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

April 29, 2015

Contents

Chapter 1

Set, Relations and Functions

1.1 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 one often proves 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 is equivalent to the statement "$P \Longrightarrow Q$":

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 statement "$Q \implies P$" is called the *converse* of the statement " $P \implies Q$". 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 1.1.1 (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 a formula consult your favorite logic book. □

Example 1.1.2. Let $\Phi(x) = x^2 + 3 \cdot x + 4$.
If $a = 2$, then

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

Notation 1.1.3. *Let $P(x)$ be a statement involving the variable x.*

(a) "for all $x : P(x)$" *is the statement that for all 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 also use "$\exists x : P(x)$", "$P(x)$ is true for some x", "There exists x with $P(x)$" or similar phrases.*

Example 1.1.4. "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 1.1.5. *Let $P(x)$ be a statement involving the variable x.*

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

$$\text{for all } x : \text{for all } y : \qquad P(x) \text{ and } P(y) \quad \implies \quad x = y$$

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

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

Example 1.1.6. "There exists at most one $x : (x^2 = 1$ and x is a real number)" is false since $1^2 = 1$ and $(-1)^2 = 1$, but $1 \neq -1$.

"There exists a unique x : $(x^3 = -1$ and x is a real number)" is true since $x = -1$ is the only element 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 $x^2 = -1$.

Theorem 1.1.7. *Let $P(x)$ be a 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. Consult A.1.2 in the appendix for the proof. □

Exercises 1.1:

#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.

1.2 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,

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

In other words:

$$\text{for all } x : \qquad x \notin S \iff \text{not-}(x \in S)$$

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 set?

Suppose that $\mathcal{D}$ is a nice set . 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 set, $\mathcal{D}$ is not nice set.

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

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

Theorem 1.2.1. *Let A and B be sets. Then*

$$\Big(A = B\Big) \qquad \iff \qquad \Big(\text{ for all } x: \quad (x \in A) \iff (x \in B)\Big)$$

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. □

Definition 1.2.2. *Let A and B be sets. We say that A is* subset *of B and write $A \subseteq B$ if*

$$\text{for all } x: \qquad (x \in A) \implies (x \in B)$$

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

Theorem 1.2.3. *Let A and B sets. Then $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.*

Proof.

$$\begin{aligned}
& A \subseteq B \text{ and } B \subseteq A \\
\iff\ & \text{for all } x: \ (x \in A \implies x \in B) \text{ and } (x \in B \implies x \in A) && \text{–definition of subset} \\
\iff\ & \text{for all } x: \ x \in A \iff x \in B \\
& \qquad -A.1.1(LR19): \Big((P \implies Q) \text{ and } (Q \implies P)\Big) \iff \Big(P \iff Q\Big) \\
\iff\ & A = B && -\ 1.2.1
\end{aligned}$$

□

Theorem 1.2.4. *Let t be an object. Then there exists a set, denote by $\{t\}$ such that*

$$\text{for all } x: \qquad x \in \{t\} \quad \iff \quad x = t$$

Proof. This is an axiom of Set Theory. □

Theorem 1.2.5. *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*

$$\text{for all } x: \qquad x \in \{s \in S \mid P(s)\} \quad \iff \quad x \in S \ \text{ and } \ P(x)$$

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 1.2.6.

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

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

Notation 1.2.7. *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 1.2.8. (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 1.2.9. *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*

$$\text{for all } x: \qquad x \in \{\Phi(s) \mid s \in S\} \qquad \Longleftrightarrow \qquad \text{there exists } s \in S: \ x = \Phi(s)$$

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 1.2.10.

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

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

We now combine the two previous theorems into one:

Theorem 1.2.11. *Let S be a set, let $P(x)$ be a statement involving the variable x and let $\Phi(x)$ 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 $\left\{\Phi(s) \;\middle|\; s \in S \text{ and } P(s)\right\}$ such that*

$$\text{for all } x: \qquad x \in \left\{\Phi(s) \;\middle|\; s \in S \text{ and } P(s)\right\} \quad \Longleftrightarrow \quad \text{there exists } s \in S: \Big(P(s) \text{ and } x = \Phi(s)\Big)$$

Proof. Define

$$(*) \qquad \left\{\Phi(s) \;\middle|\; s \in S \text{ and } P(s)\right\} := \left\{\Phi(s)) \;\middle|\; s \in \{r \in S \mid P(r)\}\right\}$$

See A.1.3 for a formal proof that this set has the required properties.

□

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 1.2.12.

$$\left\{2n \;\middle|\; n \in \mathbb{Z} \text{ and } n^2 = 1\right\} = \left\{2n \;\middle|\; n \in \{s \in \mathbb{Z} \mid s^2 = 1\}\right\} = \left\{2n \;\middle|\; n \in \{1, -1\}\right\} = \{2, -2\}$$

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

Theorem 1.2.13. *Let A and B be sets.*

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

$$\text{for all } x: \qquad x \in A \cup B \iff x \in A \text{ or } x \in B$$

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

$$\text{for all } x: \qquad x \in A \cap B \iff x \in A \text{ and } x \in B$$

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

$$\text{for all } x: \qquad x \in A \smallsetminus B \iff x \in A \text{ and } x \notin B$$

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

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

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

$$\text{for all } x: \qquad x \in \{a, b\} \iff x = a \text{ or } x = b$$

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

(b) Applying 1.2.5 with $P(x)$ being the statement "$x \in B$" we can define

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

Then for all x:

$$\begin{aligned} & x \in A \cap B \\ \iff & x \in \{a \in A \mid a \in B\} && \text{– definition of } A \cap B \\ \iff & x \in A \text{ and } x \in B && \text{– Theorem 1.2.5} \end{aligned}$$

(c) Applying 1.2.5 with $P(x)$ being the statement "$x \notin B$" we can define

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

Then for all x:

$$\begin{aligned} & x \in A \setminus B \\ \iff & x \in \{a \in A \mid a \notin B\} && \text{– definition of } A \setminus B \\ \iff & x \in A \text{ and } x \notin B && \text{– Theorem 1.2.5} \end{aligned}$$

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

$$\varnothing := D \setminus D$$

Then for all x:

$$\begin{aligned} & x \in \varnothing \\ \iff & x \in D \setminus D && \text{– definition of } \varnothing \\ \iff & x \in D \text{ and } x \notin D && \text{– (c)} \end{aligned}$$

The latter statement is false and so $x \notin \varnothing$ for all x.

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

$$\begin{aligned} & x \in \{a, b\} \\ \iff & x \in \{a\} \cup \{b\} && \text{– definition of } \{a, b\} \\ \iff & x \in \{a\} \text{ or } x \in \{b\} && \text{–(a)} \\ \iff & x = a \text{ or } x = b && \text{–1.2.4} \end{aligned}$$

□

Exercises 1.2:

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

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

#3. Let a, b and c be objects. Show that there exists a set A such that

$$\text{for all } x: \qquad x \in A \quad \Longleftrightarrow \quad (x = a \text{ or } x = b) \text{ or } x = c.$$

#4. Let A and B be sets. Prove that

(a) $A \subseteq A \cup B$.

(b) $A \cap B \subseteq A$.

(c) $A \setminus B \subseteq A$.

#5. Let A, B and C be sets. Show that there exists a set D such that

$$\text{for all } x: \qquad x \in D \quad \Longleftrightarrow \quad (x \in A \text{ or } x \in B) \text{ and } x \notin C.$$

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

1.3 Relations and Functions

Definition 1.3.1. *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.*

(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 1.3.2. *Let a, b, c, d, e and f be objects.*

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

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

Proof. (a): See Exercise 1.3.#1.

(b)

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

□

Theorem 1.3.3. *Let A and B be sets. Then there exists a set, denoted by $A \times B$, such that*

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

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

Example 1.3.4. Let $A = \{1,2\}$ and $B = \{2,3,5\}$. Then

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

Definition 1.3.5. *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*

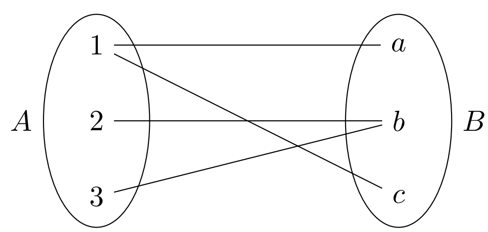
$$aRb \iff (a,b) \in T.$$

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

Example 1.3.6. (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 1.3.7. (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 $F(a)$. So*

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

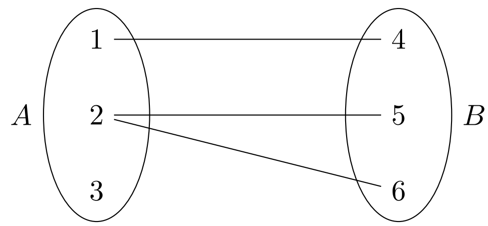
$F(a)$ is called the image of a under F. If $b = F(a)$ we will say that F maps a to b.

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

Example 1.3.8. (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 $x \in \mathbb{R}$.

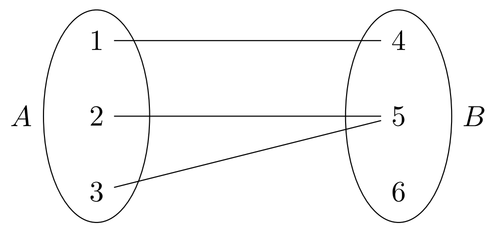
(b) $F = \big(\mathbb{R}, \mathbb{R}, \{(x^2, x^3) \mid x \in \mathbb{R}\}\big)$ is the relation with $x^2 F x^3$ for all $x \in \mathbb{R}$. For $x = 1$ we see that $1F1$ and for $x = -1$ we see that $1F-1$. So F is not a function.

(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 exist 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 1.3.9. *A and B be sets and suppose that $\Phi(x)$ is a formula involving a variable x such that for all a in A*

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

Put

$$T := \Big\{\big(a, \Phi(a)\big) \;\Big|\; a \in A\Big\} \quad \text{and} \quad F := (A, B, T).$$

Then F is a function from A to B. We denote this function by

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

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

Example 1.3.10. (1) $F : \ \mathbb{R} \to \mathbb{R}$, $r \mapsto r^2$ denotes the function from $\mathbb{R}$ to $\mathbb{R}$ with $F(r) = r^2$ for all $r \in \mathbb{R}$.

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

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

Definition 1.3.11. *Let $F : A \to B$ be a function.*

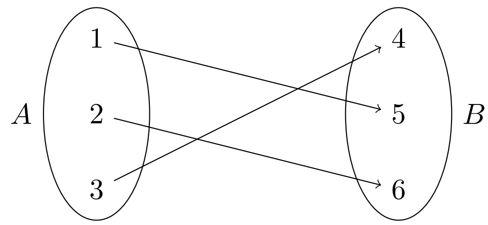
(a) *A is called the* domain *of F. B is called the* codomain *of F.*

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

(c) *F is called* surjective *(or* onto*) if for all $b \in B$ there exists (at least one) $a \in A$ with $b = F(a)$,*

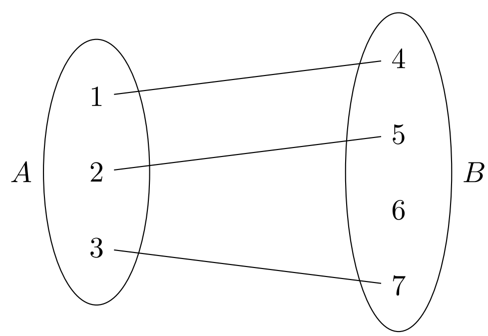
(d) *F is called* bijective *(or a* 1-1 correspondence*) if for all $b \in B$ there exists a unique $a \in A$ with $b = F(a)$*

Example 1.3.12. (1) The function



is bijection.

(2) The function



is injective but is neither surjective nor bijective.

Theorem 1.3.13. *Let $f : A \to B$ be a function.*

(a) *Then f is bijective if and only if f is a injective and surjective.*

(b) *f is injective if and only*

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

Proof. (a)

	f is bijective	
$\iff$	for all $b \in B$ there exists a unique $a \in A$ with $b = f(a)$	- Definition of bijective
$\iff$	for all $b \in B$ there exists at most one $a \in A$ with $b = f(a)$ and for all $b \in B$ there exists $a \in A$ with $b = f(a)$	- 1.1.7
$\iff$	f is injective and surjective	- Definition of injective and surjective

(b)

	f is injective	
$\iff$	for all $b \in B:$ there exists at most one $a \in A$ with $b = f(a)$	- Definition of injective
$\iff$	for all $b \in B, a, c \in A:$ $(b = f(a)$ and $b = f(c)) \implies a = c$	- Definition of 'exists at most one'
$\iff$	for all $a, c \in A:$ $f(a) = f(c) \implies a = c$	

□

Theorem 1.3.14. *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. See B.1.1 in the appendix. □

Definition 1.3.15. (a) *Let A be a set. The* identity function id_A *on A is the function*

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

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

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

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

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

Exercises 1.3:

#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. 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 $A = A_1 \cup A_2, A_1 \cap A_2 = \varnothing$, $B = B_1 \cup B_2$ and $B_1 \cap B_2 = \varnothing$. Let $\pi_1 : A_1 \to B_1$ and $\pi_2 : A_2 \to B_2$ be bijections. Define

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

Show that π is a bijection.

#3. Prove that the given function is injective

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

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

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

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

#4. Prove that the given function is surjective.

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

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

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

(d) $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q},\ (a, b) \mapsto \begin{cases} \frac{a}{b} & \text{when } b \neq 0 \\ 0 & \text{when } b = 0. \end{cases}$

#5. (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.

1.4 The Natural Numbers and Induction

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

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

It can be deduced from the Axioms of Set Theory that $\mathbb{N}$ is a set. 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:

$$
\begin{aligned}
0 &:= \varnothing \\
1 &:= \{0\} &&= 0 \cup \{0\} \\
2 &:= \{0,1\} &&= 1 \cup \{1\} \\
3 &:= \{0,1,2\} &&= 2 \cup \{2\} \\
4 &:= \{0,1,2,3\} &&= 3 \cup \{3\} \\
&\vdots \\
n+1 &:= \{0,1,2,3,\dots,n\} &&= n \cup \{n\} \\
&\vdots
\end{aligned}
$$

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

Definition 1.4.1. *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 1.4.2 (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 $0 \notin S$. As $m \in S$ this gives $m \neq 0$. Thus $k := m - 1$ is natural number. Note that $k < m$. 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 1.4.3 (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:

$$\text{for all } i \in \mathbb{N}: \qquad i < n \quad \Longrightarrow \quad P(i).$$

We will show that the two conditions in the Principal of Mathematical Induction hold for $Q(n)$ in place of $P(n)$. $Q(0)$ is statement

$$\text{for all } i \in \mathbb{N}: \qquad i < 0 \quad \Longrightarrow \quad P(i).$$

$i < 0$ is false for all $i \in \mathbb{N}$. Hence the implication $i < 0 \Longrightarrow P(i)$ is true for all $i \in \mathbb{N}$. Thus

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

Suppose now that $Q(k)$ is true for some $k \in \mathbb{N}$. Then $P(i)$ is a true for all $i \in \mathbb{N}$ with $i < k$. Then by (i), also $P(k)$ is true. If $i \in \mathbb{N}$ with $i < k + 1$, then either $i < k$ or $i = k$. In either case $P(i)$ is true. Thus $Q(k+1)$ is true. We proved

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

By $(*)$ and $(**)$ the Hypothesis of the Principal of Mathematical Induction is 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. □

Two more versions of the induction principal:

Theorem 1.4.4. *Suppose that $r \in \mathbb{Z}$ and that, for all $n \in \mathbb{Z}$ with $n \geq r$, a statement $P(n)$ is given. Also suppose that*

(i) *$P(r)$ is true, and*

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

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

Proof. See Exercise 1.4.#5. □

Theorem 1.4.5. *Suppose that $r \in \mathbb{Z}$ and that, for all $n \in \mathbb{Z}$ with $n \geq r$, a statement $P(n)$ is given. Also suppose that:*

(i) *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. See Exercise 1.4.#6. □

Exercises 1.4:

#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. Suppose that $r \in \mathbb{Z}$ and that, for all $n \in \mathbb{Z}$ with $n \geq r$, a statement $P(n)$ is given. Also suppose that

(i) $P(r)$ is true, and

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

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

#6. Suppose that $r \in \mathbb{Z}$ and that, for all $n \in \mathbb{Z}$ with $n \geq r$, a statement $P(n)$ is given. Also suppose that:

(i) 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.

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

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

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

Whenever A is set containing exactly n roses, then all roses in A have the same color.

If A is a set containing exactly one rose, then certainly all roses in A have the same color. Thus $P(1)$ is true.

Suppose now k is a positive integer such that $P(k)$ is true. So whenever D is a set containing exactly k roses then all roses in D have the same color. We need to show that $P(k+1)$ is true. So let A be any set containing exactly $k+1$-roses. Since $k \geq 1$ we have $k+1 \geq 2$. Hence A contains at least two roses and we can choose roses x and y in A with $x \neq y$. Consider the sets

$$B := A \setminus \{x\} \qquad \text{and} \qquad C := A \setminus \{y\}$$

Then B consist of all the elements of A other than x. Since A contains exactly $k+1$ roses, B contains exactly k roses. By the induction assumption $P(k)$ is true and so all roses in B have the same color. Similarly all roses in C have the same color.

Now let z be any rose in A distinct from x and y. Then $z \neq x$ and so $z \in B$. Also $z \neq y$ and so $z \in C$.

We will show that all roses in A have the same color as z. For this let a be any rose in A. We will distinguish the cases $a \neq x$ and $a = x$.

Suppose first that $a \neq x$. Then $a \in B$. Recall that $z \in B$ and all roses in B have the same color. Thus a has the same color as z.

Suppose next that $a = x$. Since $x \neq y$ this gives $a \neq y$ and so $a \in C$. Recall that $z \in C$ and all roses in C have the same color. Thus a has the same color as z.

Hence in either case a has the same color as z and so 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 positive integers n. Thus in any set of roses all the roses have the same color. So all roses have the same color. $\square$

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

1.5 Equivalence Relations

Definition 1.5.1. *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 1.5.2. (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 $2 \nleq 1$ 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 " ≠ " on any set A.

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

(4) Consider the relation S defined on $\mathbb{R}$ by

$$aSb \qquad \Longleftrightarrow \qquad a - b \in \mathbb{Z}.$$

Let $a, b, c \in \mathbb{R}$.

$a - a = 0 \in \mathbb{Z}$ and so aSa. Thus S is reflexive

If aSb, then $a - b \in \mathbb{Z}$. Hence also $-(a - b) \in \mathbb{Z}$. So $b - a \in \mathbb{Z}$. Thus bSa and so S is symmetric.

If aSb and bSc, then $a - b \in \mathbb{Z}$ and $b - c \in \mathbb{Z}$. Hence also $(a - b) + (b - c) \in \mathbb{Z}$. Thus $a - c \in \mathbb{Z}$ and S is transitive.

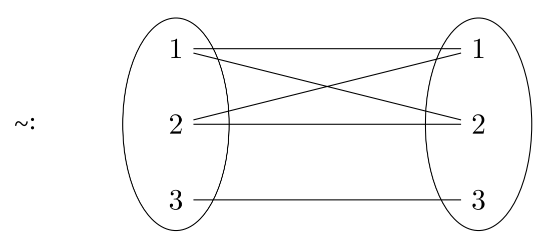
Since S is reflexive, symmetric and transitive, S is an equivalence relation.

Definition 1.5.3. *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 $\sim$.*

Example 1.5.4. (1) Consider the relation



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

$$\begin{aligned}
[1]_\sim &= \{a \in A \mid 1 \sim a\} = \{1,2\}\\
[2]_\sim &= \{a \in A \mid 2 \sim a\} = \{1,2\}\\
[3]_\sim &= \{a \in A \mid 3 \sim a\} = \{3\}
\end{aligned}$$

and so

$$A/\sim = \big\{\{1,2\},\{3\}\big\}$$

(2) Consider the relation S on $\mathbb{R}$ defined by

$$aSb \qquad \Longleftrightarrow \qquad a - b \in \mathbb{Z}.$$

By Example 1.5.2(4) S is an equivalence relation. We have

$$[0]_S = \{b \in \mathbb{R} \mid 0Sb\} = \{b \in \mathbb{R} \mid b - 0 \in \mathbb{Z}\} = \{b \in \mathbb{R} \mid b \in \mathbb{Z}\} = \mathbb{Z}$$

and

$$\begin{aligned}
[\pi]_S &= \{b \in \mathbb{R} \mid \pi S b\}\\
&= \{b \in \mathbb{R} \mid b - \pi \in \mathbb{Z}\}\\
&= \{b \in \mathbb{R} \mid b - \pi = k \text{ for some } k \in \mathbb{Z}\}\\
&= \{b \in \mathbb{R} \mid b = \pi + k \text{ for some } k \in \mathbb{Z}\}\\
&= \{\pi + k \mid k \in \mathbb{Z}\}\\
&= \{\dots, \pi - 4, \pi - 3, \pi - 2, \pi - 1, \pi, \pi + 1, \pi + 2, \pi + 3, \pi + 4, \dots\}
\end{aligned}$$

Theorem 1.5.5. *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) $[a] \cap [b] \neq \varnothing$.

(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 so $[a] \cap [b] \neq \varnothing$.

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

We will first show that $[a] \subseteq [b]$. We need to show that $d \in [b]$ for all $d \in [a]$. 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 1.2.3

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

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

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

Theorem 1.5.6. *Let $\sim$ be an equivalence relation on the set A and $a \in A$. Then there exists a unique equivalence class X of $\sim$ with $a \in X$, namely $X = [a]_\sim$.*

Proof. Let X be an equivalence class of $\sim$. We need to show that $a \in X$ if and only if $X = [a]_\sim$. By definition of 'equivalence class' we know that $X = [b]_\sim$ for some $b \in A$. We have

$$\begin{array}{lll} & a \in X & \\ \iff & a \in [b]_\sim & - \text{ Since } X = [b]_\sim \\ \iff & [a]_\sim = [b]_\sim & - \text{ 1.5.5} \\ \iff & [a]_\sim = X & - \text{ Since } X = [b]_\sim \end{array}$$

□

Exercises 1.5:

#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 1.3.5(a)) 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.

#5. Let A be a set and $\mathcal{B}$ a set of subsets of A. (So each element of $\mathcal{B}$ is a subset of A.) Suppose that for each $a \in A$ there exists a unique $B \in \mathcal{B}$ with $a \in B$. Define a relation $\sim$ on A by

$$a \sim b \iff \text{there exists } B \in \mathcal{B} \text{ with } a \in B \text{ and } b \in B.$$

Show that $\sim$ is an equivalence relation and that $\mathcal{B} = A/\sim$.

Chapter 2

Rings

2.1 Definitions and Examples

Definition 2.1.1. *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 $a = a + 0_R$ *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 2.1.2. *Let R be a ring. Then R is called* commutative *if*

(Ax 9) $ab = ba$ *for all* $a, b \in R$. *[commutative multiplication]*

Definition 2.1.3. *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 2.1.4. (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 $\mathbb{Z}_2 = \{0, 1\}$ and define an addition $\oplus$ and a multiplication $\odot$ on $\mathbb{Z}_2$ by

$\oplus$	0	1
0	0	1
1	1	0

and

$\odot$	0	1
0	0	0
1	0	1

Then $(\mathbb{Z}_2, \oplus, \odot)$ is a commutative ring with identity.

(f) Let $2\mathbb{Z}$ be the set of even integers. Then $(2\mathbb{Z}, +, \cdot)$ 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 2.1.5. Let $R = \{0, 1\}$ and $a, b \in R$. Define an addition and multiplication on R by

$+$	0	1
0	0	1
1	1	a

and

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

For which values of a and b is $(R, +, \cdot)$ a ring?

Note first that 0 is additive identity, so $0_R = 0$.

Case 1. *Suppose that $a = 1$:*

+	0	1
0	0	1
1	1	1

and

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

Then $1 + x = 1 \neq 0 = 0_R$ for all $x \in R$ and so 1 does not have an additive inverse. Hence R is not a ring.

