circ(i)+s^\circ(i) \quad = \quad (r^\circ+s^\circ)(i) \tag{4}$$

and

$$(r^\circ s)(i) = \sum_{k=0}^{i} r^\circ(k)s(i-k)$$

Note that $r^\circ(k) = 0_R$ unless $k=0$ and $s^\circ(i-k) = 0_R$ unless and $i-k=0$. Hence $r^\circ(k)s(i-k) = 0_R$ unless $k=0$ and $i-k=0$ (and so also $i=0$). Thus $(r^\circ s)(i) = 0$ if $i \neq 0$ and $(r^\circ s)(0) = r^\circ(0)s^\circ(0) = rs$. This

$$r^\circ s^\circ = (rs)^\circ \tag{5}$$

Define $\rho : R \to P, r \to r^\circ$. If $r, s \in R$ with $r^\circ = s^\circ$, then $r = r^\circ(1) = s^\circ(1) = s$ and so ρ is 1-1. By (4) and (5), ρ is a homomorphism and so (b) is proved.

Assume from now on that R has an identity.

For $k \in \mathbb{N}$ let $\delta_k \in P$ be defined by

$$\delta_k(i) := \begin{cases} 1_R & \text{if } i=k \\ 0_R & \text{if } i \neq k \end{cases} \tag{6}$$

Let $f \in P$. Then

$$(r^\circ f)(i) = \sum_{k=0}^{i} r^\circ(k)f(i-k) = r \cdot f(i) + \sum_{i=1}^{k} 0_R f(i-k) = r \cdot f(i) \tag{7}$$

and similarly

$$(fr^\circ)(i) = f(i) \cdot r \tag{8}$$

In particular, 1_R° is an identity in P. Since $\delta_0 = 1_R^\circ$ we conclude

$$\delta_0 = 1_R^\circ = 1_P \tag{9}$$

For $f = \delta_k$ we conclude that

$$(r^\circ\delta_k)(i) = (\delta_k r^\circ)(i) = \begin{cases} r & \text{if } i = k \\ 0_R & \text{if } i \neq k \end{cases} \tag{10}$$

Let $m \in \mathbb{N}$ and $a_0, \dots a_m \in R$. Then (10) implies

$$\left(\sum_{k=0}^{m} a_k^\circ \delta\right)(i) = \begin{cases} a_i & \text{if } i \leq m \\ 0_R & \text{if } i > m \end{cases} \tag{11}$$

We conclude that if $f \in P$ and $a_0, a_1, a_2, \dots a_m \in R$ then

$$f = \sum_{k=0}^{m} a_k^\circ \delta_k \quad \iff \quad m \geq \deg f \text{ and } a_k = f(k) \text{ for all } 0 \leq k \leq m \tag{12}$$

We compute

$$(\delta_k\delta_l)(i) = \sum_{j=0}^{i} \delta_k(j)\delta_l(i-j) \tag{13}$$

Since $\delta_k(j)\delta_l(i-j)$ is 0_R unless $j = k$ and $l = i - j$, that is unless $j = k$ and $i = l + k$, in which case it is 1_R, we conclude

$$(\delta_k\delta_l)(i) \quad = \quad \begin{cases} 1_R & \text{if } i = k + l \\ 0_R & \text{if } i \neq k + l \end{cases} \quad = \quad \delta_{k+l}(i) \tag{14}$$

and so

$$\delta_k\delta_l = \delta_{k+l} \tag{15}$$

Note that $x = \delta_1$. We conclude that

$$x^k = \delta_k \tag{16}$$

By (10)

$$r^\circ x = x r^\circ \quad \text{for all } r \in R \tag{17}$$

We will now verify the four conditions (i)-(iv) in the definition of a polynomial. By (b) we we can identify r with r° in R. Then R becomes a subring of P. By (9), $1_R^\circ = 1_P$. So (i) holds. By (17), (ii) holds. (iii) and (iv) follow from (12) and (16). $\square$