Case 2. *Suppose that $a = 0$ and $b = 1$:*

+	0	1
0	0	1
1	1	1

and

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

Then $(R, +, \cdot)$ is $(\mathbb{Z}_2, \oplus, \odot)$ and so R is commutative ring with identity 1.

Case 3. *Suppose that $a = 0$ and $b = 0$:*

+	0	1
0	0	1
1	1	0

and

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

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. R is commutative, but does not have an identity.

Example 2.1.6. 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 2.1.7. *Let R and S be rings. Recall from 1.3.3 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 2.1.#3. □

Example 2.1.8. Determine the addition and multiplication table of the ring $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Recall from 2.1.4(b) that $\mathbb{Z}_2 = \{0,1\}$. So

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0),(0,1),(1,0),(1,1)\}$$

+	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,1)	(1,0)	(1,1)
(0,1)	(0,1)	(0,0)	(1,1)	(1,0)
(1,0)	(1,0)	(1,1)	(0,0)	(0,1)
(1,1)	(1,1)	(1,0)	(0,1)	(0,0)

and

·	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
(0,1)	(0,0)	(0,1)	(0,0)	(0,1)
(1,0)	(0,0)	(0,0)	(1,0)	(1,0)
(1,1)	(0,0)	(0,1)	(1,0)	(1,1)

Exercises 2.1:

#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. Below are parts of the addition table and parts of the multiplication table of a ring. Complete both tables.

$+$	w	x	y	z
w	w			
x		y	z	
y		z	w	x
z		w		y

$\cdot$	w	x	y	z
w				
x		y		
y				
z				

#3. Prove Theorem 2.1.7.

2.2 Elementary Properties of Rings

Theorem 2.2.1. *Let R be ring and $a, b \in R$. Then $(a + b) + (-b) = a$.*

Proof.

$$\begin{aligned} (a+b)+(-b) &= a+(b+(-b)) && -\textbf{Ax 2} \\ &= a+0_R && -\textbf{Ax 5} \\ &= a && -\textbf{Ax 4} \end{aligned}$$

□

Theorem 2.2.2 (Additive Cancellation Law)**.** *Let R be ring and $a, b, c \in R$. Then*

$$\begin{aligned} & a = b \\ \iff \quad & c + a = c + b \\ \iff \quad & a + c = b + c \end{aligned}$$

Proof. "First Statement $\Longrightarrow$ Second Statement": Suppose that $a = b$. Then $c + a = c + b$ by the Principal of Substitution 1.1.1.

"Second Statement $\Longrightarrow$ Third Statement": Suppose that $c + a = c + b$. Applying **Ax 3** to both sides yields $a + c = b + c$.

"Third Statement $\Longrightarrow$ First Statement": Suppose that $a + c = b + c$. Then the Principal of Substitution gives $(a + c) + (-c) = (b + c) + (-c)$. Applying 2.2.1 to both sides gives $a = b$. □

Definition 2.2.3. *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 $a \in R$.

Theorem 2.2.4 (Additive Identity Law)**.** *Let R be a ring and $a, c \in R$. Then*

$$\begin{array}{rrcl} & a & = & 0_R \\ \Longleftrightarrow & c + a & = & c \\ \Longleftrightarrow & a + c & = & c \end{array}$$

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:

$$\begin{array}{rclcrcl} a & = & 0_R & \Longleftrightarrow & a & = & b \\ c + a & = & c & \Longleftrightarrow & c + a & = & c + b \\ a + c & = & c & \Longleftrightarrow & a + c & = & b + c \end{array}$$

So the Theorem follows from the Cancellation Law 2.2.2. □

Definition 2.2.5. *Let R be a ring and $c \in R$. An* additive inverse *of c is an element a in R with $c + a = 0_R$.*

Theorem 2.2.6 (Additive Inverse Law)**.** *Let R be a ring and $a, c \in R$. Then*

$$\begin{array}{rrcl} & a & = & -c \\ \Longleftrightarrow & c + a & = & 0_R \\ \Longleftrightarrow & a + c & = & 0_R \end{array}$$

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:

$$\begin{aligned} a &= -c &\iff\quad a &= b \\ c+a &= 0_R &\iff\quad c+a &= c+b \\ a+c &= 0_R &\iff\quad a+c &= b+c \end{aligned}$$

So the Theorem follows from the Cancellation Law 2.2.2. □

Definition 2.2.7. *Let R be a ring and $a, b \in R$. Then $a - b := a + (-b)$. Note here $-b \in R$ by* **Ax 5** *and so $a - b = a + (-b) \in R$ by* **Ax 1**.

Theorem 2.2.8. *Let R be ring and $a, b, c \in R$. Then*

$$\begin{aligned} & c &= b - a \\ \iff\quad & c+a &= b \\ \iff\quad & a+c &= b \end{aligned}$$

Proof.

$$\begin{aligned} & a+c &=&\ b & \\ \iff\quad & c+a &=&\ b & -\ \textbf{Ax 3} \\ \iff\quad & (c+a)+(-a) &=&\ b+(-a) & -\ \text{Additive Cancellation Law 2.2.2} \\ \iff\quad & c &=&\ b-a & -\ \text{2.2.1 and Definition of } b-a \end{aligned}$$

□

Theorem 2.2.9. *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) $b - a = 0_R$ *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 2.2.6 $0_R = -0_R$.

(b) $a - 0_R \overset{\text{Def: -}}{=} a + (-0_R) \overset{\text{(a)}}{=} a + 0_R \overset{\textbf{Ax 4}}{=} a.$

(c) We compute

$$a \cdot 0_R \overset{\textbf{Ax 4}}{=} a \cdot (0_R + 0_R) \overset{\textbf{Ax 8}}{=} a \cdot 0_R + a \cdot 0_R,$$

and so by the Additive Identity Law 2.2.4 $a \cdot 0_R = 0_R$. Similarly $0_R \cdot a = 0_R$.

(d) We have

$$ab + a \cdot (-b) \overset{\textbf{Ax 8}}{=} a \cdot (b + (-b)) \overset{\text{Def } -b}{=} a \cdot 0_R \overset{\text{(c)}}{=} 0_R.$$

So by the Additive Inverse Law 2.2.6 $-(ab) = a \cdot (-b)$.

(e) By **Ax 5**, $a + (-a) = 0_R$ and so by the Additive Inverse Law 2.2.6, $a = -(-a)$.

(f) By Theorem 2.2.8 applied with $c = 0_R$:

$$0_R = b - a \qquad \Longleftrightarrow \qquad 0_R + a = b.$$

By **Ax 4** $0_R + a = a$ and so the Principal of Substitution gives

$$0_R = b - a \qquad \Longleftrightarrow \qquad a = b.$$

(g)

$$\begin{aligned}(a+b) + \big((-a) + (-b)\big) &\overset{\textbf{Ax 3}}{=} (b+a) + \big((-a) + (-b)\big) \overset{\textbf{Ax 2}}{=} ((b+a) + (-a)) + (-b) \\ &\overset{2.2.1}{=} b + (-b) \overset{\textbf{Ax 5}}{=} 0_R.\end{aligned}$$

and so by the Additive Inverse Law 2.2.6 $-(a+b) = (-a) + (-b)$. By definition of " − ", $(-a) + (-b) = (-a) - b$.

(h)

$$\begin{aligned}-(a-b) \overset{\text{Def -}}{=} -(a + (-b)) &\overset{\text{(g)}}{=} (-a) + (-(-b)) \overset{\text{(e)}}{=} (-a) + b \\ &\overset{\textbf{Ax 3}}{=} b + (-a) \overset{\text{Def -}}{=} b - a\end{aligned}.$$

(i) $(-a) \cdot (-b) \overset{\text{(d)}}{=} a \cdot (-(-b)) \overset{\text{(e)}}{=} a \cdot b.$

(j) $a \cdot (b - c) \overset{\text{Def -}}{=} a \cdot (b + (-c)) \overset{\textbf{Ax 8}}{=} a \cdot b + a \cdot (-c) \overset{\text{(d)}}{=} ab + (-(ac)) \overset{\text{Def -}}{=} ab - ac.$
Similarly $(a - b) \cdot c = ab - ac$.

(k) Suppose now that R has an additive identity. Then

$$a + ((-1_R) \cdot a) \overset{\textbf{(Ax 10)}}{=} 1_R \cdot a + (-1_R) \cdot a \overset{\textbf{Ax 8}}{=} (1_R + (-1_R)) \cdot a \overset{\textbf{Ax 5}}{=} 0_R \cdot a \overset{\text{(c)}}{=} 0_R.$$

Hence by the Additive Inverse Law 2.2.6 $-a = (-1_R) \cdot a$. Similarly, $-a = a \cdot (-1_R)$. □

Exercises 2.2:

#1. Let R be a ring and $a, b, c, d \in R$. Prove that

$$(a - b)(c - d) = \big((ac - ad) + bd\big) - bc$$

In each step of your proof, quote exactly one Axiom, Definition or Theorem.

#2. Prove or give a counterexample:

If R is a ring with identity, then $1_R \neq 0_R$.

#3. Let R be a ring such that $a \cdot a = 0_R$ for all $a \in R$. Show that $ab = -(ba)$ for all $a, b \in R$.

#4. Let R be a ring such that $a \cdot a = a$ for all $a \in R$. Show that

(a) $a + a = 0_R$ for all $a \in R$

(b) R is commutative.

2.3 The General Associative, Commutative and Distributive Laws in Rings

$\mathbb{Z}^+$ denotes the set of positive integers:

$$\mathbb{Z}^+ := \{n \in \mathbb{N} \mid n > 0\} = \{1, 2, 3, 4, 5, \ldots\}$$

Definition 2.3.1. *Let R be a ring, $n \in \mathbb{Z}^+$ 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 $\sum_{i=1}^{n} a_i = \Bigg(\Big(\ldots\big((a_1 + a_2) + a_3\big) + \ldots + a_{n-2}\Big) + a_{n-1}\Bigg) + a_n.$

  We will also write $a_1 + a_2 + \ldots + a_n$ for $\sum_{i=1}^{n} a_n$

(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 $1 \leq k < n$ and $x, y \in R$ such that*
    (i) *x is sum of $(a_1, \ldots, a_k)$ in R,*
    (ii) *y is a sum of $(a_{k+1}, a_{k+2}, \ldots, a_n)$ in R, and*
    (iii) *$z = x + y$.*

Example 2.3.2. Let R be a ring and $a, b, c, d \in R$. Find all sums of (a), (a, b), (a, b, c) and (a, b, c, d).

Sums of (a): We have $n = 1$ and a is the only sum of (a).

Sums of (a, b): Then $n = 2$ and $k = 1$. $a + b$ is the only sum of (a, b).

Sums of (a, b, c): We have $n = 3$ and $k = 1$ or 2. $a + (b + c)$ is the only sum with $k = 1$ and $(a + b) + c$ is the only sum with $k = 2$.

Sums of (a, b, c, d): We have $n = 4$ and $k = 1$, 2 or 3. $a+(b+(c+d))$ and $a+((b+c)+d)$ are the sums with $k = 1$, $(a + b) + (c + d)$ is the sum for $k = 2$ and $(a + (b + c)) + d$ and $((a + b) + c) + d$ are the only sums for $k = 3$.

We remark that the numbers of formal sums of an $n+1$-tuple is the n-th Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{2n!}{n!(n+1)!}$$

For example the number of formal sums of a 4-tuple is $C_3 = \frac{6!}{3!4!} = \frac{6\cdot 5}{6} = 5$.

Definition 2.3.3. *Let R be a ring, $n \in \mathbb{Z}^+$ and $a_1, a_2, \dots a_n \in R$.*

(a) *$\prod_{i=1}^{k} a_n$ is defined similarly as in 2.3.1(a), just replace '$\sum$' by '$\prod$' and '$+$' by '$\cdot$'. We will also write $a_1a_2\dots a_n$ for $\prod_{i=1}^{n} a_i$.*

(b) *A product of $(a_1, \dots, a_n)$ in R is defined similarly as in 2.3.1(b), just replace 'sum' by 'product' and '$+$' by '$\cdot$'.*

(c) *Let $a \in R$. Then $na := \sum_{i=1}^{n} a = \underbrace{a + a + \dots + a}_{n-times}$ and $a^n := \prod_{i=1}^{n} a = \underbrace{aa\dots a}_{n-times}$.*

(d) *If R has an identity and $a \in R$, then $a^0 := 1_R$.*

Theorem 2.3.4 (General Associative Law, GAL)**.** *Let R be a ring, $n \in \mathbb{Z}^+$ and $a_1, a_2, \dots, a_n$ elements of R. Then any sum of $(a_1, a_2, \dots, a_n)$ in R is equal to $\sum_{i=1}^{n} a_i$ and any product of $(a_1, a_2, \dots, a_n)$ is equal to $\prod_{i=1}^{n} a_i$*

Proof. See D.1.3 □

Theorem 2.3.5 (General Commutative Law,GCL)**.** *Let R be a ring, $a_1, a_2, \dots, a_n \in R$ and*

$$f : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$$

a bijection.

(a) *Any sum of $(a_1, \dots, a_n)$ is equal to any sum of $(a_{f(1)}, \dots, a_{f(n)})$. In particular,*

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{f(i)}.$$

(b) *Suppose that* R *is commutative. Then any product of* $(a_1, \ldots, a_n)$ *is equal to any product of* $(a_{f(1)}, \ldots, a_{f(n)})$. *In particular,*

$$\prod_{i=1}^{n} a_i = \prod_{i=1}^{n} a_{f(i)}.$$

Proof. See D.2.2 □

Theorem 2.3.6 (General Distributive Law,GDL)**.** *Let* R *be a ring,* $n, m \in \mathbb{Z}^+$ *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 2.3.7. 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 the General Distributive Law:

$$(a + b + c)(d + e) = (ad + ae) + (bd + be) + (cd + ce).$$

Exercises 2.3:

#1. Prove or give a counterexample:

Let R *be a ring and* $a, b \in R$. *Then*

$$(a + b)^2 = a^2 + 2ab + b^2.$$

(Note here that according to Definition 2.3.3(c) $2d = d + d$ for any d in R.)

#2. Let R be a commutative ring with identity. Suppose that $1_R + 1_R = 0_R$. Prove that

$$(a + b)^2 = a^2 + b^2.$$

for all $a, b \in R$.

#3. Let $S := \{a, b, c, d\}$ and let + be the addition on S defined by

+	a	b	c	d
a	b	a	d	a
b	c	d	b	a
c	d	a	c	c
d	b	d	b	c

Compute all possible sums of (a, b, c, d), where 'sum' is defined as in 2.3.1(b).

2.4 Divisibility and Congruence in Rings

Definition 2.4.1. *Let R be ring and $a, b \in R$. Then we say that a* divides *b in R and write $a \mid b$ if there exists $c \in R$ with $b = ac$* □

Example 2.4.2. (1) Does $7 \mid 133$ in $\mathbb{Z}$?

Yes, since $133 = 7 \cdot 19$.

(2) Does $2 \mid 3$ in $\mathbb{Z}$?

No since $2 \cdot k$ is even, $3 \neq 2k$ for all $k \in \mathbb{Z}$.

(3) Does $2 \mid 3$ in $\mathbb{Q}$?

Yes, since $3 = 2 \cdot \frac{3}{2}$.

(4) For which $a, b, c, d \in \mathbb{R}$ does $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Bigg| \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $\mathrm{M}_2(\mathbb{R})$?

Let $\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d} \in \mathbb{R}$, then

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} = \begin{bmatrix} \tilde{a} & \tilde{b} \\ 0 & 0 \end{bmatrix}$$

Hence

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Bigg| \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \Longleftrightarrow \qquad c = 0 \text{ and } d = 0$$

For example

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Bigg| \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \not\Bigg| \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Note that

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

So $ca = b$ does not imply that $a|b$.

Theorem 2.4.3. *Let R be a ring and $a \in R$.*

(a) $a|0_R$.

(b) $0_R|a$ *if and only of* $a = 0_R$.

(c) *If R has an identity, then* $1_R|a$.

Proof. (a) By 2.2.9(c), $0_R = a \cdot 0_R$ and so $a|0_R$.

(b) By (a) applied with $a = 0_R$ we have $0_R|0_R$.

Suppose now that $a \in R$ with $0_R\,|\,a$. Then there exists $b \in R$ with $a = 0_R b$ and so by 2.2.9(c), $a = 0_R$.

(c) By definition of an identity, $a = 1_R a$, and so $1_R\,|\,a$. □

Theorem 2.4.4. *Let R be a ring and $a, b, c, u, v \in R$.*

(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) *Suppose that $a|b$ and $a|c$. Then*

$$a|(b{+}c), \quad a|(b{-}c), \quad a|(bu{+}c), \quad a|(bu{-}c), \quad a|(bu{+}cv), \quad a|(b{+}cv), \quad a|(b{-}cv), \quad a|(bu{-}c).$$

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

$$(*) \qquad b = ar \quad \text{and} \quad c = bs.$$

Hence

$$c \overset{(*)}{=} bs \overset{(*)}{=} (ar)s \overset{\textbf{Ax 2}}{=} a(rs).$$

Since R is closed under multiplication, $rs \in R$ and so $a|c$ by definition of divide.

(b) We will first show

$$(**) \qquad a|b \quad \Longrightarrow \quad a|(-b) \quad \text{and} \quad (-a)|b.$$

Suppose that a divides b. Then by definition of 'divide' there exists $r \in R$ with $b = ar$. Thus

$$-b = -(ar) \overset{2.2.9(\mathrm{d})}{=} a(-r) \quad \text{and} \quad b = ar \overset{2.2.9(\mathrm{i})}{=} (-a)(-r)$$

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 2.2.9(e), $-(-b) = b$ and so $-a|b$.

Suppose that $(-a)|b$. Then by $(**)$ applied with $-a$ in place of a, $-(-a)|b$. By 2.2.9(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

$$b + c \overset{(***)}{=} ar + as \overset{\mathbf{Ax\ 8}}{=} a(r + s) \qquad \text{and} \qquad b - c \overset{(***)}{=} ar - as \overset{2.2.9(\mathrm{j})}{=} a(r - s).$$

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

$$(+) \qquad a|b + c \qquad \text{and} \qquad a|b - c.$$

By definition of 'divide', $b \,|\, bu$. Since $a \,|\, b$ we conclude from (a) that $a \,|\, bu$. Also $a \,|\, c$ and $(+)$ implies that

$$a|(bu + c) \qquad \text{and} \qquad a|(bu - c).$$

Similarly, as $a|c$ and $c|cv$ we have $a|cv$. Also $a|b$ and $(+)$ implies

$$a|(b + cv) \qquad \text{and} \qquad a|(b - cv).$$

Moreover, since $a|bu$ and $a|cv$ we get from $(+)$ that

$$a|(bu + cv) \qquad \text{and} \qquad a|(bu - cv).$$

□

Definition 2.4.5. *Let R be a ring and $n \in R$. Then the relation '$\equiv \pmod{n}$' on R 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 2.4.6. (1) Consider the ring $\mathbb{Z}$:

$6 \equiv 4 \pmod{2}$ 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.

(2) Let R be a ring and $a, b \in R$. Then

$$\begin{array}{lll} & a \equiv b \pmod{0_R} & \\ \iff & 0_R | a - b & \text{– Definition of '}a \equiv b \pmod{0_R}\text{'} \\ \iff & a - b = 0_R & \text{– 2.4.3(b)} \\ \iff & a = b & \text{– 2.2.9(f)} \end{array}$$

So congruence modulo 0_R is the equality relation.

(3) Let R be a ring with identity and $a, b \in R$. By 2.4.3(c) we have $1_R | a - b$ and so

$$a \equiv b \pmod{1_R} \qquad \text{for all } a, b \in R$$

(4) When is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} \left(\operatorname{mod} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right)$$

in $\mathrm{M}_2(\mathbb{R})$?

$$
\begin{aligned}
&\begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} \left(\operatorname{mod} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) \\
\iff\quad & \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Bigg| \begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} && \text{– definition of ‘ $\equiv$ ’} \\
\iff\quad & \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Bigg| \begin{bmatrix} a - \tilde{a} & b - \tilde{b} \\ c - \tilde{c} & d - \tilde{d} \end{bmatrix} \\
\iff\quad & c - \tilde{c} = 0 \quad \text{and} \quad d - \tilde{d} && \text{– see Example 2.4.2(4)} \\
\iff\quad & c = \tilde{c} \quad \text{and} \quad d = \tilde{d}
\end{aligned}
$$

Theorem 2.4.7. *Let R be a ring and $n \in R$. Then the relation '$\equiv \pmod n$' is an equivalence relation on R.*

Proof. We have to show that '$\equiv \pmod n$' is reflexive, symmetric and transitive. Let $a, b, c \in R$.

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

Symmetric: Suppose that $a \equiv b \pmod n$. Then $n | (a - b)$. By 2.4.4(b) this gives $n | -(a - b)$. By 2.2.9(h) we have $-(a - b) = b - a$. Hence $n | b - a$ and so $b \equiv a \pmod n$. Thus '$\equiv \pmod n$' is symmetric.

Transitive: Suppose that $a \equiv b \pmod n$ and $b \equiv c \pmod n$. Then $n | a - b$ and $n | b - c$. Thus 2.4.4(c) shows that

$$n | (a - b) + (b - c).$$

We compute

$$
\begin{aligned}
(a - b) + (b - c) &= \big(a + (-b)\big) + \big(b + (-c)\big) && \text{– definition of ‘ $-$ ’} \\
&= \Big(\big(a + (-b)\big) + b\Big) + (-c) && \text{– } \mathbf{Ax\ 2} \\
&= \Big(\big(a + (-b)\big) + \big(-(-b)\big)\Big) + (-c) && \text{– 2.2.9(e)} \\
&= a + (-c) && \text{– 2.2.1} \\
&= a - c && \text{– definition of ‘ $-$ ’}
\end{aligned}
$$

Hence $n | a - c$ and $a \equiv c \pmod n$. Thus '$\equiv \pmod n$' is transitive. □

Definition 2.4.8. *Let R be a ring and $n \in R$. Recall from 2.4.7 that the relation '$\equiv \pmod n$' is an equivalence relation.*

(a) *For $a \in R$ we denote the equivalence class of '$\equiv \pmod{n}$' containing a by $[a]_n$. So*

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

$[a]_n$ is called the congruence class of a modulo n.

(b) *R_n denotes the set of equivalence classes of '$\equiv \pmod{n}$'. So*

$$R_n = \{ [a]_n \mid a \in R\}.$$

Theorem 2.4.9. *Let R be a ring and $a, b, n \in R$. Then the following statements are equivalent*

(a) *$a = b + nk$ for some $k \in R$*

(b) *$a - b = nk$ for some $k \in R$.*

(c) *$n | a - b$.*

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

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

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

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

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

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

(j) *$n | b - a$.*

(k) *$b - a = nl$ for some $l \in R$.*

(l) *$b = a + nl$ for some $l \in R$.*

Proof. (a) $\iff$ (b): See 2.2.8.

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

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

By 2.4.7 '$\equiv \pmod{n}$' is an equivalence relation. So Theorem 1.5.5 implies that (d)-(i) are equivalent.

Applying the fact that statements (a) to (d) are equivalent with a and b interchanged, shows that (i) to (l) are equivalent.

We proved that (a)-(d) are equivalent, that (d) to (i) are equivalent and that (i) to (l) are equivalent. Hence (a)-(l) are equivalent. □

Theorem 2.4.10. *Let R be a ring and $a, n \in R$. Then*

$$[a]_n = \{a + nl \mid l \in R\}.$$

Proof. Let $b \in R$. Then

$$\begin{aligned} & b \in [a]_n \\ \iff\ & b = a + nk \text{ for some } l \in R && - 2.4.9 \\ \iff\ & b \in \{a + nl \mid l \in R\} && - \text{Definition of } \{a + nk \mid k \in R\} \end{aligned}$$

Hence $[a]_n = \{a + nl \mid l \in R\}$ by 1.2.1. □

Example 2.4.11. (1) Consider the ring $\mathbb{Z}$.

$$[3]_5 = \{3 + 5k \mid k \in \mathbb{Z}\} = \{\ldots, -12, -7 - 2, 3, 8, 13, 18, \ldots\}.$$

(2) Consider the ring $\mathbb{Q}$:

$$[3]_5 = \{3 + 5k \mid k \in \mathbb{Q}\} = \{3 + l \mid l \in \mathbb{Q}\} = \mathbb{Q}.$$

(3) Consider the ring $\mathrm{M}_2(\mathbb{R})$.

$$\begin{aligned}
\left[\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right]_{\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}} &= \left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \cdot A \;\middle|\; A \in \mathrm{M}_2(\mathbb{R})\right\}\\
&= \left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \cdot \begin{bmatrix}a & b\\ c & d\end{bmatrix} \;\middle|\; a, b, c, d \in \mathbb{R}\right\}\\
&= \left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\ c & d\end{bmatrix} \;\middle|\; a, b, c, d \in \mathbb{R}\right\}\\
&= \left\{\begin{bmatrix}1 & 0\\ c & d\end{bmatrix} \;\middle|\; c, d \in \mathbb{R}\right\}
\end{aligned}$$

Exercises 2.4:

#1. Consider the ring $\mathrm{M}_2(\mathbb{R})$.

(a) Does $\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$ divide $\begin{bmatrix}2 & 3\\ 2 & 3\end{bmatrix}$ in $\mathrm{M}_2(\mathbb{R})$?

(b) Does $\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$ divide $\begin{bmatrix}2 & 3\\ 3 & 2\end{bmatrix}$ in $\mathrm{M}_2(\mathbb{R})$?

(c) Compute $\left[\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right]_{\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}}$.

(d) Let $a, b, c, d, \tilde{a}, \tilde{b}, \tilde{c}, \tilde{d} \in \mathbb{R}$. Show that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} \left(\operatorname{mod} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)$$

if and only of

$$a - c = \tilde{a} - \tilde{c} \qquad \text{and} \qquad b - d = \tilde{b} - \tilde{d}.$$

2.5 Congruence in the ring of integers

For a general ring it is difficult to explicitly determine all the equivalence classes of relation $\equiv \pmod n$. But thanks to the division algorithm it is fairly easy for the ring of integers.

Theorem 2.5.1 (The Division Algorithm)**.** *Let a and b be integers with $b > 0$. Then there exist unique integers q and r such that*

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

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

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

Note that $S \subseteq \mathbb{N}$. We would like to apply the Well-Ordering Axiom C.4.2 to S, so we need to verify that S is not empty. That is we need to find $x \in \mathbb{Z}$ such that $a - bx \ge 0$.

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

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

We proved that S is non-empty subset of $\mathbb{N}$. Hence by the Well-ordering Axiom C.4.2 S has a minimal element r. Thus

$$(*) \qquad r \in S \qquad \text{and} \qquad r \le s \quad \text{for all } s \in S.$$

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 $0 \le r < b$. Since $r \in S$, $r \ge 0$. Suppose for a contradiction that $r \ge b$. Then $r - b \ge 0$. Hence

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

and $q + 1 \in \mathbb{Z}$. Thus $r - b \in S$. Since r is a minimal element of S this implies $r \le r - b$, see $(*)$. It follows that $b \le 0$, a contradiction since $b > 0$ by the hypothesis of the Theorem.

This contradiction shows that $r < b$, so the existence assertion in the Theorem is proved. To show the uniqueness let $q, r, \tilde{q}$ and $\tilde{r}$ be integers with

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

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

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

By $(**)$ we have $0 \le r < b$. Multiplying with -1 gives $0 \ge -r > -b$ and so

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

By $(**)$

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

and adding the last two equations yields

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

By $(***)$ we have $b(q - \tilde{q}) = \tilde{r} - r$. Thus

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

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 $\tilde{r} = r$. □

Theorem 2.5.2 (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 2.5.#1 □

Definition 2.5.3. *Let a and b be integers with $b \neq 0$. According to the Division Algorithm there exist unique integers q and r with $a = bq + r$ and $0 \le r < |b|$. Then r is called the* remainder *of a when divided by b in $\mathbb{Z}$. q is called the integral quotient of a when divided by b in $\mathbb{Z}$.*

Example 2.5.4. (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.

Theorem 2.5.5. *Let a, b, n be integers with $n \neq 0$. Then*

$$a \equiv b \pmod{n}$$

if and only if

a and b have the same remainder when divided by n.

Proof. By the division algorithm there exist integers q_1, r_1, q_2, r_2 with

$$(*) \qquad a = nq_1 + r_1 \quad \text{and} \quad 0 \le r_1 < |n|$$

and

$$(**) \qquad b = nq_2 + r_2 \quad \text{and} \quad 0 \le r_2 < |n|.$$

So r_1 and r_2 are the remainders of a and b, respectively when divided by n in $\mathbb{Z}$.

$\Longrightarrow$: Suppose $a \equiv b \pmod{n}$. Then by 2.4.9 we have $a = b + nk$ for some integer k. Then

$$a = b + nk \overset{(*)}{=} (nq_2 + r_2) + nk = n(q_2 + k) + r_2.$$

Since $q_2 + k \in \mathbb{Z}$ and $0 \le r_2 < |n|$, we conclude that r_2 is the remainder of a when divided by n. So $r_1 = r_2$ and a and b have the same remainder when divided by n.

$\Longleftarrow$: Suppose a and b have the same remainder then divided by n. Then $r_1 = r_2$ and so

$$a - b \overset{(*)(**)}{=} (nq_1 + r_1) - (nq_2 + r_2) = n(q_1 - q_2) + (r_1 - r_2) = n(q_1 - q_2).$$

Thus $n \mid a - b$ and so $a \equiv b \pmod{n}$. □

Theorem 2.5.6. *Let n be positive integer.*

(a) *Let $a \in \mathbb{Z}$. Then there exists a unique $r \in \mathbb{Z}$ with $0 \le r < n$ 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 $0 \le r < n$. We need to show that $[a]_n = [r]_n$ if and only if $r = s$.

Since $r = n0 + r$ and $0 \le r < n$, we see that r is the remainder of r when divided by n. By 2.5.5, $[a]_n = [r]_n$ 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 2.5.7. Determine $\mathbb{Z}_5$.

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

Exercises 2.5:

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

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

#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.

#6. (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.

2.6 Modular Arithmetic in Commutative Rings

Theorem 2.6.1. *Let R be a commutative ring and $a, \tilde{a}, b, \tilde{b}$ and n elements of R. 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 2.4.9 that

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

for some $k, l \in R$. Hence

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

Since $k + l \in R$, 2.4.9 gives

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

Also

$$\begin{aligned} \tilde{a}\cdot\tilde{b} &= (a+nk)(b+nl) = ab + anl + nkb + nknl \\ &\overset{\textbf{(Ax 9)}}{=} ab + nal + nkb + nknl = ab + n(al + kb + knl), \end{aligned}$$

and, since $al + kb + knl \in R$, 2.4.9 implies

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

□

In view of 2.6.1 the following definition is well-defined.

Definition 2.6.2. *Let R be commutative ring and a, b and n elements of R. Define*

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

The function

$$R_n \times R_n \to R_n, \quad (A, B) \mapsto A \oplus B$$

is called the addition *on R_n, and the function*

$$R_n \times R_n \to R_n, \qquad (A, B) \to A \odot B$$

is called the multiplication *on R_n.*

Example 2.6.3. (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 2.6.1 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

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

Theorem 2.6.4. *Let R be a commutative ring and $n \in R$.*

(a) $(R_n, \oplus, \odot)$ *is a commutative ring.*

(b) $0_{R_n} = [0_R]_n$.

(c) $-[a]_n = [-a]_n$ *for all* $a \in R$.

(d) *If* R *has an identity, then* $[1_R]_n$ *is an identity for* R_n.

Proof. We need to verify the eight Axioms of a ring. If $d \in R$ we will just write $[d]$ for $[d]_n$.
Let $A, B, C \in R_n$. By definition of R_n there exist a, b and c in R with

$$(*) \qquad A = [a], \quad B = [b], \quad \text{and} \quad C = [c].$$

Ax 1: We have

$$\begin{aligned} A \oplus B &= [a] \oplus [b] && -(*) \\ &= [a+b] && -\text{Definition of } \oplus \end{aligned}$$

Since $a + b \in R$ we conclude that $A \oplus B \in R_n$.

Ax 2:

$$\begin{aligned} A \oplus (B \oplus C) &= [a] \oplus ([b] \oplus [c]) && -(*) \\ &= [a] \oplus [b+c] && -\text{Definition of } \oplus \\ &= [a + (b+c)] && -\text{Definition of } \oplus \\ &= [(a+b)+c] && -\textbf{Ax 2} \\ &= [a+b] \oplus [c] && -\text{Definition of } \oplus \\ &= ([a] \oplus [b]) \oplus [c] && -\text{Definition of } \oplus \\ &= (A \oplus B) \oplus C. && -(*) \end{aligned}$$

Ax 3:

$$\begin{aligned} A \oplus B &= [a] \oplus [b] && -(*) \\ &= [a+b] && -\text{Definition of } \oplus \\ &= [b+a] && -\textbf{Ax 2} \\ &= [b] \oplus [a] && -\text{Definition of } \oplus \\ &= B \oplus A. && -(*) \end{aligned}$$

Ax 4: Define

$$0_{R_n} := [0_R] \tag{$**$}$$

Then

$$\begin{aligned} A \oplus 0_{R_n} &= [a] \oplus [0_R] && - (*) \text{ and } (**) \\ &= [a + 0_R] && - \text{Definition of } \oplus \\ &= [a] && - \textbf{Ax 4} \\ &= A && - (*) \end{aligned}$$

and

$$\begin{aligned} 0_{R_n} \oplus A &= [0_R] \oplus [a] && - (**) \text{ and } (*) \\ &= [0_R + a] && - \text{Definition of } \oplus \\ &= [a] && - \textbf{Ax 4} \\ &= A && - (*) \end{aligned}$$

and so **Ax 4** holds.

Ax 5: Put

$$-A := [-a] \tag{$***$}$$

Then $-A \in R_n$ and

$$\begin{aligned} A \oplus -A &= [a] \oplus [-a] && - (*) \text{ and } (***) \\ &= [a + (-a)] && - \text{Definition of } \oplus \\ &= [0_R] && - \textbf{Ax 4} \\ &= 0_{R_n} && - (**) \end{aligned}$$

and so **Ax 4** holds.

Ax 6: Similarly to **Ax 1** we have $A \odot B = [a] \odot [b] = [ab]$ and so $A \odot B \in R_n$.

Ax 7: Similarly to **Ax 2** we can use the definition of $\odot$ and the fact that multiplication in R is associative to compute

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

Ax 8:

$$\begin{aligned} A \odot (B \oplus C) &= [a] \odot ([b] \oplus [c]) && -(*) \\ &= [a] \odot [b+c] && - \text{Definition of } \oplus \\ &= [a(b+c)] && - \text{Definition of } \odot \\ &= [ab+bc] && - \textbf{Ax 8} \\ &= [ab] \oplus [ac] && - \text{Definition of } \oplus \\ &= [a] \odot [b]) \oplus ([a] \odot [c]) && - \text{Definition of } \odot \\ &= (A \odot B) \oplus (A \odot C) && -(*) \end{aligned}$$

and similarly

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

Ax 9 Similarly to **Ax 3** we can use the definition of $\odot$ and the fact that multiplication in R is commutative to compute

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

Ax 10 Suppose R is a ring with identity. Put

(+) $$1_{R_n} := [1_R]_n$$

Similarly to **Ax 4** we can use the definition of $\odot$ and the fact that 1_R is a multiplicative identity in R to compute

$$A \odot 1_{R_n} = A \odot [1_R] = [a] \odot [1_R] = [a1_R] = [a] = A,$$

and

$$1_{R_n} \odot A = [1_R] \odot A = [1_R] \odot [a] = [1_R a] = [a1_R] = A.$$

□

Theorem 2.6.5. *Let R be a commutative ring, $a, n \in R$ and $k \in \mathbb{Z}^+$. Then $[a]_n^k = [a^k]_n$*

Proof. The proof is by induction on k. We have $[a]^1 = [a] = [a^1]$ and so statement holds for $k = 1$. Suppose the statement holds for k, that is

$$(*) \qquad [a^k] = [a^k]$$

Then

$$\begin{aligned} [a]^{k+1} &= [a]^k \odot [a] && - \text{Definition of } [a]^{k+1},\ 2.3.3 \\ &= [a^k] \odot [a] && - \text{Induction assumption } (*) \\ &= [a^k a] && - \text{Definition of } \odot,\ 2.6.2 \\ &= [a^{k+1}] && - \text{Definition of } a^{k+1} \end{aligned}$$

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

Notation 2.6.6. *Let R be a ring and $a, b, n \in R$. . 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.*

Remark 2.6.7. *Consider the expression*

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

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

$$\begin{gathered} [2^5 + 3 \cdot 7]_n \\ [2^5]_n \oplus [3 \cdot 7]_n \\ [2^5]_n \oplus ([3]_n \odot [7]_n) \\ [2]_n^5 \oplus [3 \cdot 7]_n \\ [2]_n^5 \oplus ([3]_n \odot [7]_n) \end{gathered}$$

But thanks to Theorem 2.6.1 and Theorem 2.6.5 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 2.6.1 and Theorem 2.6.5.

Example 2.6.8. (1) Compute $[13^{34567}]_{12}$ in $\mathbb{Z}_{12}$.

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

In simplified notation this becomes: In $\mathbb{Z}_{12}$:

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

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 $\mathbb{Z}_{50}$.

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 $0 \le 46 < 50$, 2.5.6(a) shows that the remainder in question is 46.

(4) Let Fun($\mathbb{R}$) be the set of functions from $\mathbb{R}$ to $\mathbb{R}$. Define an addition and multiplication on Fun(R) by

$$(f+g)(a) = f(a) + g(a) \qquad \text{and} \qquad (fg)(a) = f(a)g(a).$$

for all $f, g \in$ Fun(R) and $a \in \mathbb{R}$. Given that (Fun($\mathbb{R}$), +·) is a ring (see Exercise 2.6.#1) . Compute

$$[\sin x]^2_{\cos x}.$$

In $\text{Fun}(\mathbb{R})_{\cos x}$:

$$\sin^2 x = 1 - \cos^2 x = 1 - 0^2 = 1$$

So $[\sin x]^2_{\cos x} = [1]_{\cos x}$.

Exercises 2.6:

#1. Let R be a ring and I a set. Let $\text{Fun}(I, R)$ be the set of functions from I to R. For $f, g \in \text{Fun}(I, R)$ let $f + g$ and $f \cdot g$ be the functions from I to R defined by

$$(f + g)(i) = f(i) + g(i) \quad \text{and} \quad (f \cdot g)(i) = f(i) \cdot g(i).$$

for all $i \in I$. Show that

(a) $(\text{Fun}(I, R), +, \cdot)$ is a ring.

(b) If R has an identity, then $\text{Fun}(I, R)$ has an identity.

(c) If R is commutative, then $\text{Fun}(I, R)$ is commutative.

2.7 Subrings

Definition 2.7.1. *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 2.7.2 (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 the Additive Identity Law 2.2.4 implies that

$$(*) \qquad 0_S = 0_R.$$

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 $s+t = 0_R$. The Additive Inverse Law 2.2.6 shows that $t = -s$. 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** holds 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.

We proved that **Ax 1**-**Ax 8** hold for S and thus S is a ring. Hence, by definition, S is a subring of R. □

Example 2.7.3. (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 2.1.4 $\mathbb{Z}$, $\mathbb{Q}$ 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 R be a ring and $n \in R$. Put $nR := \{nk \mid k \in R\}$. Show that nR is subring of R.

We will verify the four conditions of the Subring Theorem for $S = nR$.

Observe first that since $nR = \{nk \mid k \in R\}$,

$$(*) \qquad a \in nR \qquad \Longleftrightarrow \qquad \text{there exists } k \in R \text{ with } a = nk.$$

Let $a, b \in nR$. Then by $(*)$

$$(**) \qquad a = nk \quad \text{and} \quad b = nl,$$

for some $k, l \in R$.

(I): $0 = n0$ and so $0 \in nR$ by $(*)$

(II): $a + b \overset{(**)}{=} nk + nl = n(k + l)$. Since $k + l \in R$, $(*)$ shows $a + b \in R$. So nR is closed under addition.

(III): $ab \overset{(**)}{=} (nk)(nl) = n(knl)$. Since $nkl \in R$, $(*)$ shows $ab \in R$. So nR is closed under multiplication.

(IV): $-a \overset{(**)}{=} -(nk) = n(-k)$. Since $-k \in R$, $(*)$ shows $-a \in R$. So nR is closed under negatives.

Thus all four conditions of the Subring Theorem Hold and so nR is a subring of R.

(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. Moreover,

+	0	2
0	0	2
2	2	0

,

·	0	2
0	0	0
2	0	0

and

x	0	2
$-x$	0	2

So $\{0, 2\}$ is closed under addition, multiplication and negatives. Thus $\{0, 2\}$ is a subring of $\mathbb{Z}_4$ by Subring Theorem.

Exercises 2.7:

#1. Which of the following nine sets are subrings of $M_2(\mathbb{R})$? Which ones have an identity? (You don't need to justify your answers)

(1) $\left\{ \begin{bmatrix} 0 & r \\ 0 & 0 \end{bmatrix} \middle| \, r \in \mathbb{Q} \right\}$.

(2) $\left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \middle| \, a, b, c \in \mathbb{Z} \right\}$.

(3) $\left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \middle| \, a \in \mathbb{Z}, b \in \mathbb{Q} \right\}$.

(4) $\left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \middle| \, a \in \mathbb{Q}, b \in \mathbb{Z} \right\}$.

(5) $\left\{ \begin{bmatrix} a & a \\ b & b \end{bmatrix} \middle| \, a, b \in \mathbb{R} \right\}$.

(6) $\left\{ \begin{bmatrix} a & 0 \\ a & 0 \end{bmatrix} \middle| \, a \in \mathbb{R} \right\}$.

(7) $\left\{ \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix} \middle| \, a \in \mathbb{R} \right\}$.

(8) $\left\{ \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \middle| \, a \in \mathbb{R} \right\}$.

(9) $\left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \middle| \, a \in \mathbb{R} \right\}$.

#2. Let $\mathbb{Z}[i]$ denote the set $\{a + bi \mid a, b \in \mathbb{Z}\}$. Show that $\mathbb{Z}[i]$ is a subring of $\mathbb{C}$.

#3. Let R be a ring and S and T subrings of R. Show that $S \cup T$ is a subring of R if and only if $S \subseteq T$ or $T \subseteq S$.

2.8 Units in Rings

Definition 2.8.1. *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*

$$uu^{-1} = 1_R \quad \text{and} \quad u^{-1}u = 1_R$$

(b) *Let $u, v \in R$. Then v is called an (multiplicative) inverse of u if $uv = 1_R$ and $vu = 1_R$.*

(c) *Let $e \in R$. Then e is called an (multiplicative) identity of R, if $ea = a$ and $ae = a$ for all $a \in R$.*

Example 2.8.2. (1) Units in $\mathbb{Z}$: Let u be a unit in $\mathbb{Z}$. Then $uv = 1$ for some $v \in \mathbb{Z}$. Thus $u = \pm 1$.

(2) Units in $\mathbb{Q}$: Let u be a non-zero rational number. Then $u = \frac{n}{m}$ for some $n, m \in \mathbb{Z}$ 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.

(3) Units in $\mathbb{Z}_8$: By 2.5.6 $\mathbb{Z}_8 = \{0, 1, 2, 3, 4, 5, 6, 7\}$ and so $\mathbb{Z}_8 = \{0, \pm 1, \pm 2, \pm 3, 4\}$. We compute

$\cdot$	0	± 1	± 2	± 3	4
0	0	0	0	0	0
± 1	0	$\boxed{\pm 1}$	± 2	± 3	4
± 2	0	± 2	4	± 2	0
3	0	± 3	± 2	$\boxed{\pm 1}$	4
4	0	4	0	4	0

So $\pm 1, \pm 3$ (that is $1, 3, 5, 7$) are the units in $\mathbb{Z}_8$.

Theorem 2.8.3. (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. In particular, u has at most one inverse in R.

Proof. (a)

$$e \overset{(*)}{=} ee' \overset{(**)}{=} e'.$$

(b)

$$y \overset{\textbf{(Ax 10)}}{=} 1_R y \overset{(+)}{=} (xu)y \overset{\textbf{Ax 7}}{=} x(uy) \overset{(++)}{=} x1_R \overset{\textbf{(Ax 10)}}{=} x.$$

□

Theorem 2.8.4 (Multiplicative Inverse Law)**.** *Let R be a ring with identity and $u, v \in R$. Suppose u is a unit. Then*

$$\begin{array}{rrcl} & v & = & u^{-1} \\ \iff & vu & = & 1_R \\ \iff & uv & = & 1_R \end{array}$$

Proof. Recall first that by definition of a 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 (*) $uu^{-1} = 1_R$.

$$vu = 1_R \qquad \text{and} \qquad uu^{-1} = 1_R$$

and 2.8.3(b) applied with $x = v$ and $y = u^{-1}$ gives $v = u^{-1}$. Thus $uv = uu^{-1} \overset{(*)}{=} 1_R$.

'Third Statement $\Longrightarrow$ First Statement': Suppose that $uv = 1_R$. By (**) $u^{-1}u = 1_R$. Hence

$$u^{-1}u = 1_R \qquad \text{and} \qquad uv = 1_R$$

and 2.8.3(b) applied with $x = u^{-1}$ and $y = v$ gives $u^{-1} = v$. □

Theorem 2.8.5. *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}, $aa^{-1} = 1_R$ and $a^{-1}a = 1_R$. Hence also $a^{-1}a = 1_R$ and $aa^{-1} = 1_R$. Thus a^{-1} is a unit and by the Multiplicative Inverse Law 2.8.4, $a = (a^{-1})^{-1}$.

(b) See Exercise 2.8.#5. □

Definition 2.8.6. *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 2.8.7 (Multiplicative Cancellation Law for Integral Domains)**.** *Let R be an integral domain and* $a, b, c \in R$ *with* $a \neq 0_R$*. Then*

$$\begin{array}{rrcl} & ab & = & ac \\ \iff & b & = & c \\ \iff & ba & = & ca \end{array}$$

Proof. 'First Statement $\Longrightarrow$ Second Statement:'' Suppose $ab = ac$. Then

$$\begin{aligned} a(b-c) &= ab - ac && 2.2.9(j) \\ &= ab - ab && \text{Principal of Substitution}, ab = ac \\ &= 0_R && 2.2.9(f) \end{aligned}$$

Since R is an integral domain, **(Ax 11)** holds. As $a(b-c) = 0_R$ this implies $a = 0_R$ or $b - c = 0_R$. By assumption $a \neq 0_R$ and so $b - c = 0_R$. Thus by 2.2.9(f), $b = c$.

'Second Statement $\Longrightarrow$ Third Statement:'' If $b = c$, then $ba = ca$ by the Principal of Substitution.

'Third Statement $\Longrightarrow$ First Statement:'' Since integral domains are commutative, we have $ab = ba$ and $ac = ca$. Thus $ba = ca$ implies $ab = ac$. □

Definition 2.8.8. *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 2.8.9. Which of the following rings are fields? Which are integral domains?

(1) $\mathbb{Z}$.

(2) $\mathbb{Q}$.

(3) $\mathbb{R}$.

(4) $\mathbb{Z}_3$.

(5) $\mathbb{Z}_8$.

(6) $\mathrm{M}_2(\mathbb{R})$.

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.

(1) 2 does not have an inverse in $\mathbb{Z}$. So $\mathbb{Z}$ is an integral domain, but not a field.