Lemma F.3.2. *Let R and P be rings and $x \in P$. Suppose that Conditions (i)-(iv) in 4.1.1 hold under the convention that $f_0 x^0 := f_0$ for all $f_0 \in R$. Then R and P have identities and $1_R = 1_P$.*

Proof. Since $x \in P$, 4.1.1(iii) shows that $x = \sum_{i=0}^m e_i x^i$ for some $m \in \mathbb{N}$ and $e_0, e_1, \ldots e_n \in \mathbb{R}$. Let $r \in R$. Then

$$rx = r \sum_{i=0}^{n} e_i x^i = \sum_{i=0}^{n} (re_i) x^i.$$

So 4.1.1(iv) shows that $re_1 = r$. Since $rx = xr$ by 4.1.1(ii) a similar argument gives $e_1 r = e$ and so e_1 is an identity in R and $e_1 = 1_R$. Now let $f \in P$. Then $f = \sum_{i=0}^n f_i x^i$ for some $n \in \mathbb{N}$ and $f_0, \ldots, f_n \in R$. Thus

$$f \cdot 1_R = (\sum_{i=0}^{n} f_i x^i) \cdot 1_R = \sum_{i=0}^{n} (f_i 1_R) x^i = \sum_{i=0}^{n} f_i x^i = f$$

Similarly, $1_R \cdot f = f$ and so 1_R is an identity in P. $\square$

Appendix G

Cardinalities

G.1 Cardinalities of Finite Sets

Notation G.1.1. *For $a, b \in \mathbb{Z}$ set $[a \dots b] := \{c \in \mathbb{Z} \mid a \leq c \leq b\}$.*

Lemma G.1.2. *Let $A \subsetneq [1 \dots n]$. Then there exists a bijection $\alpha : [1 \dots n] \to [1 \dots n]$ with $\alpha(A) \subseteq [1 \dots n-1]$.*

Proof. Since $A \neq [1 \dots n]$ there exists $m \in [1 \dots n]$ with $m \notin A$. Define $\alpha : [1 \dots n] \to [1 \dots n]$ by $\alpha(n) = m$, $\alpha(m) = n$ and $\alpha(i) = i$ for all $i \in [1 \dots n]$ with $n \neq i \neq m$. It is easy to verify that α is bijection. Since $\alpha(m) = n$ and $m \notin A$, $\alpha(a) \neq n$ for all $a \in A$. So $n \notin \alpha(A)$ and so $\alpha(A) \subseteq [1 \dots n] - 1$. ☐

Lemma G.1.3. *Let $n \in \mathbb{N}$ and let $\beta : [1 \dots n] \to [1 \dots n]$ be a function. If β is 1-1, then β is onto.*

Proof. The proof is by induction on n. If $n = 1$, then $\beta(1) = 1$ and so β is onto. Let $A = \beta([1 \dots n-1])$. Since $\beta(n) \notin A$, $A \neq [1 \dots n]$. Thus by G.1.2 there exists a bijection $\alpha : [1 \dots n]$ with $\alpha(A) \subseteq [1 \dots n-1]$. Thus $\alpha\beta([1 \dots n-1]) \subseteq [1 \dots n-1]$. By induction $\alpha\beta([1 \dots n-1] = [1 \dots n-1]$. Since $\alpha\beta$ is 1-1 we conclude that $\alpha\beta(n) = n$. Thus $\alpha\beta$ is onto and $\alpha\beta$ is a bijection. Since α is also a bijection this implies that β is a bijection. ☐

Definition G.1.4. *A set A is* finite *if there exists $n \in \mathbb{N}$ and a bijection $\alpha : A \to [1 \dots n]$.*

Lemma G.1.5. *Let A be a finite set. Then there exists a unique $n \in \mathbb{N}$ for which there exists a bijection $\alpha : A \to [1 \dots n]$.*

Proof. By definition of a finite set G.1.4 there exist $n \in \mathbb{N}$ and a bijection $\alpha : A \to [1 \dots n]$. Suppose that also $m \in \mathbb{N}$ and $\beta : A \to [1 \dots m]$ is a bijection. We need to show that $n = m$ and may assume that $n \leq m$. Let $\gamma : [1 \dots n] \to [1 \dots m], i \to i$ and $\delta := \gamma \circ \alpha \circ \beta^{-1}$. Then γ is a 1-1 function from $[1 \dots m]$ to $[1 \dots m]$ and so by G.1.3, δ is onto. Thus also γ is onto. Since $\gamma([1 \dots n]) = [1 \dots n]$ we conclude that $[1 \dots n] = [1 \dots m]$ and so also $n = m$. ☐

Definition G.1.6. *Let A be a finite set. Then the unique $n \in \mathbb{N}$ for which there exists a bijection $\alpha : A \to [1 \dots n]$ is called the* cardinality *or* size *of A and is denoted by $|A|$.*

Theorem G.1.7. *Let A and B be finite sets.*

(a) *If $\alpha : A \to B$ is 1-1 then $|A| \leq |B|$, with equality if and only if α is onto.*

(b) *If $\alpha : A \to B$ is onto then $|A| \geq |B|$, with equality if and only if α is 1-1.*

(c) *If $A \subseteq B$ then $|A| \leq |B|$, with equality if and only if $|A| = |B|$.*

Proof. (a) If α is onto then α is a bijection and so $|A| = |B|$. So it suffices to show that if $|A| \geq |B|$, then α is onto. Put $n = |A|$ and $m = |B|$ and let $\beta : A \to [1 \dots n]$ and $\gamma : B \to [1 \dots m]$ be bijection. Assume $n \geq m$ and let $\delta : [1 \dots m] \to [1 \dots n]$ be the inclusion map. Then $\delta\gamma\alpha\beta^{-1}$ is a 1-1 function form $[1 \dots n]$ to $[1 \dots n]$ and so by G.1.3 its onto. Hence δ is onto, $n = m$ and δ is bijection. Since also γ is bijection, this forces $\alpha\beta^{-1}$ to be onto and so also α is onto.

(b) Since α is onto there exists $\beta : B \to A$ with $\alpha\beta = \mathrm{id}_B$. Then β is 1-1 and so by (a), $|B| \leq |A|$ and β is a bijection if and only if $|A| = |B|$. Since α is a bijection if and only if β is, (b) is proved.

(c) Follows from (a) applied to the inclusion map $A \to B$. □

Proposition G.1.8. *Let A and be B be finite sets. Then*

(a) *If $A \cap B = \emptyset$, then $|A \cup B| = |A| + |B|$.*

(b) $|A \times B| = |A| \cdot |B|$.

Proof. (a) Put $n = |A|$, $m = |B|$ and let $\beta : A \to [1 \dots n]$ and $\gamma : B \to [1 \dots m]$ be bijections. Define $\gamma : A \cup B \to [1 \dots n + m]$ by

$$\gamma(c) = \begin{cases} \alpha(c) & \text{if } c \in A \\ \beta(c) + n & \text{if } c \in B \end{cases}$$

Then it is readily verified that γ is a bijection and so $|A \cup B| = n + m = |A| + |B|$.

(b) The proof is by induction on $|B|$. If $|B| = 0$, then $B = \emptyset$ and so also $A \times B = \emptyset$. If $|B| = 1$, then $B = \{b\}$ for some $b \in B$ and so the map $A \to A \times B, a \to (a, b)$ is a bijection. Thus $|A \times B| = |A| = |A| \cdot |B|$. Suppose now that (b) holds for any set B of size k. Let C be a set of size $k + 1$. Pick $c \in C$ and put $B = C \setminus \{c\}$. Then $C = B \cup \{c\}$ and so (a) implies $|B| = k$. So by induction $|A \times B| = |A| \cdot k$. Also $|A \times \{c\} = |A|$ and so by (a)

$$|A \times C| = |A \times B| + |A \times \{c\}| = |A| \cdot k + |A| = |A| \cdot (k + 1) = |A||C|$$

(b) now follows from the principal of mathematical induction 0.4.2. □

Bibliography

[1] T.W. Hungerford *Abstract Algebra, An Introduction* second edition, Brooks/Cole **1997**.