(2) The inverse of a non-zero rational numbers is rational. So $\mathbb{Q}$ is an integral domain and a field.

(3) The inverse of a non-zero real numbers is real. So $\mathbb{R}$ is an integral domain and a field.

(4) ± 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.

(5) By Example 2.8.2 the units in $\mathbb{Z}_8$ are ± 1 and ± 3. Thus 2 is not a unit and so $\mathbb{Z}_8$ is not a field. Note that $2 \cdot 4 = 8 = 0$ in $\mathbb{Z}_8$ and so $\mathbb{Z}_8$ is not an integral domain

(6) 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. Since

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1_{\mathrm{M}_2(\mathbb{R})}$$

$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not a unit so **(Ax 12)** fails. Also $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0_{\mathrm{M}_2(\mathbb{R})}$ and so **(Ax 11)** fails. Thus $\mathrm{M}_2(\mathbb{R})$ fails all conditions of a field and integral domain, except for $1_R \neq 0_R$.

Theorem 2.8.10. *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 to verify Ax 11 in 2.8.6. 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

$$0_F \overset{2.2.9(c)}{=} a^{-1} \cdot 0_F \overset{(*)}{=} a^{-1} \cdot (a \cdot b) \overset{\mathbf{Ax\ 7}}{=} (a^{-1} \cdot a) \cdot b \overset{\text{Def: } a^{-1}}{=} 1_F \cdot b \overset{\mathbf{(Ax\ 10)}}{=} b.$$

So $b = 0_F$.

We have proven that $a \neq 0_F$ implies $b = 0_F$. So $a = 0_F$ or $b = 0_F$. Hence Ax 11 holds and F is an integral domain. □

Theorem 2.8.11. *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 in R. Put

$$S := \{ar \mid r \in R\}.$$

and define

$$f: \ R \to S, \quad r \mapsto ar.$$

We will show that f is a bijection. Let $b, c \in R$ with $f(b) = f(c)$. Then $ab = ac$. As $a \neq 0_R$ the Multiplicative Cancellation Law for Integral Domains 2.8.7 gives $b = c$. Thus f is injective. Let $s \in S$. The definition of S implies that $s = ar$ for some $r \in R$. Then $f(r) = ar = s$ and f is surjective. 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 2.8.12. *Let R be a ring with identity, $a \in R$ and $n \in \mathbb{Z}^+$. Then*

$$a^{-n} := (a^{-1})^n.$$

Exercises 2.8:

#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. Find all units in $\mathrm{Fun}(\mathbb{R}, \mathbb{R})$.

#3. An element e of a ring is said to be an **idempotent** if $e^2 = e$.

(a) Find four idempotents in $\mathrm{M}_2(\mathbb{R})$.

(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.

#4. Prove or give a counter example:

(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.

#5. (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.

#6. Let R be a ring with identity. If ab and a are units in R, prove that b is a unit.

#7. 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 $b = c$.)

#8. Let R be a ring with identity and $a, b, c \in R$. Suppose that a is a unit in R. Show that

$$\begin{aligned} & ab = ac \\ \iff & b = c \\ \iff & ba = ca \end{aligned}$$

2.9 The Euclidean Algorithm for Integers

Theorem 2.9.1. *Let a and b be integers and suppose that $b|a$ and $a \neq 0$. Then*

$$1 \leq |b| \leq |a|.$$

Proof. Since $a|b$ we have $a = bk$ for some k in $\mathbb{Z}$. Since $a \neq 0$ we get $b \neq 0$ and $k \neq 0$. Hence $|b|$ and $|k|$ are positive integers and so $1 \leq |b|$ and $1 \leq |k|$ Hence also $|b| \cdot 1 \leq |b| \cdot |k|$ and so

$$1 \leq |b| = |b| \cdot 1 \leq |b| \cdot |k| = |bk| = |a|.$$

□

Definition 2.9.2. (a) *Let R be a ring and $a, b, c \in R$. We say that c is a common divisor of a and b in R provided that*

$$c|a \quad \text{and} \quad c|b.$$

(b) *Let a, b and d be integers. We say that d is a greatest common divisor of a and b in $\mathbb{Z}$, and we write*

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

provided that

(i) *d is a common divisor of a and b in $\mathbb{Z}$; and*

(ii) *if c is a common divisor of a and b in $\mathbb{Z}$, then $c \leq d$.*

Example 2.9.3. (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.

Theorem 2.9.4. *Let a, b, q, r and d be integers with*

$$a = bq + r \qquad \text{and} \qquad d = \gcd(b, r).$$

Then

$$d = \gcd(a, b).$$

Proof. We need to verify the two conditions (i) and (ii) of the gcd.

(i): Since $d = \gcd(b, r)$ we know that d is a common divisor of b and r. As $a = bq + r$ we conclude that d divides a, see 2.4.4(c). Thus d is a common divisor of a and b.

(ii) Let c be a common divisor of a and b. Since $a = bq + r$ we have $r = a - bq$. Hence $c \mid r$, see 2.4.4(c). Thus c is a common divisor of b and r. Since $d = \gcd(b, r)$ this gives $c \leq d$. □

Theorem 2.9.5 (Euclidean Algorithm)**.** *Let a and b be integers not both 0 and let E_{-1} and E_0 be the equations*

$$\begin{array}{ccccccc} E_{-1} & : & a & = & a \cdot 1 & + & b \cdot 0 \\ E_0 & : & b & = & a \cdot 0 & + & b \cdot 1 \end{array},$$

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

$$E_k \ : \ r_k \ = \ a \cdot x_k \ + \ b \cdot y_k \ .$$

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

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

(Note here that such t_{i+1}, q_{i+1} exist by the division algorithm 2.5.2)

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}, that is

$$r_{i+1} := r_{i-1} - r_i q_{i+1}, \qquad x_{i+1} := x_{i-1} - x_i q_{i+1}, \qquad y_{i+1} := y_{i-1} - x_i q_{i+1}.$$

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) *$d = \gcd(a, b)$.*

(c) *There exist $x, y \in \mathbb{Z}$ with $d = ax + by$.*

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

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

Hence

$$r_{i+1} := r_{i-1} - r_i q_{i+1}, \qquad x_{i+1} := x_{i-1} - x_i q_{i+1}, \qquad y_{i+1} := y_{i-1} - x_i q_{i+1}.$$

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}

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

So

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

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 2.9.4 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.

By equation E_{m-1} we have

$$r_{m-1} = ax_{m-1} + by_{m-1}.$$

If $r_{m-1} > 0$, then

$$d = r_{m-1} = ax_{m-1} + by_{m-1},$$

and if $r_{m-1} < 0$, then

$$d = -r_{m-1} = a(-x_{m-1}) + b(-y_{m-1}).$$

In either case (c) holds. □

Example 2.9.6. Let $a = 1492$ and $b = 1066$. Then

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

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

Theorem 2.9.7. *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 2.9.5 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 2.4.4(c), d divides e. Hence 2.9.1 shows that $d \leq |d| \leq |e| = e$. Thus d is the smallest possitive integer of the form $au + bv$ with $u, v \in \mathbb{Z}$. □

Theorem 2.9.8. *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 2.9.5 $d = ax + by$ for some $x, y \in \mathbb{Z}$. So if c is a common divisor of a and b, then 2.4.4(c) 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 2.9.1, $c \leq |c| \leq |d| = d$. Hence by definition, d is a greatest common divisor of a and b. □

Theorem 2.9.9. *Let a, b integers not both 0 with $1 = \gcd(a, b)$. Let c be an integer with $a \mid bc$. Then $a \mid c$.*

Proof. Since $1 = \gcd(a, b)$, 2.9.5 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 2.4.4(c), a also divides $a(xc) + (cb)y$. Thus $a \mid c$. □

Exercises 2.9:

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

#2. If $a|c$ and $b|c$, 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 $4|(a^2 - 1)$.

(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 $\gcd(a,b)|c$.

#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 $24|(a^2 - 1)$.

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

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

2.10 Integral Primes

Definition 2.10.1. *An integer p is called a prime if $p \notin \{0, 1, -1\}$ and the only divisors of p in $\mathbb{Z}$ are 1, -1, p and $-p$.*

Theorem 2.10.2. (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 $|p| = \gcd(a,p)$) or ($p \nmid a$ and $1 = \gcd(a,p)$).*

(c) *Let p and q be primes with $p|q$. 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 2.9.1

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

Moreover,

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

Thus the following statements are equivalent:

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

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

Case 1: Suppose $p|a$.

Since $p \mid p$, we conclude that p is a common divisor of a and p. Thus by 2.4.4(b) also $|p|$ is a common divisor of a and p. As $d = \gcd(a, p)$ this gives $|p| \leq d$. By definition of a prime we have $p \notin \{0, \pm 1\}$, so we $|p| > 1$. Hence also $d > 1$ and thus $d \neq 1$. Together with $(+)$ we get $d = |p|$. So $p \mid a$ and $|p| = \gcd(a, p)$. 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 $(+)$ $d = 1$. Thus $p \nmid a$ and $1 = \gcd(a, p)$. 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, we know that $p \notin \{\pm 1\}$ and so $p \in \{\pm q\}$. □

Theorem 2.10.3. *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 $p|ab$, then $p|a$ or $p|b$.*

Proof. '$\Longrightarrow$': Suppose p is prime and $p|ab$ for some integers a and b. Suppose that $p \nmid a$. Then 2.10.2 gives $1 = \gcd(a,p)$. Since $p|ab$, 2.9.9 now implies that $p|b$. So $p|a$ or $p|b$.

'$\Longleftarrow$': For the converse, see Exercise 2.10#3. □

Exercises 2.10:

#1. Let n be an integer with $n \notin \{0,1,-1\}$. Prove that there exists a positive prime integer p with $p|n$.

#2. Let p be an integer other than $0,\pm1$. 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 $r = \pm1$ or $s = \pm1$.

#3. Let p be an integer other than $0,\pm1$ 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.

#4. Prove that $1 = \gcd(a,b)$ if and only if there does not exist a prime integer p with $p|a$ and $p|b$.

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

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

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

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

#6. Let a and b be integers. Then $a|b$ 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.

2.11 Isomorphism and Homomorphism

Definition 2.11.1. *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*

$$f(a+b) = f(a) \oplus f(b) \qquad [f \text{ respects addition}]$$

and

$$f(a \cdot b) = f(a) \odot f(b) \qquad [f \text{ respects multiplication}]$$

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 bijective. surjective*

(c) $(R,+,\cdot)$ *is called* isomorphic *to* $(S,\oplus,\odot)$, *if there exists an isomorphism from* $(R,+,\cdot)$ *to* $(S,\oplus,\odot)$.

Example 2.11.2. (1) Consider

$$g: \quad \mathbb{R} \to \mathbb{R}, \quad a \mapsto -a.$$

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 a bijection.

(2) Let R and S be rings and consider

$$h: \quad R \to S, \quad r \mapsto 0_S.$$

Let $a, b \in R$. Then

$$h(a+b) = 0_S = 0_S + 0_S = h(a) + h(b) \qquad \text{and} \qquad h(ab) = 0_S = 0_S 0_S = h(a)h(b).$$

So h is a homomorphism. h is injective if and only if $R = \{0_R\}$ and h is surjective if and only if $S = \{0_S\}$. Hence h is an isomorphism if and only if $R = \{0_R\}$ and $S = \{0_S\}$.

(3) Let S be a ring and R a subring of S. Consider

$$\mathrm{id}_{R,S}: \quad R \to S, \quad r \mapsto r.$$

Let $a, b \in R$. Then

$$\mathrm{id}_{R,S}(a+b) = a + b = \mathrm{id}_{R,S}(a) + \mathrm{id}_{R,S}(b) \qquad \text{and} \qquad \mathrm{id}_{R,S}(ab) = ab = \mathrm{id}_{R,S}(a)\mathrm{id}_{R,S}(b)$$

and so $\mathrm{id}_{R,S}$ is a homomorphism. Note that $\mathrm{id}_{R,S}$ is injective. Moreover, $\mathrm{id}_{R,S}$ is surjective if and only if $R = S$. Hence $\mathrm{id}_R := \mathrm{id}_{R,R}$ is an isomorphism.

(4) Let R be a ring and $n \in R$. Consider the function

$$k: \quad R \to R_n, \quad a \mapsto [a]_n.$$

Let $a, b \in R$. By definition of the addition and multiplication in R_n

$$k(a+b) = [a+b]_n = [a]_n \oplus [b]_n = k(a) \oplus k(b) \qquad \text{and} \qquad k(ab) = [ab]_n = [a]_n \odot [b]_n = k(a) \odot k(b).$$

So k is homomorphism. Note that

$$k(n) = [n]_n = [0_R]_n = k(0_R).$$

If $n \neq 0_R$ we conclude that k is not injective.

Suppose $n = 0_R$. Then by Example 2.4.6(2) $a \equiv b \pmod{n}$ if and only of $a = b$. Thus k is injective.

Let $A \in R_n$. The definition of R_n shows that $A = [a]_n$ for some $a \in R$. Hence $k(a) = A$ and so k is surjective.

Example 2.11.3. Consider the function

$$f: \quad \mathbb{C} \to \mathrm{M}_2(\mathbb{R}), \quad r + si \mapsto \begin{bmatrix} r & s \\ -s & r \end{bmatrix}$$

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

$$\begin{aligned}
f\Big(ab\Big) &= f\Big((r+si)(\tilde{r}+\tilde{s}i)\Big)\\
&= f\Big((r\tilde{r}-s\tilde{s})+(r\tilde{s}+s\tilde{r})i\Big)\\
&= \begin{bmatrix} r\tilde{r}-s\tilde{s} & r\tilde{s}+s\tilde{r}\\ -(r\tilde{s}+s\tilde{r}) & r\tilde{r}-s\tilde{s}\end{bmatrix}\\
&= \begin{bmatrix} r & s\\ -s & r\end{bmatrix}\begin{bmatrix} \tilde{r} & \tilde{s}\\ -\tilde{s} & \tilde{r}\end{bmatrix}\\
&= f(r+si)f(\tilde{r}+\tilde{s}i)\\
&= f(a)f(b).
\end{aligned}$$

Thus 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 injective.

Since $1\neq 0$ we have that $\begin{bmatrix} 1 & 0\\ 0 & 0\end{bmatrix} \neq \begin{bmatrix} r & s\\ -s & r\end{bmatrix}$ for all $r,s\in\mathbb{R}$. Thus f is not surjective.

Put

$$S := \left\{ \begin{bmatrix} r & s\\ -s & r\end{bmatrix} \;\middle|\; r,s\in\mathbb{R} \right\}.$$

Using the Subring theorem it is straight forward to check that S is a subring of $\mathrm{M}_2(\mathbb{R})$. Alternatively, Theorem 2.11.11 below also shows that S is a subring of $\mathrm{M}_2(\mathbb{R})$. It follows that

$$\widetilde{f}: \quad \mathbb{C}\to S, \quad r+is\mapsto \begin{bmatrix} r & s\\ -s & r\end{bmatrix}.$$

is an isomorphism of rings. Thus

$$\mathbb{C} \qquad \text{and} \qquad \left\{ \begin{bmatrix} r & s\\ -s & r\end{bmatrix} \;\middle|\; r,s\in\mathbb{R} \right\}.$$

are isomorphic rings.

Notation 2.11.4. (a) *'$f : R \to S$ is a ring homomorphism' stands for the more precise statement '$(R, +, \cdot)$ 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 two conditions for a homomorphism become*

$$f(a+b) = f(a) + f(b) \quad \text{and} \quad f(ab) = f(a)f(b).$$

Remark 2.11.5. *Let $R = \{r_1, r_2, \dots, r_n\}$ be a ring with n elements. Suppose that the addition and multiplication table is given by*

A:

$+$	r_1	$\dots$	r_j	$\dots$	r_n
r_1	a_{11}	$\dots$	a_{1j}	$\dots$	a_{1n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r_i	a_{i1}	$\dots$	a_{ij}	$\dots$	a_{in}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r_n	a_{n1}	$\dots$	a_{nj}	$\dots$	a_{nn}

and M:

$\cdot$	r_1	$\dots$	r_j	$\dots$	r_n
r_1	b_{11}	$\dots$	b_{1j}	$\dots$	b_{1n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r_i	b_{i1}	$\dots$	b_{ij}	$\dots$	b_{in}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r_n	b_{n1}	$\dots$	b_{nj}	$\dots$	b_{nn}

So $r_i + r_j = a_{ij}$ and $r_i r_j = b_{ij}$ for all $1 \le i, j \le n$.

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':

	r'_1	$\dots$	r'_j	$\dots$	r'_n
r'_1	a'_{11}	$\dots$	a'_{1j}	$\dots$	a'_{1n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r'_i	a'_{i1}	$\dots$	a'_{ij}	$\dots$	a'_{in}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r'_n	a'_{n1}	$\dots$	a'_{nj}	$\dots$	a'_{nn}

and M':

	r'_1	$\dots$	r'_j	$\dots$	r'_n
r'_1	b'_{11}	$\dots$	b'_{1j}	$\dots$	b'_{1n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r'_i	b'_{i1}	$\dots$	b'_{ij}	$\dots$	b'_{in}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r'_n	b'_{n1}	$\dots$	b'_{nj}	$\dots$	b'_{nn}

(a) *f is a homomorphism if and only if A' and M' are the tables for the addition and multiplication of the elements $r'_1, \dots, r'_n$ in S, that is $r'_i + r'_j = a'_{ij}$ and $r'_i r'_j = b'_{ij}$ for all $1 \le i, j \le n$.*

(b) *f is injective if and only if $r'_1, \dots, r'_n$ are pairwise distinct.*

(c) *f is surjective if and only if $S = \{r'_1, r'_2, \dots, r'_n\}$.*

(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 $1 \le i, j \le n$. Since $r_i + r_j = a_{ij}$ and $r_i r_j = b_{ij}$ this holds if and only if

$$f(a_{ij}) = f(r_i) + f(r_j) \quad \text{and} \quad f(b_{ij}) = f(r_i)f(r_j)$$

for all $1 \le i, j \le n$. Since $f(r) = r'$, this is equivalent to

$$a'_{ij} = r'_i + r'_j \quad \text{and} \quad b'_{ij} = r'_i r'_j$$

for all $1 \le i, j \le n$

(b) f is injective if and only if (for or all $a, b \in R$) $f(a) = f(b)$ implies $a = b$ and so if and only if $a \ne b$ implies $f(a) \ne f(b)$. Since for each $a \in R$ there exists a unique $1 \le i \le n$ with $a = r_i$, f is injective if and only (for all $1 \le i, j \le n$) $i \ne j$ implies $f(r_i) \ne f(r_j)$, that is $i \ne j$ implies $r'_i \ne r'_j$.

(c) f is surjective 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 surjective if and only if $S = \{r'_1, \ldots, r'_n\}$.

(d) Follows from (a)-(c). □

Example 2.11.6. Let R be the ring with additions and multiplication table

⊞	0	1
0	1	0
1	0	1

and

⊡	0	1
0	0	1
1	1	1

Note here that R is indeed a ring, see Example 2.1.6. Then the function

$$f: \quad R \to \mathbb{Z}_2, \quad 0 \mapsto 1, \quad 1 \to 0$$

is an isomorphism.

Replacing 0 by 1 and 1 by 0 in the above tables we obtain

	1	0
1	0	1
0	1	0

and

	1	0
1	1	0
0	0	0

.

Note that these are addition and multiplication tables for $\mathbb{Z}_2$ and so by 2.11.5 f is an isomorphism.

Theorem 2.11.7. *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* $a, b \in R$.

Proof. (a) We have

$$\begin{aligned} f(0_R) + f(0_R) &= f(0_R + 0_R) && - f \text{ respects addition} \\ &= f(0_R) && - \textbf{Ax 4} \text{ for } R. \end{aligned}$$

So the Additive Identity Law 2.2.4 for S implies that $f(0_R) = 0_S$.

(b) We compute

$$\begin{aligned} f(a) + f(-a) &= f(a + (-a))) && - f \text{ respects addition} \\ & f(0_R) && - \textbf{Ax 5} \text{ for } R. \\ &= 0_S && - \text{by (a)} \end{aligned}$$

So the Additive Inverse Law 2.2.6 for S implies that $f(-a) = -f(a)$.

(c)

$$f(a-b) \overset{\text{Def} -}{=} f(a + (-b)) \overset{\text{f hom}}{=} f(a) + f(-b) \overset{\text{(b)}}{=} f(a) + (-f(b)) \overset{\text{def} -}{=} f(a) - f(b).$$

□

Theorem 2.11.8. *Let $f : R \to S$ be a homomorphism of rings. Suppose that R has an identity and that f is surjective. Then*

(a) *S is a ring with identity and* $f(1_R) = 1_S$.

(b) *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 will first show that $f(1_R)$ is an identity in S. For this let $s \in S$. Since f is surjective, $s = f(r)$ for some $r \in R$. Thus

$$s \cdot f(1_R) = f(r)f(1_R) \overset{\text{f hom}}{=} f(r1_R) \overset{\textbf{(Ax 10)}}{=} f(r) = s,$$

and similarly $f(1_R) \cdot s = s$. So $f(1_R)$ is an identity in S. By 2.8.3(a) a ring has at most one identity and so $f(1_R) = 1_S$.

(b) Let u be a unit in R. We will first show that $f(u^{-1})$ is an inverse of $f(u)$:

$$f(u)f(u^{-1}) \overset{\text{f hom}}{=} f(uu^{-1}) \overset{\text{def inv}}{=} f(1_R) \overset{\text{(a)}}{=} 1_S.$$

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 2.8.4 $f(u)^{-1}$ is the unique inverse of $f(u)$ and so $f(u^{-1}) = f(u)^{-1}$. □

Example 2.11.9. Find all surjective homomorphisms from $\mathbb{Z}_6$ to $\mathbb{Z}_2 \times \mathbb{Z}_3$.

We start with setting up some convenient notation. For $a, b \in \mathbb{Z}$ and h a function from $\mathbb{Z}_6$ to $\mathbb{Z}_2 \times \mathbb{Z}_3$ define

$$[a] := [a]_6, \qquad h[a] := h\big([a]_6\big), \quad \text{and} \quad [a,b] := \big([a]_2, [b]_3\big).$$

Let $a, b, c, d \in \mathbb{Z}$. Then

$$[a,b] + [c,d] = \big([a]_2, [b]_3\big) + \big([c]_2, [d]_3\big) = \big([a]_2 + [c]_2, [b]_3 + [d]_3\big) = \big([a+c]_2, [b+d]_3\big) = [a+c, b+d].$$

Thus

$$(*) \qquad [a,b] + [c,d] = [a+c, b+d] \quad \text{and similarly} \quad [a,b] \cdot [c,d] = [a \cdot c, b \cdot d]$$

'Uniqueness of the surjective homomorphism':

Let $f : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ be a surjective homomorphism. We will compute $f[r]$ for $0 \le r \le 5$ and thereby prove that f is unique determined. Since f is an surjective homomorphism, we get from 2.11.8(a) 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 2.11.7(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 2.5.6 $\mathbb{Z}_6 = \{[0],[1],[2],[3],[4],[5]\}$. Hence f is uniquely determined.

'Existence of the surjective homomorphism':

Define a function $g : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ by

$$(**) \qquad g[r] = [r,r] \quad \text{for all } 0 \le r \le 5.$$

We will show that g is a isomorphism, and so also surjective homomorphism. For this we first show that $g[m] = [m,m]$ for all $m \in \mathbb{Z}$. Indeed, by the Division Algorithm, $m = 6q + r$ for some $q, r \in \mathbb{Z}$ with $0 \le r < 6$. Then by 2.4.9 $[m]_6 = [r]_6$ 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

$$(***) \qquad g[m] = g[r] = [r,r] = [m,m].$$

Thus

$$g[n+m] \overset{(***)}{=} [n+m, n+m] \overset{(*)}{=} [n,n] + [m,m] \overset{(***)}{=} g[n] + g[m],$$

and

$$g[nm] \overset{(***)}{=} [nm, nm] \overset{(*)}{=} [n,n][m,m] \overset{(***)}{=} g[n]g[m].$$

So g is a homomorphism of rings. Since $\mathbb{Z}_2 = \{[0]_2, [1]_2\}$ and $\mathbb{Z}_3 = \{[0]_3, [1]_3, [2]_3\}$ we have

$$\begin{aligned}\mathbb{Z}_2 \times \mathbb{Z}_3 = \{(x,y) \mid x \in \mathbb{Z}_2, y \in \mathbb{Z}_3\} &= \{[0,0],[0,1],[0,2],[1,0],[1,1],[1,2]\} \\ &= \{[0,0],[4,4],[2,2],[3,3],[1,1],[5,5]\} \\ &= \{g[0],\ g[4],\ g[2],\ g[3],\ g[1],\ g[5]\}\end{aligned}$$

and so g is a surjective. Note that g is also injective. Thus g is an isomorphism and so

$$\mathbb{Z}_6 \quad \text{is isomorphic to} \quad \mathbb{Z}_2 \times \mathbb{Z}_3.$$

Example 2.11.10. 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 surjective, there exists $r \in R$ with $f(r) = s$. Thus

$$s + s = f(r) + f(r) \overset{\text{f hom}}{=} f(r+r) \overset{(*)}{=} f(0_R) \overset{2.11.7(a)}{=} 0_S$$

Since $[1]_4 + [1]_4 = [2]_4 \neq [0]_4$ we conclude that $\mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Theorem 2.11.11. *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 2.7.2. 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 2.11.7(a) $f(0_R) = 0_S$. By **Ax 4** $0_R \in R$ 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) $-x \overset{(**)}{=} -f(a) \overset{2.11.7(b)}{=} f(-a)$. By **Ax 5** $-a \in R$. So $-x \in \operatorname{Im} f$ by (*). □

Definition 2.11.12. 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 2.11:

#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 $M_2(\mathbb{Z}_2)$. Show that S is isomorphic to the ring R from Exercise 2.1.#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 $f : \mathbb{Z}_m \to \mathbb{Z}_{nm}, [x]_m \mapsto [xn]_{nm}$. 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 injective.

(d) If $n > 1$, then f is not surjective.

#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.

#9. Show that the two rings are not isomorphic.

(a) $2\mathbb{Z}$ and $\mathbb{Z}$.

(b) $\mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ and $\mathrm{M}_2(\mathbb{R})$.

(c) $\mathbb{Z}_2 \times \mathbb{Z}_{14}$ and $\mathbb{Z}_{16}$.

(d) $\mathbb{Q}$ and $\mathbb{R}$.

(e) $\mathbb{Z} \times \mathbb{Z}_2$ and $\mathbb{Z}$.

(f) $\mathbb{Z}_4 \times \mathbb{Z}_4$ and $\mathbb{Z}_{16}$.

#10. Let $f : R \to S$ and $g : S \to T$ be homomorphism of rings.

(a) Show that $g \circ f : R \to T$ is a homomorphism of rings.

(b) If f and g are isomorphisms, show that $g \circ f$ is an isomorphism.

(c) Suppose f is an isomorphism. For $s \in S$ let s' be the unique element of R with $f(s') = s$. Show that the function $h : S \to R, s \mapsto s'$ is an isomorphism of rings.

#11. Let $f : R \to S$ be an isomorphism of rings. If R is an integral domain, show that S is an integral domain.

2.12 Associates in commutative rings

Definition 2.12.1. *Let R be 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 $a \sim_R b$ if there exists a unit u in R with $au = b$. We will usually just write $a \sim b$ for the more precise $a \sim_R b$.*

Remark 2.12.2. *Until now we have used '$\sim$' to denote an arbitrary relation. From now on the symbol '$\sim$' will be reserved for the 'associated' on a ring R.*

Theorem 2.12.3. *Let a, b be integers. Then $1 = \gcd(a, b)$ if and only if there exist integers u and v with $1 = au + bv$.*

Proof. If $1 = \gcd(a, b)$, then $1 = au + bv$ for some $u, v \in \mathbb{Z}$ by the Euclidean Algorithm 2.9.5.

Conversely, suppose that $1 = au + bv$ for some $u, v \in \mathbb{Z}$. Since 1 is the smallest positive integer this shows that 1 is the smallest positive integer of the form $au + bv, u, v \in \mathbb{Z}$. From 2.9.7 we conclude that $1 = \gcd(a, b)$. □

Theorem 2.12.4. *Let n be a non-zero integer and $a \in \mathbb{Z}$. Then $1 = \gcd(a, n)$ if and only if $[a]_n$ is a unit in $\mathbb{Z}_n$.*

Proof.

$$\begin{array}{llll} & 1 = \gcd(a,b) & & \\ \iff & 1 = au + nv & \text{for some } u, v \in \mathbb{Z} & - 2.12.3 \\ \iff & [1]_n = [au]_n & \text{for some } u \in \mathbb{Z} & - 2.4.9 \\ \iff & [1]_n = [a]_n[u]_n & \text{for some } u \in \mathbb{Z} & - \text{Definition of multiplication in } \mathbb{Z}_n \\ \iff & [1]_n = [a]_n U & \text{for some } U \in \mathbb{Z}_n & - \text{Definition of } \mathbb{Z}_n \\ \iff & 1_{\mathbb{Z}_n} = [a]_n U & \text{for some } U \in \mathbb{Z}_n & - 1_{\mathbb{Z}_n} = [1]_n \text{ by 2.6.4} \\ \iff & 1_{\mathbb{Z}_n} = [a]_n U \text{ and } 1_{\mathbb{Z}_n} = U[a]_n & \text{for some } U \in \mathbb{Z}_n & - \mathbb{Z}_n \text{ is commutative} \\ \iff & [a]_n \text{ is a unit in } \mathbb{Z}_n & & - \text{Definition of a unit} \end{array}$$

□

Example 2.12.5. (1) Let $n \in \mathbb{Z}$. Find all associates of n in $\mathbb{Z}$.

By 2.8.2 the units in $\mathbb{Z}$ are ± 1. So the associates of n are $n \cdot \pm 1$, that is $\pm n$.

(2) Find all associates of 0, 1, 2 and 5 in $\mathbb{Z}_{10}$.

By 2.5.6 $\mathbb{Z}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ 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	$\boxed{1}$	2	$\boxed{1}$	2	5

and so by 2.12.4 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	$\pm 2, \pm 3 \cdot 2$	$\pm 2, \pm 4$
± 3	$\pm 3, \pm 3 \cdot 4$	$\pm 1, \pm 3$
± 4	$\pm 4, \pm 3 \cdot 4$	$\pm 2, \pm 4$
5	$\pm 5, \pm 3 \cdot 5$	5

Theorem 2.12.6. *Let R be a 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)** $a1_R = a$ and so $a \sim a$ by definition of '$\sim$'. Thus $\sim$ is reflexive.

Symmetric: Let $a, b \in R$ with $a \sim b$. By by definition of '$\sim$' this means that 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})

$$bu^{-1} = (au)u^{-1} \overset{\textbf{Ax 2}}{=} a(uu^{-1}) \overset{\text{def } u^{-1}}{=} a1_R \overset{\textbf{(Ax 10)}}{=} a.$$

By 2.8.5 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 2.8.5 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. □

Example 2.12.7. Determine the equivalence classes of $\mathbb{Z}_{10}$ with respect to $\sim$.

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 2.12.5

$$\begin{aligned} [0]_\sim &= \{0\} \\ [1]_\sim &= \{\pm 1, \pm 3\} \\ [2]_\sim &= \{\pm 2, \pm 4\} \\ [5]_\sim &= \{5\} \end{aligned}$$

By 2.5.6 $\mathbb{Z}_{10} = \{0, 1, \ldots, 9\} = \{0, \pm 1, \pm 2, \pm 3, \pm 4, 5\}$. So for each $x \in \mathbb{Z}_{10}$ there exists $y \in \{0, 1, 2, 5\}$ with $x \in [y]_\sim$. Thus by 1.5.5 $[x]_\sim = [y]_\sim$. So $[0]_\sim, [1]_\sim, [2]_\sim, [5]_\sim$ are all the equivalence classes of $\sim$.

Theorem 2.12.8. *Let R be a ring with identity and $a, b \in R$ with $a \sim b$. Then $a|b$ and $b|a$ in R.*

Proof. Since $a \sim b$, $au = b$ for some unit $u \in R$. So $a|b$.

By 2.12.6 the relation $\sim$ is symmetric and so $a \sim b$ implies $b \sim a$. Hence we can apply the result of the previous paragraph applied with a and b interchanged and conclude that $b|a$. □

Theorem 2.12.9. *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): If $1_R \sim r$ then 2.12.8 gives $r|1_R$.

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

(c) $\Longrightarrow$ (d): Suppose that $rs = 1_R$ for some $s \in R$. Since R is commutative we get $sr = 1_R$. So r is a unit.

(d) $\Longrightarrow$ (a): Suppose r is a unit. By **(Ax 10)**, $1_R r = r$. Since r is a unit this gives $1_R \sim r$ by definition of '$\sim$'. □

Theorem 2.12.10. *Let R be a ring with identity and $a, b, c, d \in R$.*

(a) *Suppose $a \sim b$. Then $a|c$ if and only if $b|c$.*

(b) *Suppose $c \sim d$. Then $b|c$ if and only if $b|d$.*

(c) *Suppose $a \sim b$ and $c \sim d$. Then $a|c$ if and only if $b|d$.*

Proof. (a) Suppose that $a \sim b$.

$\Longleftarrow$: Suppose that $b|c$. Since $a \sim b$ we know that $a|b$, see 2.12.8. From $a|b$ and $b|c$ we get $a|c$, since $|$ is transitive by 2.4.4(a)).

$\Longrightarrow$: Since $a \sim b$ and '$\sim$' is symmetric (see 2.12.6) we have $b \sim a$. So we can apply the result of previous paragraph applied with a and b interchanged. Thus $a|c$ implies $b|c$.

(b) Suppose that $c \sim d$.

$\Longrightarrow$: Suppose that $b|c$. Since $c \sim d$ know that $c|d$ and so $b|d$ as $|$ is transitive.

$\Longleftarrow$: Since $c \sim d$ and '$\sim$' is symmetric we have $d \sim c$. So we can apply the result of previous paragraph applied with c and d interchanged. Thus $b|d$ implies $b|c$.

(c) Suppose that $a \sim b$ and $c \sim d$. By (a) $a|c$ if and only if $b|c$. By (b) the latter holds if and only if $b|d$. □

Definition 2.12.11. *Let R be a 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 2.12:

#1. Let $R = \mathbb{Z}_{18}$.

(a) Find all units in R.

(b) Determine the equivalence classes of the relation $\sim$ on R.

#2. Let R be a 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 $b|c$.

#5. Is A associated to B in $\mathrm{M}_2(\mathbb{R})$?

(a) $A = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 10 & 6 \\ 15 & 9 \end{bmatrix}$.

(b) $A = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 12 & 20 \\ 15 & 25 \end{bmatrix}$.

Chapter 3

Polynomial Rings

3.1 Addition and Multiplication

Definition 3.1.1. *Let R and P be rings and $x \in P$. Then P is called a* polynomial ring *in x with coefficients in R provided that the following four conditions hold:*

(i) *R has identity 1_R, P has an identity 1_P and 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*

$$f = \sum_{i=0}^{n} f_i x^i \ \ (= f_0 + f_1 x + \ldots + f_n x^n).$$

(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 $n < i \leq m$.

Remark 3.1.2. *Let P be a polynomial ring in x with coefficients in the ring R.*

(a) *The elements of P are called polynomials in x with coefficients in R.* **Polynomials are not functions***. See section 3.4 for the connections between polynomials and polynomial functions.*

(b) *x is a fixed element of P. x* **is not a variable***.*

Theorem 3.1.3. *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 $ab^n = b^n a$ for all $n \in \mathbb{N}$*

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) Suppose that

$$(*) \qquad ab = ba.$$

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

$$(**) \qquad ab^k = b^k a.$$

We compute

$$\begin{aligned} ab^{k+1} &= a(b^k b) && -\text{ definition of } b^{k+1} \\ &= (ab^k)b && -\textbf{ Ax 7} \\ &= (b^k a)b && -\ (**) \\ &= b^k(ab) && -\textbf{ Ax 7} \\ &= b^k(ba) && -\ (*) \\ &= (b^k b)a && -\textbf{ Ax 7} \\ &= b^{k+1}a. && -\text{ definition of } b^{k+1} \end{aligned}$$

Thus (b) also holds for $n = k+1$. So by the Principal Of Induction, (b) holds for all $n \in \mathbb{N}$. □

Theorem 3.1.4. *Let R be a ring with identity and P a polynomial ring with coefficients in R with respect to x. Then $1_R = 1_P$. In particular, $x = 1_R x$.*

Proof. Let $f \in P$. Then by Condition 3.1.1(iii) on polynomial ring there exist $n \in \mathbb{N}$ and $f_0, f_1, \ldots f_n \in R$ with

$$(*) \qquad f = \sum_{i=0}^{n} f_i x^i.$$

Let $1 \le i \le n$. By Condition 3.1.1(ii) on polynomial ring $1_R x = x 1_R$ and so by 3.1.3(b)

$$(**) \qquad 1_R x^i = x^i 1_R.$$

Thus

$$(***) \qquad (f_i x^i)1_R \overset{\textbf{Ax 7}}{=} f_i(x^i 1_R) \overset{(**)}{=} f_i(1_R x^i) \overset{\textbf{Ax 2}}{=} (f_i 1_R)x^i \overset{\textbf{(Ax 10)}}{=} f_i x^i$$

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 and so $1_R = 1_P$. Since $x \in P$ this gives $1_R x = 1_P x = x$. □

Theorem 3.1.5. *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) $f + g = \displaystyle\sum_{i=0}^{\max(n,m)} (f_i + g_i) x^i$.

(b) $-f = \sum_{i=0}^{n} (-f_i) x^n$.

(c) $fg = \displaystyle\sum_{i=0}^{n} \left(\sum_{j=0}^{m} f_i g_j x^{i+j}\right) = \sum_{k=0}^{n+m} \left(\sum_{i=\max(0,k-m)}^{\min(n,k)} f_i g_{k-i}\right) x^k = \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} f_i g_{k-i}\right) x^k$.

Proof. (b) (a) Put $p := \max(n, m)$. Then $f_i = 0_R$ for all $n < i \le p$ and $g_i = 0_R$ for all $m < i \le p$. 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

$$\begin{aligned} f + g &= \left(\textstyle\sum_{i=0}^{p} f_i x^i\right) + \left(\textstyle\sum_{i=0}^{p} g_i x^i\right) && -(*) \\ &= \textstyle\sum_{i=0}^{p} (f_i x^i + g_i x^i) && -\text{GCL and GAL} \\ &= \textstyle\sum_{i=0}^{p} (f_i + g_i) x^i && -\ \textbf{Ax 8} \end{aligned}$$

So (a) holds.

(b) Since R is a subring of P we have $0_R = 0_P$. Using (a) we compute

$$f + \sum_{i=0}^{n} (-f_i) x^i = \sum_{i=0}^{n} (f_i + (-f_i)) x^i = \sum_{i=0}^{n} 0_P x^i = \sum_{i=0}^{n} 0_P = 0_P.$$

and so $-f = \sum_{i=0}^{n} (-f_i) x^i$ by the Additive Inverse Law.

(c) Let $a \in R$ and $b \in \mathbb{N}$. By Hypothesis $ax = xa$ and so by 3.1.3(b)

$$(**) \qquad ax^n = x^n a.$$

We now can compute fg.

(∗∗∗)

$$\begin{aligned}
fg &= \left(\sum_{i=0}^{n} f_i x^i\right)\cdot\left(\sum_{j=0}^{m} g_j x^j\right) && \text{– (ii) and (iii)}\\
&= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i x^i)(g_j x^j)\right) && \text{– GDL}\\
&= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i(x^i g_j))x^j\right) && \text{– GAL}\\
&= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i(g_j x^i))x^j\right) && \text{– } x^i g_j = g_j x^i \text{ by } (**)\\
&= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i g_j)(x^i x^j)\right) && \text{– GAL}\\
&= \sum_{i=0}^{n}\left(\sum_{j=0}^{m}(f_i g_j)x^{i+j}\right) && \text{– } x^i x^j = x^{i+j}\text{, by 3.1.3(a)}
\end{aligned}$$

Let $i, j, k \in \mathbb{Z}$ with $k = i + j$. We will show that

(+) $$0 \le i \le n \text{ and } 0 \le j \le m \iff 0 \le k \le n+m \text{ and } \max(0, k-m) \le i \le \min(k, n)$$

Suppose first that $0 \le i \le n$ and $0 \le j \le m$. Then $0 \le k = i + j \le n + m$. Since $j \le m$ we have $m - j > 0$ and so $k - m = i + j - m = i - (m - j) \le i$. As $0 \le i$ this gives $\max(0, k - m) \le i$. Since $j \ge 0$ we have $i \le i + j = k$. As $i \le n$ we get $i \le \min(k, n)$.

Suppose next that $0 \le k \le n + m$ and $\max(0, k - m) \le i \le \min(k, n)$. Then $0 \le i \le n$. Since $i \le k$ we get $0 \le k - i = j$. As $k \le n + m$ and $i \le n$ we have $j = k - i \le (n + m) - n \le m$. Thus $(***)$ holds.

Put

$$A := \{0, 1, \dots, n\} \times \{0, \dots, m\}, \quad B := \{(k, i) \in \mathbb{Z} \times \mathbb{Z} \mid 0 \le k \le n + m,\ \max(0, k - m) \le i \le \min(k, n)\}.$$

It follows that the function

$$A \to B, \quad (i, j) \mapsto (i + j, i)$$

is a bijection with inverse

$$B \to A, \quad (k, i) \mapsto (i, k - i).$$

Hence the substitution $k = i + j$ (and so $j = k - i$) and the GCL and GAL imply that

(++)

$$\begin{aligned}
\sum_{i=0}^{n}\left(\sum_{j=0}^{m} f_i g_j x^{i+j}\right) &= \sum_{k=0}^{n+m}\left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i} x^k\right)\\
&= \sum_{k=0}^{n+m}\left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i}\right) x^k \quad \text{–GDL}
\end{aligned}.$$

Suppose $0 \leq i < \max(0, k-m)$. Then then $k - i > m$ and so $g_{k-i} = 0_R$. Hence $f_i g_{k-i} = f_i 0_R = 0_R$ (by 2.2.9(c)).

Suppose $\min(k,n) < i \leq k$. Then $\min(n,k) \neq k$ and so $\min(n,k) = n$. Hence $n < i$, so $f_i = 0_R$. Thus $f_i g_{k-i} = 0_R g_{k-i} = 0_R$. 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

$$(+++) \qquad \sum_{k=0}^{n+m} \left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i} \right) x^k = \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} f_i g_{k-i} \right) x^k.$$

Combining $(***)$, $(++)$ and $(+++)$ gives (c). □

Example 3.1.6. Let P be a polynomial ring in x with coefficients in $\mathbb{Z}_6$. Let

$$f = 1 + 2x + 3x^2 \qquad \text{and} \qquad g = 1 + 4x + 5x^2 + 2x^3$$

Compute $f + g$ and $f \cdot g$ in P.

$$\begin{aligned} f + g &= 1 + 2x + 3x^2 \\ &\quad + 1 + 4x + 5x^2 + 2x^3 \\ &= 2 + 6x + 8x^2 + 2x^3 \\ &= 2 \qquad + 2x^2 + 2x^3 \end{aligned}$$

$$\begin{aligned} fg &= (1 + 2x + 3x^2)(1 + 4x + 5x^2 + 2x^3) \\ &= (1 \cdot 1) + (1 \cdot 4 + 2 \cdot 1)x + (1 \cdot 5 + 2 \cdot 4 + 3 \cdot 1)x^2 \\ &\qquad + (1 \cdot 2 + 2 \cdot 5 + 3 \cdot 4)x^3 + (2 \cdot 2 + 3 \cdot 5)x^4 + (3 \cdot 2)x^5 \\ &= 1 + 6x + 16x^2 + 24x^3 + 19x^4 + 6x^5 \\ &= 1 \qquad + \ 4x^2 \qquad + \ 3x^4 \end{aligned}$$

Definition 3.1.7. *Let R be a ring with identity.*

(a) *$R[x]$ denotes the polynomial ring in x with coefficients in R constructed in F.3.1. So the elements of $R[x]$ are the infinite sequence*

$$(a_i)_{i=0}^{\infty} = (a_0, a_1, a_2, \ldots, a_i, \ldots)$$

such that $a_i \in R$ for all $i \in \mathbb{N}$ and there exists $n \in \mathbb{N}$ with $a_i = 0_R$ for all $i > n$. Also

$$x = (0_R, 1_R, 0_R, 0_R, \ldots, 0_R, \ldots)$$
$$(a_0, a_1, a_2 \ldots, a_i, \ldots) + (b_0, b_1, b_2 \ldots, b_i, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots, a_i + b_i, \ldots)$$

and

$$(a_0, a_1, a_2 \ldots, a_i, \ldots) \cdot (b_0, b_1, b_2 \ldots, b_i, \ldots)$$
$$= (a_0 b_0, a_0 b_1 + a_1 b_0, a_0 b_2, a_1 b_1, a_2 b_0, \ldots, a_0 b_i + a_1 b_{i-1} + a_{i-1} b_1 + a_i b_0, \ldots)$$

(b) *Let* $f \in R[x]$ *and let* $n \in \mathbb{N}$ *and* $a_0, a_1, \ldots a_n \in R$ *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 3.1.1)*

(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) *Let* $f \in R[x]$. *If* $f = 0_R$ *define* $\deg f := 0_R$ *and* $\text{lead}(f) = 0_R$. *If* $f = \sum_{i=0}^{n} f_i x^i$ *with* $f_i \in R$ *and* $f_n \neq 0$, *define* $\deg f := n$ *and* $\text{lead}(f) = f_n$.

Theorem 3.1.8. *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* $\text{lead}(f) = 0_R$.

(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* $f = \text{lead}(f)$.

(d) $f = \sum_{i=0}^{\deg f} f_i x^i$. *Here, for* $f = 0_R$, *the empty sum* $\sum_{i=0}^{-\infty} f_i x^i$ *is defined to by* 0_R.

Proof. This follows straightforward from the definition of $\deg f$ and lead$f$ and we leave the details to the reader. □

Theorem 3.1.9. *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* $\text{lead}(fg) = \text{lead}(f)\text{lead}(g)$.

  (2) $\deg(fg) < \deg f + \deg g$, $\text{lead}(f)\text{lead}(g) = 0_R$, $f \neq 0_R$ *and* $g \neq 0_R$.

In particular, $\deg fg \leq \deg f + \deg g$.

Proof. Put $n := \deg f$ and $m := \deg g$. By 3.1.8(d) we have

$$f = \sum_{i=0}^{n} f_i x^n \qquad \text{and} \qquad g = \sum_{i=0}^{m} g_i x^i.$$

(a) By 3.1.5(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) If $f = 0_R$, then also $-f = 0_R$ and so $\deg f = -\infty = \deg(-f)$. Suppose $f \neq 0_R$. Then $f_n \neq 0_R$ and so also $-f_n \neq 0_R$. Also $-f = -(\sum_{i=0}^{n} f_i x^i) = \sum_{i=0}^{n}(-f_i)x^i$. Since $-f_n \neq 0_R$ this gives $\deg(-f) = n = \deg f$.

(c) Suppose first that $f = 0_R$. Then $fg = 0_R g = 0_R$. Hence $\deg f = -\infty$, $\deg(fg) = -\infty$, $\text{lead} f = 0_R$ and $\text{lead}(fg) = 0_R$. Hence

$$\deg(fg) = -\infty = -\infty + \deg g = \deg f + \deg g \text{ and } \text{lead}(fg) = 0_R = 0_R \cdot \text{lead}(g) = \text{lead}(f)\text{lead}(g).$$

So (c:1) holds in this case. Similarly, (c:1) holds if $g = 0_R$.

So suppose $f \neq 0_R \neq g$ By 3.1.5(c),

$$fg = \sum_{k=0}^{n+m} \left(\sum_{i=\max(0,k-m)}^{\min(k,n)} f_i g_{k-i} \right) x^k.$$

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

$$(fg)_{n+m} = \sum_{i=n}^{n} f_i g_{n+m-i} = f_n g_m = \text{lead}(f)\text{lead}(g).$$

Suppose that $\text{lead}(f)\text{lead}(g) \neq 0_R$. Then $\deg(f + g) = n + m$ and $\text{lead}(fg) = \text{lead}(f)\text{lead}(g)$. Thus (c:1) holds.

Suppose that $\text{lead}(f)\text{lead}(g) = 0_R$. Then $\deg(f + g) < n + m$ and (c:2) holds. □

Theorem 3.1.10. *Let R be a commutative ring with identity. Then $R[x]$ is commutative ring with identity.*

Proof. By definition of a polynomial ring $R[x]$ is a ring with identity. So we just need to show that $R[x]$ is commutative. Let $f, g \in R[x]$ and put $n - \deg f$ and $m = \deg g$. Then

$$\begin{aligned}
fg &= \left(\sum_{i=0}^{n} f_i x^i\right)\left(\sum_{j=0}^{m} g_j x^j\right) \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m} f_i g_j x^{i+j} && \text{– Theorem 3.1.5} \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m} g_j f_i x^{j+i} && \text{– R commutative} \\
&= \sum_{j=0}^{m}\sum_{i=0}^{n} g_j f_i x^{j+i} && \text{– GCL, GAL} \\
&= \left(\sum_{j=0}^{m} g_j x^j\right)\left(\sum_{i=0}^{n} f_i x^i\right) && \text{– Theorem 3.1.5} \\
&= gf
\end{aligned}$$

We proved that $fg = gf$ for all $f, g \in R[x]$ and so $R[x]$ is commutative. □

Theorem 3.1.11. *Let R be field or an integral domain. Then*

(a) $\deg(fg) = \deg f + \deg g$ *and* $\operatorname{lead}(fg) = \operatorname{lead}(f)\operatorname{lead}(g)$ *for all* $f, g \in R[x]$.

(b) $\deg(rf) = \deg f$ *and* $\operatorname{lead}(rf) = r\operatorname{lead}(f)$ *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 2.8.10 any field is an integral domain. So in any case R is an integral domain. Let $f, g \in R[x]$. We will first show that

(*) If $\operatorname{lead}(f)\operatorname{lead}(g) = 0_R$ then $f = 0_R$ or $g = 0_R$.

Indeed since R is an integral domain, $\operatorname{lead}(f)\operatorname{lead}(g) = 0_R$ implies $\operatorname{lead}(f) = 0$ or $\operatorname{lead}(g) = 0_R$. 3.1.8 now shows $f = 0_R$ or $g = 0_R$.

(a) By 3.1.9(c)

(1) $\deg(fg) = \deg f + \deg g$ and $\operatorname{lead}(fg) = \operatorname{lead}(f)\operatorname{lead}(g)$, or

(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 the first case, (a) holds. The second case contradicts (*) and so does not occur.

(b) Let $r \in R$ with $r \neq 0_R$. By 3.1.8 $\deg r = 0$ and $\operatorname{lead} r = r$. Using (b) we conclude that

$$\deg(rf) = \deg r + \deg f = 0 + \deg f = \deg f \quad \text{and} \quad \operatorname{lead}(rf) = \operatorname{lead}(r)\operatorname{lead}(f) = r\operatorname{lead}(f).$$

(c) By 3.1.10, $R[x]$ is a commutative ring with identity. Since R is an integral domain $1_R \neq 0_R$ and thus $1_{R[x]} = 1_R \neq 0_R = 0_{R[x]}$. Let $fg \in R[x]$ with $fg = 0_R$. Then by (a) $\operatorname{lead}(f)\operatorname{lead}(g) = \operatorname{lead}(fg) = \operatorname{lead}(0_R) = 0_R$ and by (*), $f = 0_R$ or $g = 0_R$. Hence $R[x]$ is an integral domain. □

Theorem 3.1.12 (Division Algorithm)**.** *Let R be ring with identity and $f, g \in R[x]$ such that $g \neq 0_R$ and $\operatorname{lead}(g)$ is unit in R. Then there exist uniquely determined $q, r \in R[x]$ with*

$$f = gq + r \quad \textit{and} \quad \deg r < \deg g.$$

Proof. Fix $g \in R[x]$ such that $g \neq 0_R$ and $\operatorname{lead}(g)$ is unit in R. 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 $q, r \in R[x]$ with $f = gq + r$ and $\deg r < \deg g$.

We will use complete induction to show that $P(n)$ holds for all $n \in \mathbb{N}$. So 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 $f \in R[x]$ with $\deg f \leq k$. Note that $f = g \cdot 0_R + f$. If $\deg f < \deg g$ then then $P(k)$ holds for f with $q := 0_R$ and $r := f$.

So we may assume that $\deg f \geq \deg g$. Put $m := \deg g$, then $m \geq \deg f \geq k$. Since $g \neq 0_R$ we have $m = \deg g \in \mathbb{N}$, $g_m \neq 0_R$ and $g_m = \operatorname{lead}(f)$. By hypothesis $\operatorname{lead}(g)$ is a unit in R and so g_m as an inverse g_m^{-1}. Define

$$(*) \qquad \tilde{f} := f - g \cdot g_m^{-1} f_k x^{k-m}.$$

We compute

$$\begin{array}{rrcrccc}
g: & g_m x^m & + & g_{m-1}x^{m-1} & + & \ldots & + \\
f: & f_k x^k & + & f_{k-1}x^{k-1} & + & \ldots & + \\
g \cdot g_m^{-1} f_k x^{k-m}: & g_m g_m^{-1} f_k x^k & + & g_{m-1} g_m^{-1} f_k x^{k-1} & + & \ldots & + \\
\hline
\tilde{f}: & & & (f_{k-1} - g_{m-1} g_m^{-1} f_k) x^{k-1} & + & \ldots & +
\end{array}$$

The above calculation shows that $\deg \tilde{f} \leq k-1$. By the induction assumption, $P(k-1)$-holds and so there exist $\tilde{q}$ and $\tilde{r} \in R[x]$ with

$$(**) \qquad \tilde{f} = g\tilde{q} + \tilde{r} \quad \text{and} \quad \deg \tilde{r} < \deg g.$$

We compute

$$(***) \qquad \begin{array}{rcll}
f & = & \tilde{f} + g \cdot g_m^{-1} f_k x^{k-m} & - (*) \\
& = & (g\tilde{q} + \tilde{r}) + g \cdot g_m^{-1} f_k x^{k-m} & -(**) \\
& = & (g\tilde{q} + g \cdot g_m^{-1} f_k x^{k-m}) + \tilde{r} & - \textbf{Ax 2}, \textbf{Ax 3} \\
& = & g \cdot (\tilde{q} + g_m^{-1} f_k x^{k-m}) + \tilde{r} & - \textbf{Ax 8}
\end{array}$$

Put $q := \tilde{q} + g_m^{-1} f_k x^{k-m}$ and $r := \tilde{r}$. Then by $(***)$ $f = qg + r$ and by $(**)$, $\deg r = \deg \tilde{r} < \deg g$. Thus $P(k)$ is proved.

By the Principal of Complete Induction 1.4.3 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 $q_i, r_i \in R[x]$ with

$$(+) \qquad f = gq_i + r_i \quad \text{and} \quad \deg r_i < \deg g.$$

Then

$$gq_1 + r_1 = gq_2 + r_2$$

and so

$$(++) \qquad g \cdot (q_1 - q_2) = r_2 - r_1.$$

Suppose $q_1 - q_2 \neq 0_R$ Then $\deg(q_1 - q_2) \geq 0$ and $\mathrm{lead}(q_1 - q_2) \neq 0_R$. Since $\mathrm{lead}(g)$ is a unit in R this implies $\mathrm{lead}(g)\mathrm{lead}(q_1 - q_2) \neq 0_R$, see Exercise 2.8#8. Thus

$$\begin{aligned}
\deg g &\leq \deg g + \deg(q_1 - q_2) && - \deg(q_1 - q_2) \geq 0 \\
&= \deg(g \cdot (q_1 - q_2)) && - \mathrm{lead}(g)\mathrm{lead}(q_1 - q_2) \neq 0_R, 3.1.9(\text{c:}1) \\
&= \deg(r_1 - r_2) && - (++) \\
&\leq \max(\deg r_1, \deg r_2) && - 3.1.9 \\
&< \deg g && - (+)
\end{aligned}$$

This contradiction shows $q_1 - q_2 = 0_R$. Hence, by (++) also $r_2 - r_1 = g \cdot (q_1 - q_2) = g \cdot 0_R = 0_R$. Thus $q_1 = q_2$ and $r_1 = r_2$, see 2.2.9(f). □

Definition 3.1.13. *Let R be a ring and $f, g \in R[x]$ such that* $\mathrm{lead}(f)$ *is a unit in R. Let $q, r \in R[x]$ 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 in R.*

Example 3.1.14. 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.

$$\begin{array}{r|rrrrrrrrr}
 & -x^2 & + & x & - & 1 & & & & \\
\hline
-x^2 + x - 1 & x^4 & + & x^3 & & & - & x & + & 1 \\
 & x^4 & - & x^3 & + & x^2 & & & & \\
\hline
 & & - & x^3 & - & x^2 & - & x & + & 1 \\
 & & - & x^3 & + & x^2 & - & x & & \\
\hline
 & & & & & x^2 & & & + & 1 \\
 & & & & & x^2 & - & x & + & 1 \\
\hline
 & & & & & & & x & &
\end{array}$$

Thus

$$x^4 + x^3 - x + 1 = (-x^2 + x - 1) \cdot (-x^2 + x - 1) + x.$$

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 3.1:

#1. Let P be a polynomial ring in x with coefficients in R. Perform the indicated operation in P and simplify your answer:

(a) $(3x^4 + 2x^3 - 4x^2 + x + 4) + (4x^3 + x^2 + 4x + 3)$ if $R = \mathbb{Z}_5$.

(b) $(x + 1)^3$ if $R = \mathbb{Z}_3$

(c) $(x - 1)^5$ if $R = \mathbb{Z}_5$.

(d) $(x^2 - 3x + 2)(2x^3 - 4x + 1)$ if $R = \mathbb{Z}_7$.

(e) $\left(x + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right)\left(x - \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right)$ if $R = \mathrm{M}_2(\mathbb{R})$.

#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. Give an example in $\mathbb{Z}[x]$ to show that the Division algorithm maybe false if the leading coefficent of g is not a unit.

3.2 Divisibility in $F[x]$

In a general 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:

Theorem 3.2.1. *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 2.4.1 $f = gq$ for some $q \in F[x]$. Thus $f = gq + 0_F$. Since $\deg 0_F = -\infty < \deg g$, Definition 3.1.13 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 2.5.3 $f = gq + 0_F$ for some $q \in F[x]$. Thus $f = gq$ and so Definition 2.4.1 shows that $g|f$. □

Theorem 3.2.2. *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 2.2.9(c), $g = fh = f0_R = 0_R$, contrary to the assumption. Thus $h \neq 0_R$ and so $\deg h \geq 0$. Since R is a field or an integral domain we can apply 3.1.11(a) and conclude

$$\deg g = \deg fh = \deg f + \deg h \geq \deg f.$$

□

Theorem 3.2.3. *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 3.1.8(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 2.12.9.

(e) $\Longrightarrow$ (a): Since f is a unit, $1_F = fg$ for some $g \in F[x]$. Since F is a field we conclude from 3.1.11(a) that

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

and so also $\deg f = \deg g = 0$. □

Theorem 3.2.4. *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 2.12.10.

(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 2.4.3 that $f = 0_F$. Hence $f = 0_F = g$ and so also $\deg g = -\infty = \deg f$ and thus (c) holds. Similarly, (c) holds if $f = 0_F$.

Assume that $f \neq 0_F$ and $g \neq 0_F$. Since $f \mid g$ and $g \mid f$ we conclude from 3.2.2 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 $f \mid g$. If $f = 0_F$, then $\deg g = \deg f = -\infty$ and so $g = 0_F$. Hence $f = g$ and so $f \sim g$ since $\sim$ is an equivalence relation and so reflexive, see 2.12.6.

Thus we may assume $f \neq 0_F$. Since $f \mid g$ we have $g = fh$ for some $h \in F[x]$. Thus by 3.1.11(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 3.2.3, h is a unit. So $g \sim f$ by definition of $\sim$.

(d) $\Longrightarrow$ (a): This holds since $\sim$ is symmetric by 2.12.6. □

Definition 3.2.5. *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 $\check{f} := f \cdot \operatorname{lead}(f)^{-1}$; $\check{f}$ is called the* monic polynomial associated *to f. If $f = 0_F$ put $\check{f} = 0_F$.*

Example 3.2.6. Let $f = 3x^4 + 2x^3 + 4x^2 + x + 2 \in \mathbb{Z}_5[x]$. Then $\operatorname{lead}(f)^{-1} = 3^{-1} = 2$ and

$$\check{f} = (3x^4 + 2x^3 + 4x^2 + x + 2) \cdot 2 = 6x^4 + 4x^3 + 8x^2 + 2x + 4 = x^4 + 4x^3 + 3x^2 + 2x + 4.$$

Theorem 3.2.7. *Let F be a field and $f, g \in F[x]$.*

(a) $f \sim \check{f}$.

(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 2.12.6 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 3.2.3 $\operatorname{lead}(f)$ is a unit in $F[x]$. Hence also $\operatorname{lead}(f)^{-1}$ is a unit. As $\check{f} = f \cdot \operatorname{lead}(f)^{-1}$, this shows that $f \sim \check{f}$.

(b) By definition of $f \sim g$ we have $fu = g$ for some unit u in $F[x]$. By 3.2.3 we have $0_F \neq u \in F$. Hence

$$1_F \overset{g \text{ monic}}{=} \operatorname{lead}(g) \overset{fu=g}{=} \operatorname{lead}(fu) \overset{u \in \mathbb{F}, 3.1.11(b)}{=} \operatorname{lead}(f)u \overset{f \text{ monic}}{=} 1_F u \overset{\textbf{(Ax 10)}}{=} u$$

and so $u = 1_F$ and $g = fu = f1_F = f$.

(c) Suppose $f \neq 0_F$. Then

$$\operatorname{lead}(\check{f}) = \operatorname{lead}\big(f \cdot \operatorname{lead}(f)^{-1}\big) \overset{3.1.11(b)}{=} \operatorname{lead}(f)\operatorname{lead}(f)^{-1} = 1_F.$$

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 $f \sim g$. By (a) $f \sim \check{f}$. As $\sim$ is symmetric and transitive this gives $\check{f} \sim f$ and $\check{f} \sim g$. Since both $\check{f}$ and g are monic we conclude from (b) that $\check{f} = g$.

(d) By (a) $f \sim \check{f}$ and so by 3.2.4 $\deg f = \deg \check{f}$.

(e) By (a) $f \sim \check{f}$ and $g \sim \check{g}$. Thus by 1.5.5

$$(*) \qquad [f]_\sim = [\check{f}]_\sim \quad \text{and} \quad [g]_\sim = [\check{g}]_\sim.$$

Using this we get

$$\begin{array}{rcll} & f & \sim g & \\ \Longleftrightarrow & [f]_\sim & = [g]_\sim & - \ 1.5.5 \\ \Longleftrightarrow & [\check{f}]_\sim & = [\check{g}]_\sim & - \ (*) \\ \Longleftrightarrow & \check{f} & \sim \check{g} & - \ 1.5.5 \end{array}$$

□

Definition 3.2.8. *Let F be a field and $f, g, d \in F[x]$. 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.*

Theorem 3.2.9. *Let F be a field and $f, g, q, r, d, u \in F[x]$. 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. We will verify the three conditions on $d = \gcd(f, g)$.

(i): By definition of a greatest common divisor, $d|g$ and $d|r$. Since $f = gq + ru$ we conclude from 2.4.4(c) that $d|f$. Thus d is a common divisor of f and g.

(ii): Let c be any common divisor of f and g in $F[x]$. Since $f = gq + ru$ and u is a unit we have $r = f \cdot u^{-1} - g \cdot qu^{-1}$. Since c divides f and g we conclude from 2.4.4(c) that $d|r$. So c is a common divisor of g and r. As d is a greatest common divisor of g and r this gives $\deg c \le \deg d$.

(iii): Since $d = \gcd(g, r)$ we know that d is monic.

Thus d is a greatest common divisor of f and g. □

Theorem 3.2.10 (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*

$$\begin{array}{rcccccc} E_{-1} & : & f & = & f \cdot 1_F & + & g \cdot 0_F \\ E_0 & : & \breve{g} & = & f \cdot 0_F & + & g \cdot \mathrm{lead}(g)^{-1} \end{array},$$

Let $i \in \mathbb{N}$ and suppose inductively we defined equations $E_k, -1 \le k \le i$ 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

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

Suppose that $t_{i+1} \neq 0_F$. Then E_{i+1} is 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 $\mathrm{lead}(t_{i+1})^{-1}$. Continue the algorithm with $i+1$ in place of i.

Suppose that $t_{i+1} = 0_F$ and define $d := r_i, u := x_i$ and $v := y_i$. Then

$$d, u, v \in \mathbb{F}[x], \qquad d = \gcd(f, g), \qquad d = fu + gv,$$

and the algorithm stops.

Proof. For $i \in \mathbb{N}$ let $P(i)$ be the following statement:

(a) For $-1 \le k \le i$ 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;

(b) for $-1 \le k \le i$ the equation E_k is true;

(c) r_i is monic;

(d) for all $1 \le k \le i$, $\deg r_k < r_{k-1}$; and

(e) If $d \in F[x]$ with $d = \gcd(r_{i-1}, r_i)$ then $d = \gcd(f, g)$.

We will first show that $P(0)$ holds. Define

$$r_{-1} := f, x_{-1} := 1_F, y_{-1} := 0_F, r_0 := \check{g}, x_0 := 0_F, \text{and} y_0 = \text{lead}(g)^{-1}.$$

Then for $k = -1$ and $k = 0$, E_k is the equation $r_k = f \cdot x_k + g \cdot y_k$ and so (a) holds for $i = 0$. Also E_{-1} and E_0 are true, so (b) holds for $i = 0$. Note that $r_0 = \check{g}$ is monic and so (c) holds for $i = 0$. There is no integer k with $1 \leq k \leq 0$ and thus (d) holds for $i = 0$. Assume $d \in F[x]$ with $d = \gcd(r_{-1}, r_0)$. Then $d = \gcd(f, \check{g})$. Note that $g = f \cdot 0_R + \check{g} \cdot \text{lead}(g)$. As $\text{lead}(g)$ is a unit in $F[x]$ we conclude from 3.2.9 that $d = \gcd(f, g)$. Thus (e) holds for $i = 0$. Hence $P(0)$ holds.

Suppose now that $i \in \mathbb{N}$ and that $P(i)$ holds. Then the equations

$$\begin{array}{lcllcl} E_{i-1} & : & r_{i-1} & = & f \cdot x_{i-1} & + \quad g \cdot y_{i-1} \quad \text{and} \\ E_i & : & r_i & = & f \cdot x_i & + \quad g \cdot y_i. \end{array}$$

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 $u_{i+1} = (\text{lead} t_{i+1})^{-1}$. Multiplying the preceding equation with u_{i+1} gives the true equation

$$E_{i+1} \quad : \quad (r_{i-1} - r_i q_{i+1}) u_{i+1} \quad = \quad f \cdot (x_{i-1} - x_i q_{i+1}) u_{i+1} \quad + \quad g \cdot (y_{i-1} - y_i q_{i+1}) u_{i+1}.$$

Define

$$r_{i+1} := (r_{i-1} - r_i q_{i+1}) u_{i+1}, x_{i+1} := (x_{i-1} - x_i q_{i+1}) u_{i+1}, \text{and} y_{i+1} := (y_{i-1} - y_i q_{i+1}) u_{i+1}.$$

Then 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 (a) and (b) hold for $i+1$ in place of i.

By $(*)$ we have $t_{i+1} = r_{i-1} - r_i q_{i+1}$ and so

$$r_{i+1} = (r_{i-1} - r_i q_{i+1}) u_{i+1} = t_{i+1} u_{i+1} = t_{i+1} \text{lead}(t_{i+1})^{-1} = \check{t}_{i+1}.$$

Hence

$$r_{i+1} = \check{t}_{i+1}.$$

Thus r_{i+1} is monic and (c) holds. Moreover, $t_{i+1} = r_{i+1} \text{lead}(t_{i+1})$ and (*) gives

$$r_{i-1} = r_i q_i + r_{i+1} \text{lead}(t_{i+1}).$$

Hence, if $d \in F[x]$ with $d = \gcd(r_i, r_{i+1})$, we conclude from 3.2.9 that $d = \gcd(r_{i-1}, r_i)$. As $P(i)(e)$ holds, this gives $d = \gcd(f, g)$ and so (e) 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 (d) ensures that the algorithm stops in finitely many steps.

Suppose next that $t_{i+1} = 0_F$. We will show that $r_i = \gcd(r_i, 0_F)$. Clearly r_i is a common divisor of r_i and 0_F. If c is a common divisor of r_i and 0_F in $F[x]$, then $c \mid r_i$ and 3.2.2 shows that $\deg c \leq \deg r_i$. By P(i)(c) we know that r_i is monic. So indeed $r_i = \gcd(r_i, 0_F)$. As $t_{i+1} = 0_F$, $(*)$ implies that $r_{i-1} = r_i q_i + 0_F$ and so 3.2.9 shows that $r_i = \gcd(r_{-i}, r_i)$. As $P(i)(e)$ 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 and $r_i, x_i, v_i \in F[x]$. So putting $d := r_i, u := x_i$ and $v := y_i$ we have

$$d, u, v \in F[x], \quad d = \gcd(f, g) \quad \text{and} \quad fu + gv.$$

□

Example 3.2.11. 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 is an 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 $r_0 = \breve{g} = 3g = 6x^3 + 3x^2 + 6x + 9 = x^3 + 3x^2 + x + 4$.

$$\begin{array}{rcrcccc} E_{-1} & : & 3x^4 + x^3 + 2x^2 + x + 1 & = & f \cdot 1 & + & g \cdot 0 \\ E_0 & : & x^3 + 3x^2 + x + 4 & = & f \cdot 0 & + & g \cdot 3 \end{array},$$

$$\begin{array}{r|ccccccccc} & 3x & & & & & & & & \\ \hline x^3 + 3x^2 + x + 4 & 3x^4 & + & 4x^3 & + & 2x^2 & + & x & + & 1 \\ \hline & 3x^4 & + & 9x^3 & + & 3x^2 & + & 2x & & \\ \hline & & & & & -x^2 & & -x & + & 1 \end{array}$$

Subtracting $3x$ times E_0 from E_{-1} we get

$$-x^2 - x + 1 \quad = \quad f \cdot 1 \quad + \quad g \cdot -9x \qquad \mid \quad E_{-1} - E_0 \cdot 3x$$

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 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

$$\gcd(f,g) = 1 = f \cdot (x+2) + g \cdot (x^2 + 2x + 3)$$

Remark 3.2.12. *Let F be a field and $f,g,d \in F[x]$ with $d = \gcd(f,g)$. Then $f \neq 0_F$ and $g \neq 0_F$.*

Proof. Suppose for a contradiction that $f = 0_F$ and $g = 0_F$. Choose $n \in \mathbb{N}$ with $n > \deg d$. Then x^n is a common divisor of f and g and $\deg x^n = n > \deg g$, a contradiction to the definition of 'gcd'. □

Theorem 3.2.13. *Let F be a field and $f,g \in F[x]$ not both 0_F.*

(a) *There exists a unique $d \in F[x]$ with $d = \gcd(f,g)$.*

(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 3.2.10 there exists $u,v,d \in F[x]$ such that $d = \gcd(f,g)$ and $d = fu + gv$. This proves the existence of d and also proves (b).

To prove (c) let c be any common divisor of a and b. Since $d = fu + gv$ we conclude from 2.4.4(c) 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 $\deg e \leq \deg d$ and $\deg e \leq \deg d$. Thus $\deg d = \deg e$. Since $e|d$ we conclude from 3.2.4 that $d \sim e$. As d and e are monic this implies that $d = e$, see 3.2.7(b). Thus d is the unique greatest common divisor of f and g. □

Theorem 3.2.14. *Let F be a field and $f, g \in F[x]$. Then $1_F = \gcd(f, g)$ if and only if there exist $u, v \in F[x]$ with $fu + gv = 1_F$.*

Proof. $\Longrightarrow$: Suppose that $1_F = \gcd(f, g)$. By 3.2.12 f and g are not both 0_F and so 3.2.13(c) shows 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 2.4.4(c)). Hence $\deg c \leq \deg 1_F$ by 3.2.2. Thus $1_F = \gcd(f, g)$. □

Theorem 3.2.15. *Let F be a field and $f, g, h \in F[x]$. Suppose that $1_F = \gcd(f, g)$ and $f \mid gh$. Then $f \mid h$.*

Proof. Since $1_F = \gcd(f, g)$ we conclude from 3.2.14 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, 2.4.4(c) now implies that $f \mid h$. □

Exercises 3.2:

#1. Let F be a field and $a, b \in F$ with $a \neq b$. Show that $1_F = \gcd(x + a, x + b)$.

#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 $f \mid g$ 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 $1_F = \gcd(f, g)$. If $f \mid h$ and $g \mid h$, prove that $fg \mid h$.

#5. Let F be a field and $f, g, h \in F[x]$. Suppose that $g \neq 0_F$ and $1_F = \gcd(f, g)$. Show that $\gcd(fh, g) = \gcd(h, g)$.

#6. Let F be a field and $f, g, d \in F[x]$ such that $h \neq 0_F$ and $d = \gcd(f, g)$.

(a) Show that there exist $\hat{f}, \hat{g} \in F[x]$ with $f = \hat{f}d$ and $g = \hat{g}d$.

(b) Show that $\gcd(\hat{f}, \hat{g}) = 1_F$.

#7. Let F be a field and $f, g, h \in F[x]$ with $f \mid gh$. Show that there exist $\tilde{g}, \tilde{h} \in F[x]$ with $\tilde{g} \mid g, \tilde{h} \mid h$ and $f = \tilde{g}\tilde{h}$.

3.3 Irreducible Polynomials

Definition 3.3.1. *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 $g \mid f$, then*

$$g \sim 1_F \quad \text{or} \quad g \sim f.$$

(c) *f is called* reducible *provided that*

   (i) *$f \neq 0_F$, and*

   (ii) *there exists $g \in F[x]$ with*

$$g \mid f, \quad g \nsim 1_F, \quad \text{and} \quad g \nsim f.$$

Remark 3.3.2. *Let F be a field and $f \in F[x]$. Then the following statements are equivalent:*

(a) *f is not constant.*

(b) $\deg f \geq 1$.

(c) *$f \neq 0_F$ and $f \nsim 1_F$.*

Proof. We will show that the negation of the three statements are equivalent.

f is constant if and only if $f \in F$ and if and only if $\deg f \leq 0$, see 3.1.8(c).

Since $\deg f \in \mathbb{N}^*$, we have $\deg f \leq 0$ if and only if $\deg f < 1$ and if and only of $\deg f = -\infty$ or $\deg f = 0$.

Also $\deg f = -\infty$ or $\deg f = 0$ if and only if $f = 0_F$ or $f \sim 1_F$, see 3.1.8(c) and 3.2.3. □

Theorem 3.3.3. *Let F be a field and $f \in F[x]$. 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. By definition of 'reducible' we conclude that $f \neq 0_F$ and there exists $g \in F[x]$ 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 2.4.3). As $g \neq 0_F$ and $g \nsim 1_F$, Remark 3.3.4 shows that g is not constant. Since $g|f$ we have $\deg g \leq \deg f$, see 3.2.2. Suppose that $\deg g = \deg f$. Since $g|f$ we get from 3.2.4 that $g \sim f$, a contradiction. Thus $\deg g \neq \deg f$ and and so $\deg g < \deg f$. Hence g is a non-constant polynomials of lower degree than f which divides f. So (b) holds.

(b) $\Longrightarrow$ (c): Let g be a non-constant polynomial of lower degree than f with $g \mid f$. Then $\deg g > 0$, $\deg g < \deg f$ and $f = gh$ for some $h \in F[x]$. From $\deg g < \deg f$ we get $\deg f \neq -\infty$ and $f \neq 0_F$. As $f = gh$ we conclude that $h \neq 0_F$. By 3.1.11(a) $\deg f = \deg g + \deg h$ and since $\deg g > 0$ this gives $\deg h < \deg f$. We proved that $f = gh$, $\deg g < \deg f$ and $\deg h < \deg f$. Thus (c) holds.

(c) $\Longrightarrow$ (d): Suppose $f = gh$ with $\deg g < \deg f$ and $\deg h < \deg f$. By 3.1.11 $\deg f = \deg g + \deg h$. 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 $g|f$. As g and h are non-constant we have $g \nsim 1_F$, $\deg g \geq 1$ and $\deg h \geq 1$, see Remark 3.3.4 By 3.1.11(a) we know that $\deg f = \deg g + \deg h$ and so $\deg f > \deg g \geq 1$. Thus $f \neq 0_F$ and $\deg f \neq \deg g$. By 3.2.4 the latter statement gives $g \nsim f$.

We proved that $f \neq 0_F$, $g|f$, $g \nsim 1_F$ and $g \nsim f$. Thus the definition of 'irreducible' shows that f is not irreducible. So (f) holds.

(f) $\Longrightarrow$ (a): Suppose f is not constant and f is not irreducible. Since f is not irreducible, the statement

$$\text{If } f \in F[x] \text{ with } g|f, \text{ then } g \sim 1_F \text{ or } g \sim f$$

must be false. Hence there exists $g \in F[x]$ with $g|f$, $g \nsim 1_F$ and $g \nsim f$. The definition of 'reducible' now shows that f is reducible. Thus (a) holds. □

Remark 3.3.4. *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 $f \in F[x]$. Then 3.3.3(a),(f) shows that

$$(*) \qquad f \text{ is reducible} \quad \text{if and only if} \quad f \text{ non-constant and } f \text{ is not irreducible.}$$

(a): Let f be non-constant polynomial in $F[x]$. Then $(*)$ shows that f is reducible if and only if f is not irreducible.

(b): By definition irreducible polynomials are not constant and by $(*)$ reducible polynomials are not constant. Thus constant polynomials are neither irreducible nor constant. □

Theorem 3.3.5. *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 we get that

$$d \sim 1_F \qquad \text{or} \qquad d \sim p.$$

We treat these two cases separately:

Suppose that $d \sim 1_F$. Since both d and 1_F are monic we conclude from 3.2.7 that $d = 1_F$. As $p|gh$ this implies $p|h$, see 3.2.15

Suppose that $d \sim p$. As $d|g$ this gives $p|g$, see 2.12.10(a).

(b) $\Longrightarrow$ (c): Suppose (b) holds and let $g, h \in F[x]$ with $p = gh$. Note that $p1_F = p = gh$ and so $p|gh$. From (b) we conclude $p|g$ or $p|h$. Since the situation is symmetric in g and h we may assume $p|g$. As p is not constant we have $p \neq 0_F$. Since $p = gh$ this gives $g \neq 0_F$ and $h \neq 0_F$. As $p|g$ we have $\deg p \leq \deg g$ by 3.2.2. On the other hand by 3.1.11(a), $\deg p = \deg gh = \deg g + \deg h \geq \deg g$. Thus $\deg g = \deg p$ and $\deg h = 0$. So h is constant.

(a) $\Longleftrightarrow$ (c): Since p is not constant, p is irreducible if and only if p is not reducible, see 3.3.4(a). By 3.3.3 p is reducible if and only if p is the product of two non-constant polynomials. Thus p is irreducible if and only if p is not the product of two non-constant polynomials. The latter statement is equivalent to (c). □

Theorem 3.3.6. *Let F be a field and let p be an irreducible polynomial in $F[x]$. If $a_1, \dots, a_n \in F[x]$ and $p|a_1 a_2 \dots a_n$, 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 $p|a_1 \dots a_k a_{k+1}$. By 3.3.5, $p|a_1 \dots a_k$ 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 theorem holds for $k+1$ and so by the Principal of Mathematical Induction (1.4.2) the theorem holds for all positive integers n. □

Theorem 3.3.7. *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 2.12.8. So suppose that $p|q$. Since q is irreducible we have $p \sim 1_F$ or $p \sim q$. As p is irreducible, p is not constant and 3.3.4 shows that $p \nsim 1_F$. Thus $p \sim q$. □

Theorem 3.3.8. *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. Since $f \sim g$ we know that $\deg f = \deg g$ and that f and g have the same divisor, see 3.2.4 and 2.12.10(b).

f is irreducible if and only if f is non-constant and not reducible. This holds if and only if $\deg f \geq 1$ and f is not divisible by a non-constant polynomial of lower degree, see 3.3.3. The latter statement holds if and only if $\deg g \geq 1$ and g is not divisible by a polynomial of lower degree, and so if and only if g is irreducible. □

Theorem 3.3.9 (Factorization Theorem)**.** *Let F be a field and f a non-constant polynomial in $F[x]$. Then f is the product of irreducible polynomials in $F[x]$.*

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 is also non-constant we conclude from 3.3.3 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. Since $f = gh$ this shows that f is the product of irreducible polynomials.

Theorem 3.3.10 (Unique Factorization Theorem)**.** *Let F be a field and f a non-constant polynomial in $F[x]$. Suppose that 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*

$$f = p_1 p_2 \ldots p_n \quad \text{and} \quad f = q_1 q_2 \ldots q_m.$$

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 $\pi : \{1, \ldots n\} \mapsto \{1, \ldots m\}$ 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. The proof is by complete induction on n. So let k be a positive integer and suppose that the theorem holds whenever $n < k$. We will show that the theorem holds for $n = k$. So suppose 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 3.3.5 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 $f = (p_1 \dots p_{k-1})p_k$ we see that p_k divides f. By (*) $f = q_1 \dots q_m$ and so p_k divides $q_1 \dots q_m$. Hence by 3.3.6, $p_k | q_j$ for some $1 \leq j \leq m$. As p_k and q_j are irreducible we get from 3.3.7 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

$$\big((p_1 u)p_2 \dots p_{k-1}\big)q_m = (p_1 \dots p_{k-1})(q_m u) = p_1 \dots p_{k-1}p_k = f = (q_1 \dots q_{m-1})q_m.$$

By 3.1.11(c) $F[x]$ is an integral domain. Since q_m is irreducible, q_m is not constant and so $q_m \neq 0_F$. Hence the Multiplicative Cancellation Law for Integral Domains 2.8.7 gives

$$(p_1 u)p_2 \dots p_{k-1} = q_1 \dots q_{m-1}.$$

Since u is a unit, $p_1 u \sim p_1$. As p_1 is irreducible we conclude that also $p_1 u$ is irreducible, see by 3.3.8. The induction assumption now implies that $k - 1 = m - 1$ and that, after reordering the q_i's,

$$p_1 u \sim q_1, \quad p_2 \sim q_2, \quad \dots \quad p_{k-1} \sim q_{k-1}.$$

From $k - 1 = m - 1$ we get $k = m$. As $p_1 \sim p_1 u$ and $p_1 u \sim q_1$ 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 the theorem holds for $n = k$. By the principal of complete induction, the theorem holds for all positive integers n. □

Exercises 3.3:

#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 $\gcd(f,g) = p_1^{m_1}p_2^{m_2}\dots p_n^{m_n}$.

3.4 Polynomial function

Theorem 3.4.1. *Let R and S be commutative rings with identities, let $\alpha : R \to S$ 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

$$\begin{aligned}
\beta(f) &= \beta\left(\sum_{i=0}^{\deg f} f_i x^i\right) && -3.1.8\text{(d)} \\
&= \sum_{i=0}^{\deg f} \beta(f_i x^i) && -\beta \text{ respects addition} \\
&= \sum_{i=0}^{\deg f} \beta(f_i)\beta(x)^i && -\beta \text{ respects multiplication} \\
&= \sum_{i=0}^{\deg f} \alpha(f_i)s^i. && -(*)
\end{aligned}$$

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

$$\alpha_s : \quad R[x] \to S, \quad f \mapsto \sum_{i=0}^{\deg f} \alpha(f_i)s^i.$$

By hypothesis $\alpha(1_R) = 1_S$. 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 $m = \deg f + \deg g$.

$$\begin{array}{rcll}
\alpha_s(f+g) & = & \alpha_s\left(\sum_{i=0}^{n}(f_i+g_i)x^i\right) & \text{– 3.1.5(a) with } R[x] \text{ in place of } P \\
& = & \sum_{i=0}^{n}\alpha(f_i+g_i)s^i & \text{– definition of } \alpha_s \\
& = & \sum_{i=0}^{n}\Big(\alpha(f_i)+\alpha(g_i)\Big)s^i & \text{– Since } \alpha \text{ respects addition} \\
& = & \left(\sum_{i=0}^{\deg f}\alpha(f_i)s^i\right)+\left(\sum_{i=0}^{\deg g}\alpha(g_i)s^i\right) & \text{– 3.1.5(a) with } (S,S,x) \text{ in place of } (R,P,x) \\
& = & \alpha_s(f)+\alpha_s(g) & \text{– definition of } \alpha_s\text{, twice}
\end{array}$$

$$\begin{array}{rcll}
\alpha_s(fg) & = & \alpha_s\left(\sum_{k=0}^{m}\left(\sum_{i=0}^{k} f_i g_{k-i}\right)x^k\right) & \text{– 3.1.5(a) with } R[x] \text{ in place of } P \\
& = & \sum_{k=0}^{m}\alpha\left(\sum_{i=0}^{k} f_i g_{k-i}\right)s^k & \text{– definition of } \alpha_s \\
& = & \sum_{k=0}^{m}\left(\sum_{i=0}^{k}\alpha(f_i)\alpha(g_{k-i})\right)s^k & \text{– } \alpha \text{ respects addition and multiplication} \\
& = & \left(\sum_{i=0}^{\deg f}\alpha(f_i)s^i\right)\cdot\left(\sum_{j=0}^{\deg g}\alpha(g_j)s^j\right) & \text{– 3.1.5(a) with } (S,S,x) \text{ in place of } (R,P,x) \\
& = & \alpha_s(f)\cdot\alpha_s(g) & \text{– definition of } \alpha_s\text{, twice}
\end{array}$$

So α_s is a homomorphism and the theorem is proved. □

Example 3.4.2. Let R and S be commutative rings with identities, $\alpha : R \to S$ a ring homomorphism with $\alpha(1_R) = 1_S$ and $s \in S$. Compute α_s in each the following cases:

(1) $S = R$ and $\alpha(r) = r$.

$$\alpha_s(f) = \sum_{i=0}^{\deg f} \alpha(f_i)s^i = \sum_{i=0}^{\deg f} f_i s^i.$$

(2) $S = R[x]$, $\alpha(r) = r$ and $s = x$.

$$\alpha_s(f) = \sum_{i=0}^{\deg f} \alpha(f_i)s^i = \sum_{i=0}^{\deg f} f_i x^i = f.$$

So α_s is identity function on $R[x]$.

(3) $n \in R$, $S = R_n[x]$, $\alpha(r) = [r]_n$ and $s = x$.

Note first that by Example 2.11.2(4) the function $\alpha : R \to R_n[x], r \mapsto [r]_n$ really 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.

For example if $R = \mathbb{Z}$ and $n = 3$, then

$$\begin{aligned} \alpha_x(6x^3 + 5x^2 + 10x + 9) &= [6]_3\, x^3 + [5]_3\, x^2 + [10]_3\, x + [9]_3 = [0]_3 x^3 + [2]_3 x^2 + [1]_3 x + [0]_2 \\ &= \big(\text{in } \mathbb{Z}_3[x]\big) \quad 2x^2 + x. \end{aligned}$$

Definition 3.4.3. *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) $\mathrm{Fun}(R) = \mathrm{Fun}(R, R)$.

Theorem 3.4.4. *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* $\tau : R \to \mathrm{Fun}(I, R), r \mapsto r^*$ *is a homomorphism. If* $I \neq \varnothing$, *then* τ *is injective.*

Proof. (a)-(d): See Exercise 1 on Homework 5 or F.1.2 in the Appendix.

(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 1.3.14 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 1.3.14 and so $\tau(ab) = \tau(a)\tau(b)$ by definition of τ.
Thus τ is a homomorphism .

Suppose in addition that $I \neq \varnothing$. To show that τ is injective let $a, b \in R$ with $\tau(a) = \tau(b)$. Then $a^* = b^*$. Since $I \neq \varnothing$ we can pick $i \in I$. Then

$$a = a^*(i) = b^*(i) = b$$

and so τ is injective. □

Notation 3.4.5. *Let R be a commutative ring with identity and $f \in R[x]$. For $f = \sum_{i=0}^{\deg f} f_i x^i \in F[x]$ let f^* be the function*

$$f^* : \quad R \to R, \quad r \mapsto \sum_{i=0}^{\deg f} f_i r^i.$$

f^* is called the polynomial function *on R induced by f.*

Remark 3.4.6. *Let R be a commutative ring with identity.*

(a) *Let $\mathrm{id} : R \to R, r \mapsto r$ be the identity function on R and for $r \in R$ let $\mathrm{id}_r : R[x] \to R$ be the homomorphism from 3.4.1. Then*

$$f^*(r) = \mathrm{id}_r(f)$$

for all $f \in R[x]$ and $r \in R$.

(b) *Let $f \in R$ be constant polynomial. Then the definitions of $f^* \in \mathrm{Fun}(R)$ in 3.4.5 and in 3.4.3 coincide.*

Proof. (a): By Example 3.4.2(1) $\mathrm{id}_r(f) = \sum_{i=0}^{\deg f} f_i r^i$ and so $\mathrm{id}_r(f) = f^*(r)$.

(b) Since $f \in F$ we have $f = fx^0$ and $f^*(r) = fr^0 = f1_R = f$ 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 3.4.7. 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 3.4.8. *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 $R[x] \to \mathrm{Fun}(R)$, $f \mapsto f^*$ is a ring homomorphism.*

Proof. (a) By definition f^* is a function from R to R. Hence $f^* \in \mathrm{Fun}(R)$.

(b)

$$\begin{aligned}(f+g)^*(r) &= \mathrm{id}_r(f+g) && -3.4.6(a)\\ &= \mathrm{id}_r(f)+\mathrm{id}_r(g) && -\ \mathrm{id}_r \text{ is a homomorphism}\\ &= f^*(r)+g^*(r) && -3.4.6(a), \text{twice}\end{aligned}$$

and similarly

$$\begin{aligned}(fg)^*(r) &= \mathrm{id}_r(fg) && -3.4.6(a)\\ &= \mathrm{id}_r(f)\mathrm{id}_r(g) && -\ \mathrm{id}_r \text{ is a homomorphism}\\ &= f^*(r)g^*(r) && -3.4.6(a), \text{twice}\end{aligned}$$

(c) Let $r \in R$. Then

$$\begin{aligned}(f+g)^*(r) &= f^*(r)+g^*(r) && -\ \text{(b)}\\ &= (f^*+g^*)(r) && -\ \text{Definition of addition in Fun}(R)\end{aligned}$$

So $(f+g)^* = f^*+g^*$. Similarly

$$\begin{aligned}(fg)^*(r) &= f^*(r)g^*(r) && -\ \text{(b)}\\ &= (f^*g^*)(r) && -\ \text{Definition of multiplication in Fun}(R)\end{aligned}$$

and so $(fg)^* = f^*g^*$.

(d) Follows from (c). □

Theorem 3.4.9. *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 $t \in R$.

$$\begin{aligned}
f^*(a) &\overset{(*)}{=} (q\cdot(x-a)+r)^*(a) \overset{3.4.8(b)}{=} (q\cdot(x-a))^*(a)+r^*(a)\\
&\overset{3.4.8(b)}{=} q^*(a)\cdot(x-a)^*(a)+r^*(a) \overset{\text{Def }(x-a)^*}{=} q^*(a)(a-a)+r^*(a)\\
&\overset{(**)}{=} q^*(a)(a-a)+r \overset{2.2.9(f)}{=} q^*(a)\cdot 0_F+r\\
&\overset{2.2.9(c)}{=} 0_F+r \overset{\textbf{Ax 4}}{=} r
\end{aligned}$$

□

Definition 3.4.10. *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 3.4.11 (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

$$\begin{aligned}
& x-a\,|\,f \\
\iff\ & r = 0_F && -\ 3.2.1\\
\iff\ & f^*(a) = 0_F && -f^*(a) = r \text{ by } 3.4.9\\
\iff\ & a \text{ is a root of } f && -\ \text{Definition of root}
\end{aligned}$$

□

Theorem 3.4.12. *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

$$f^*(a) = (gh)^*(a) \overset{3.4.8(c)}{=} g^*(a)h^*(a) = 0_R \cdot h^*(a) = 0_R.$$

Thus a is a root of f. So (a) holds.

(b) : Suppose that R is field or an integral domain. By 2.8.10 all fields are integral domains. Thus R is an integral domain and so **(Ax 11)** holds. Hence

	a is a root of f	
$\iff$	$f^*(a) = 0_R$	– definition of root
$\iff$	$(gh)^*(a) = 0_R$	$-f = gh$
$\iff$	$g^*(a)h^*(a) = 0_R$	–3.4.8(b)
$\iff$	$g^*(a) = 0_R$ or $h^*(a) = 0_R$	–**(Ax 11)**
$\iff$	a is a root of g or a is a root of h	–definition of root, twice

□

Example 3.4.13. (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 unique 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, 3.4.12 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 $\mathbb{Z}_8$

$$0^2 - 1 = -1, (\pm 1)^2 - 1 = 1 - 1 = \boxed{0}, (\pm 2)^2 - 1 = 4 - 1 = 3, (\pm 3)^2 = 9 - 1 = 8 = \boxed{0}, 4^2 - 1 = 15 = -1.$$

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 3.4.14 (Root Theorem)**.** *Let F be a field and $f \in F[x]$ a non-zero polynomial. Then there exist $m \in \mathbb{N}$, elements $a_1, \dots, a_m \in F$ and $q \in F[x]$ such that*

(a) $q \neq 0_F$ *and* $\deg f = \deg q + m$*, in particular,* $m \le \deg f$*,*

(b) $f = q \cdot (x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_m)$,

(c) *q has no roots in F, and*

(d) *$\{a_1, a_2, \dots, a_m\}$ 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 3.4.11, $x - a | f$ and so

$$f = (x-a)\cdot g = g \cdot (x-a) \tag{$*$}$$

for some $g \in F[x]$. By 3.1.11

$$\deg f = \deg g + \deg(x-a) = \deg g + 1 \tag{$**$}$$

and so $\deg g = k - 1 < k$. Hence by the induction assumption there exist $n \in \mathbb{N}$, elements $a_1, \dots, a_n \in F$ and $q \in F[x]$ such that

(A) $q \neq 0_F$ and $\deg g = \deg q + n$,

(B) $g = q \cdot (x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_n)$,

(C) q has no roots in F,

(D) $\{a_1, a_2, \dots, a_n\}$ is the set of roots of g.

Put

$$m := n + 1 \qquad \text{and} \qquad a_m := a. \tag{$***$}$$

Then

$$\deg f \overset{(**)}{=} \deg g + 1 \overset{(A)}{=} \deg q + n + 1 \overset{(***)}{=} \deg q + m,$$

so (a) holds.

We have

$$f \overset{(*)}{=} g \cdot (x - a_m) \overset{(C)}{=} q \cdot (x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_n) \cdot (x - a_m).$$

and since $n = m - 1$ we see that (b) holds.

By (C) q has no roots and so (c) holds.

Let $b \in F$. Since $f = g \cdot (x - a_m)$, 3.4.12 shows that b is a root of f if and only if b is a root of g or g is a root of $x - a_m$. By (D) the roots of g are $a_1, a_2, \dots a_n$ and by 3.4.13(1) the root of $x - a_m$ is a_m. Thus the set of roots of f is $\{a_1, a_2 \dots, a_n, a_m\} = \{a_1, \dots, a_m\}$. Hence also (d) is proved. □

Remark 3.4.15. *$x^2 - 1$ has four roots in $\mathbb{Z}_8$, namely ± 1 and ± 3, see Example 3.4.13(3). So in rings without* **(Ax 11)** *a polynomial can have more roots than its degree.*

Theorem 3.4.16. *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

□

Exercises 3.4:

#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.

(d) Find an example for a field F and $f \in F[x]$ such that f is reducible and f has no root in F.

#2. Let F be an infinite field.

(a) Let $f, g \in F[x]$ with $f^* = g^*$. Show that $f = g$. *Hint:* What are the roots of $f - g$?

(b) Show that the function $F[x] \to \mathrm{Fun}(F), f \mapsto f^*$ is an injective homomorphism.

#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 3.4.14 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) $f = x^{m_f} \cdot (f^\bullet)^\bullet$

(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.

3.5 The Congruence Relation

Notation 3.5.1. *Let F be a field and $f, g, p \in F[x]$. Recall from Definition 2.4.5 that the relation $\equiv \pmod p$ is defined by*

$$f \equiv g \pmod p \qquad \text{if} \qquad p \mid f - g.$$

By 2.4.7 this relation is an equivalence relation. By 2.4.8 $[f]_p$ is denotes the equivalence class of '$\equiv \pmod p$' containing f, and $[f]_p$ is called the congruence class *of f modulo p. So*

$$[f]_p = \{g \in F[x] \mid f \equiv g \pmod p\}$$

$F[x]_p$ denotes the set of congruence classes modulo p in $F[x]$. We will also use the notation $F[x]/(p)$ for $F[x]_p$. So

$$F[x]/(p) = F[x]_p = \left\{ [f]_p \;\middle|\; f \in F[x] \right\}$$

Example 3.5.2. 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 3.2.1 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

$$
\begin{array}{r|ccccccc}
 & x & + & 1 & & & & \\ \hline
x^2+x+1 & x^3 & & & + & x & + & 1 \\
 & x^3 & + & x^2 & + & x & & \\ \hline
 & & & x^2 & & & + & 1 \\
 & & & x^2 & + & x & + & 1 \\ \hline
 & & & & & x & &
\end{array}
$$

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 3.5.3. *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) $[f]_p \cap [g]_p \neq \varnothing$.

(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. By 2.4.9 the 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

$$
\begin{array}{lcl}
f = pq_1 + r_1 & \text{and} & \deg r_1 < \deg p \\
g = pq_2 + r_2 & \text{and} & \deg r_2 < \deg p
\end{array}
$$

(m) $\Longrightarrow$ (l): 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 (l) holds with $l = q_2 - q_1$.

(a) $\Longrightarrow$ (m): Suppose $f = g + pk$ for some $k \in F[x]$. Then $f = (pq_2 + r_2) + pk = p\cdot(q_2+k) + r_2$. 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. Hence $r_1 = r_2$ and (m) holds. □

Theorem 3.5.4. *Let F be a field and $p \in F[x]$ with $p \neq 0_F$.*

(a) *Let $f \in F[x]$. 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.*

(b) *The function*

$$\rho: \quad \left\{ r \in F[x] \,\middle|\, \deg r < \deg p \right\} \to F[x]/(p), \quad r \mapsto [r]_p$$

is a bijection.

(c) $F[x]/(p) = \left\{ [r]_p \,\middle|\, r \in F[x], \deg r < \deg p \right\}$

Proof. (a): 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 3.5.3, $[f]_p = [r]_p$ if and only f and s have the same remainder when divided by n, and so if and only if $s = r$.

(b) The uniqueness assertion in (a) shows that ρ is injective. Let $A \in F[x]/(p)$. By definition of $F[x]/p$ the exists $f \in F[x]$ with $A = [f]_p$. By (a) there exists $r \in F[x]$ with $[f]_p = [r]_p$ and $\deg r < \deg p$. Then $\rho(r) = [r]_p = [f]_p = A$ and so ρ is surjective.

(c) This holds since ρ is surjective. □

Example 3.5.5. 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 3.5.4(c) shows that

$$\begin{aligned}\mathbb{Z}_3[x]/(x^2+1) &= \left\{ [f]_p \mid f \in \mathbb{Z}_2[x], \deg f < 2 \right\} \\ &= \left\{ [0]_p,\ [1]_p,\ [2]_p,\ [x]_p,\ [x+1]_p,\ [x+2]_p,\ [2x]_p,\ [2x+1]_p,\ [2x+2]_p \right\}.\end{aligned}$$

(b) Any polynomial of degree less than 3 in $\mathbb{Q}[x]$ can be uniquely written as $a + bx + cx^2$ with $a, b, c \in \mathbb{Q}$. Thus

$$\mathbb{Q}[x]/(x^3 - x + 1) = \{[a + bx + cx^2]_{x^3-x+1} \mid a,b,c \in \mathbb{Q}\}.$$

Exercises 3.5:

#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 $g \equiv 0_F \pmod{p}$.

3.6 Congruence Class Arithmetic

Remark 3.6.1. *Let F be a field and $p \in F[x]$. Recall from 2.6.2 that we defined an addition and multiplication on $F[x]/(p)$ by*

$$[f]_p + [g]_p = [f+g]_p \quad \text{and} \quad [f]_p \cdot [g]_p = [f \cdot g]_p$$

for all $f, g \in F[x]$.

Example 3.6.2. 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 3.5.4(c) 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

$$[x+1][x+1] = [(x+1)(x+1)] = [x^2+1] = [(x+1)+(x^2+x)] = [x+1]$$

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 3.6.3. *Let F be a field and $p \in F[x]$.*

(a) *$F[x]/(p)$ is a commutative ring with identity $[1_F]_p$.*

(b) *The function*

$$\sigma: \quad F[x] \to F[x]/(p), \quad f \mapsto [f]_p.$$

is an surjective homomorphism of rings.

(c) *Put $\hat{F} = \{[a]_p | a \in F\}$. Then $\hat{F}$ is a subring of $F[x]/(p)$.*

(d) *Suppose p is not constant. 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) This is a special case of 2.6.4.

(b) This is a special case of Example 2.11.2(4).

(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 9 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 injective and surjective homomorphism. By (b), 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 surjective.
By 3.5.4(b) the function

$$\rho: \quad \left\{ r \in F[x] \,\middle|\, \deg r < \deg p \right\} \to F[x]/(p), \quad r \mapsto [r]_p$$

is a bijection and so injective. Let $a \in F$. Since p is not constant, $\deg p \geq 1$ and so $\deg a \leq 0 < \deg p$. Thus F is contained in the domain on ρ. Since $\tau(a) = [a]_p = \rho(a)$ this shows that also τ is injective. Thus (d) holds. $\square$

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 3.6.4. *Let F be a field and p be a non-constant polynomial in $F[x]$. 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 and* $1_R = 1_F$.

Proof. As in 3.6.3 put $\hat{F} := \{[a]_p \mid a \in F\}$. Define

$$S := F[x]/(p) \smallsetminus \hat{F} \qquad \text{and} \qquad R := F \cup S.$$

(So for $a \in F$ we removed $[a]_p$ from $F[x]/(p)$ and replaced it by a.) Define

$$\Phi : \quad R \to F[x]/(p), \quad r \mapsto \begin{cases} [r]_p & \text{if } r \in F \\ r & \text{if } r \in S \end{cases}$$

Note that $R = F \cup S$, $S \cap F = \varnothing$, $F[x]/(p) = \hat{F} \cup S$ and $\hat{F} \cap S = \varnothing$. By 3.6.3 the function $F \to \hat{F}, a \to [a]_p$ is bijection. Also $\mathrm{id}_S : S \to S, s \to s$ is a bijection. It follows that Φ is a bijection see Exercise 1.3#2.

Next we define an addition $\oplus$ and a multiplication $\odot$ on R by

$$(1) \qquad r \oplus s = \Phi^{-1}(\Phi(r) + \Phi(s)) \quad \text{and} \quad r \odot s := \Phi^{-1}(\Phi(r)\Phi(s))$$

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 $r, s \in R$. Hence 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 3.6.3 $F[x]/(p)$ 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. □

Notation 3.6.5. *Let* R *and* S *be commutative rings with identities. Suppose that* S *is a subring of* R *and* $1_S = 1_R$. *Let* $f \in S[x]$ *and* $r \in R$. *We identify the polynomial*

$$f = \sum_{i=0}^{n} f_i x^i \quad \textit{in } S[x]$$

with the polynomial

$$g = \sum_{i=0}^{n} f_i x^i \quad \text{in } R[x]$$

Note that with this identification, $S[x]$ becomes a subring of $R[x]$.

Define

$$f(r) := \sum_{i=0}^{n} f_i s^i.$$

Note that $f(r) = g^*(r) = \iota_s(f)$ where ι is the ring homomorphism $\iota : S \to R, s \to s$.

Notation 3.6.6. *Let F be a field and p a non-constant polynomial in $F[x]$. Let R and α be as in 3.6.4. 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 3.6.7. *Let F be a field and p a non-constant polynomial in $F[x]$. Let α and Φ be as in 3.6.4.*

(a) *Let $f \in F[x]$. Then $\Phi\big(f(\alpha)\big) = [f]_p$.*

(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 $p | f$ in $F[x]$.*

(f) *α is a root of p in $F_p[\alpha]$.*

Proof. (a)

$$\begin{aligned}
\Phi\big(f(\alpha)\big) &= \Phi\left(\sum_{i=0}^{\deg f} f_i \alpha^i\right) && \text{– Definition of } f(\alpha) \\
&= \sum_{i=0}^{\deg f} \Phi(f_i)\Phi(\alpha)^i && \text{– } \Phi \text{ is a homomorphism} \\
&= \sum_{i=0}^{\deg f} [f_i]_p [x]_p^i && \text{– 3.6.4} \\
&= \left[\sum_{i=0}^{\deg f} f_i x^i\right]_p = [f]_p. && \text{– } f \to [f]_p \text{ is a homomorphism by 3.6.3}
\end{aligned}$$

(b)

$$\begin{aligned} & f(\alpha) = g(\alpha) \\ \iff\quad & \Phi(f(\alpha)) = \Phi(g(\alpha)) && - \Phi \text{ is injective} \\ \iff\quad & [f]_p = [g]_p && - (a) \end{aligned}$$

(c) Let $\beta \in F_p[\alpha]$ and $f \in F[x]$. Then

$$\begin{aligned} & f(\alpha) = \beta \\ \iff\quad & \Phi(f(\alpha)) = \Phi(\beta) && - \Phi \text{ is injective} \\ \iff\quad & [f]_p = \Phi(\beta) && - (a) \end{aligned}$$

Since $\Phi(\beta) \in F[x]/(p)$, 3.5.4 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 $b_0, \dots b_{n-1} \in F$ 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)

$$\begin{aligned} & f(\alpha) = 0_F \\ \iff\quad & f(\alpha) = 0_F(\alpha) && - \text{definition of } 0_F(\alpha) \\ \iff\quad & [f]_p = [0_F] && -(b) \\ \iff\quad & p \mid f - 0_F && - 3.5.3 \\ \iff\quad & p \mid f && - 2.2.9(\text{b}) \end{aligned}$$

(f) Note that $p|p$ and so $p(\alpha) = 0_F$ by (e). Thus α is a root of p in $F_p[\alpha]$. □

Example 3.6.8. Let $p = x^2 + x + 1 \in \mathbb{Z}_2[x]$. Determine the addition and multiplication table of $\mathbb{Z}_{2,p}[\alpha]$.

By 3.6.7(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 2.5.6 $\mathbb{Z}_2 = \{0, 1\}$ and so

$$\mathbb{Z}_{2,p}[\alpha] = \{0 + 0\alpha,\ 0 + 1\alpha,\ 1 + 0\alpha,\ 1 + 1\alpha\} = \{0,\ 1,\ \alpha, 1 + \alpha\}.$$

Note that $\alpha + \alpha = 2\alpha = 0\alpha = 0$ and so we get

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

Since $\alpha+\alpha=0$ we have $-\alpha=\alpha$. By 3.6.7(f) $p(\alpha)=0$. Hence $1+\alpha+\alpha^2=0$ and thus

$$\alpha^2=-1-\alpha=1+\alpha.$$

$\cdot$	0	1	α	$1+\alpha$
0	0	0	0	0
1	0	1	α	$1+\alpha$
α	0	α	$1+\alpha$	1
$\alpha+1$	0	$1+\alpha$	1	α

Note here that by the Distributive Law, Column '$1+\alpha$' is the sum of Column '1' and Column 'α'. Also Row '$1+\alpha$' is the sum of Row '1' and Row 'α'

Exercises 3.6:

#1. Let $p=x^3+x^2+1\in\mathbb{Z}_2[x]$. Determine the addition and multiplication table of $\mathbb{Z}_{2,p}[\alpha]$. Is $\mathbb{Z}_{2,p}[\alpha]$ a field?

#2. Let $p=x^2-3\in\mathbb{Q}[x]$. Each element of $\mathbb{Q}_p[\alpha]$ can be uniquely written in the form $b+c\alpha$, with $b,c\in\mathbb{Q}$ (Why?). Determine the rules of addition and multiplication in $\mathbb{Q}_p[\alpha]$. In other words, for $b,c,d,e\in\mathbb{Q}$ find $r,s,u,v\in\mathbb{Q}$ with

$$(b+c\alpha)+(d+e\alpha)=r+s\alpha \qquad \text{and} \qquad (b+c\alpha)(d+e\alpha)=u+v\alpha.$$

3.7 $F_p[\alpha]$ when p is irreducible

In this section we determine when $F_p[\alpha]$ is a field.

Theorem 3.7.1. *Let F be a field, p and non-constant polynomial in $F[x]$ and f any polynomial 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

	$f(\alpha)$ is a unit in $F_p[\alpha]$	
$\iff$	$f(\alpha)\beta = 1_F$ for some $\beta \in F_p[\alpha]$	$-F_p[\alpha]$ is commutative, 2.12.9
$\iff$	$f(\alpha)g(\alpha) = 1_F$ for some $g \in F[x]$	–By 3.6.7(c) $\beta = g(\alpha)$ for some $g \in F[x]$
$\iff$	$(fg)(\alpha) = 1_F(\alpha)$ for some $g \in F[x]$	– 3.4.8
$\iff$	$[fg]_p = [1_F]_p$ for some $g \in F[x]$	– 3.6.7(b)
$\iff$	$1_F = fg + ph$ for some $g, h \in F[x]$	– 3.5.3(a)(i)
$\iff$	$\gcd(f, p) = 1_F$	– 3.2.14

(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 $f(\alpha)$. □

Theorem 3.7.2. *Let F be a field and p a non-constant polynomial in $F[x]$. 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): Suppose p is irreducible. By 3.6.4(c) $F_p[\alpha]$ is a commutative ring with additive identity 0_F and multiplicative identity 1_F. Since F is a field, $1_F \neq 0_F$. Thus it remains to show that every non-zero element in $F_p[\alpha]$ is a unit. So let $\beta \in F_p[\alpha]$ with $\beta \neq 0_F$. By 3.6.7(c), $\beta = f(\alpha)$ for some $f \in F[x]$. Then $f(\alpha) \neq 0_F$ and 3.6.7(e), gives $p \nmid f$. Since p is irreducible, Exercise 3.3#4 shows that $\gcd(f, p) = 1_F$. Hence by Theorem 3.7.1 $\beta = f(\alpha)$ is a unit in $F_p[\alpha]$.

(b) $\Longrightarrow$ (c): If $F_p[\alpha]$ is a field, then by Theorem 2.8.10 $F_p[\alpha]$ is an integral domain.

(c) $\Longrightarrow$ (a): Suppose $F_p[\alpha]$ is an integral domain and let $g, h \in F[x]$ with $p|gh$. We will show that $p|g$ or $p|h$. By 3.6.7(e) α is a root of p and so $p(\alpha) = 0_F$. Since $p|gh$ we conlude from 3.4.12(a) that α is a root of gh. Hence

$$0_F = (gh)(\alpha) \overset{3.4.8}{=} g(\alpha)h(\alpha).$$

Since **(Ax 11)** holds in integral domains this gives $g(\alpha) = 0_F$ or $h(\alpha) = 0_F$. By 3.6.7(f) this implies that $p|g$ or $p|h$.

We proved that $p|gh$ implies $p|g$ or $p|h$. Thus 3.3.5 shows that p is irreducible. □

Theorem 3.7.3. *Let F be a field and p an irreducible polynomial in $F[x]$. Then F is a subring of $F_p[\alpha]$, $F_p[\alpha]$ is a field and α is a root of p in $F_p[\alpha]$.*

Proof. By 3.6.4 F is a subring of $F_p[\alpha]$. Since p is irreducible, 3.7.2 implies that $F_p[\alpha]$ is field. By 3.6.7 α is a root of p in $F_p[\alpha]$. □

Example 3.7.4. Put $K := \mathbb{R}_{x^2+1}[\alpha]$. Determine the addition and multiplication in K and show that K is a field.

By 3.6.7(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 3.6.7, 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

$$(a+b\alpha)(c+d\alpha) = ac+(bc+ad)\alpha+bd\alpha^2 = ac+(bc+ad)\alpha+bd(-1) = (ac-bd)+(ad+bc)\alpha.$$

Note that x^2+1 has no roots in $\mathbb{R}$ and so by 3.4.16 x^2+1 is irreducible. Hence 3.7.2 shows that 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 $\mathbb{C}$.

Theorem 3.7.5. *Let F be a field and $f \in F[x]$.*

(a) *Suppose f is not constant. Then there exists a field K such that F is a subring of K and f has a root in K.*

(b) *There exist a field L $n \in \mathbb{N}$, and elements $c, a_1, a_2 \dots, a_n$ in L such that F is a subring of L and*

$$f = c \cdot (x-a_1) \cdot (x-a_2) \cdot \ldots \cdot (x-a_n)$$

Proof. (a) By 3.3.9, f is a product of irreducible polynomials. In particular, there exists an irreducible polynomial p in $F[x]$ dividing f. By 3.7.3 $K = F_p[\alpha]$ is a field containing F and α is a root of p in K. Since $p|f$, 3.4.12 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 that (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 $\deg f \geq 1$. So f is not constant and by (a) there exists a field K with F as a subring and a root a of f in K. By the Factor Theorem 3.4.11 $x-a$ divides f in $K[x]$ and so $f = (x-a)\cdot g = g \cdot (x-a)$ for some $g \in K[x]$. Then $k+1 = \deg f = \deg g + \deg(x-a) = \deg g + 1$. So $\deg g = k$ and by the induction assumption there exists a field L and elements $c, a_1, \dots a_k$ in L such that K is a subring of L and

$$g = c \cdot (x-a_1) \cdot \ldots \cdot (x-a_k).$$

Put $a_{k+1} = a$. Then

$$f = g \cdot (x - a) = c \cdot (x - a_1) \cdot \ldots \cdot (x - a_k) \cdot (x - a_{k+1}).$$

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 3.7:

#1. In each part explain why $t \in F_p[\alpha]$ is a unit and find its inverse.

(a) $t = -3 + 2\alpha$, $F = \mathbb{Q}$, $p = x^2 - 2$

(b) $t = 1 + \alpha + \alpha^2$, $F = \mathbb{Z}_3$, $p = x^2 + 1$

(c) $t = 1 + \alpha + \alpha^2$, $F = \mathbb{Z}_2$, $p = x^3 + x + 1$

#2. Determine whether $F_p[\alpha]$ is a field.

(a) $F = \mathbb{Z}_3$, $p = x^3 + 2x^2 + x + 1$.

(b) $F = \mathbb{Z}_5$, $p = 2x^3 - 4x^2 + 2x + 1$.

(c) $F = \mathbb{Z}_2$, $p = x^4 + x^2 + 1$.

#3. (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 $\mathbb{Q}_{x^2-3}[\alpha]$.

#4. Let $p = x^3 + x^2 + 1 \in \mathbb{Z}_2[x]$.

(a) Determine the addition and multiplication table of $\mathbb{Z}_{2,p}[\alpha]$.

(b) Is $\mathbb{Z}_{2,p}[\alpha]$ a field?

(c) Show that $x^3 + x + 1$ has three distinct roots in $\mathbb{Z}_{2,p}[\alpha]$

Chapter 4

Ideals and Quotients

4.1 Ideals

Definition 4.1.1. *Let I be a subset of the ring R.*

(a) *We say that I* absorbs *R if*

$$ra \in I \quad \text{and} \quad ar \in I \qquad \text{for all } a \in I, r \in R$$

(b) *We say that I is an* ideal *of R if I is a subring of R and I absorbs R.*

Theorem 4.1.2 (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 4.1.1 S absorbs R and S is a subring. Thus (iii) holds and by the Subring Theorem 2.7.2 also (i), (ii) and (iv) hold.

$\Longleftarrow$: Suppose that (i)-(iv) hold. From (iii) we get that $ab \in I$ for all $a, b \in I$. Together with (i), (ii) and (iv) this shows that the four conditions of the Subring Theorem 2.7.2 hold for I. Thus I is a subring of R. By (iii), I absorbs R and so I is an ideal in R. □

Example 4.1.3. (1) Let R be a ring, then both $\{0_R\}$ and R are ideals in R.

(2) $\{3n \mid n \in \mathbb{Z}^+\}$ is an ideal in $\mathbb{Z}$.

(3) $\mathbb{Z}$ is not an ideal in $\mathbb{Q}$.

(4) Let F be a field and $a \in F$. Then $\{f \in F[x] \mid f^*(a) = 0_F\}$ is an ideal in $F[x]$.

(5) 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.

(6) Let R and S be rings. Let I be an ideal in R and J an ideal in S. Then $I \times J$ is an ideal in R. In particular, both $R \times \{0_S\}$ and $\{0_R\} \times S$ are ideals in $R \times S$.

Proof. See Exercise #1 □

Definition 4.1.4. *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 $a \in R$.*

Theorem 4.1.5. *Let R be a commutative ring and $a \in R$. Then aR is an ideal in R. Moreover, if R has an identity, 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. To show that aR is an ideal in R 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

$$\begin{aligned} 0_R &= a0_R \in aR, \\ b + c &= as + at = a(s+t) \in aR, \\ rb &= br = (as)r = a(sr) \in aR \\ -b &= -(as) = a(-s) \in aR. \end{aligned}$$

So by 4.1.2 aR is an ideal in R.

Suppose now that R has an identity. Then $a = a \cdot 1_R$ and so $a \in aR$.

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 $aR \subseteq I$. □

Definition 4.1.6. *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 4.1.7. *Let R be a commutative ring and let $a, b, n \in R$. Then*

$$a \equiv b \pmod{n} \iff a \equiv b \pmod{nR}$$

Proof.

$$\begin{aligned} & a \equiv b \pmod{n} \\ \iff \quad & a - b = nk \quad \text{for some } k \in R && - \ 2.4.9 \\ \iff \quad & a - b \in nR && - \text{Definition of } nR \\ \iff \quad & a \equiv b \pmod{nR} && - \text{Definition of } \equiv \pmod{I} \quad 4.1.10 \end{aligned}$$

□

Theorem 4.1.8. *Let I be an ideal in the ring 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 2.2.9 $a - a = 0_R$ 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 2.2.9 $b - a = -(a - b)$. 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 4.1.9. *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 4.1.10. *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 2.2.8.

(b) $\iff$ (c) and (j) $\iff$ (k): Principal of Substitution.

(c) $\iff$ (d) and (i) $\iff$ (j): This holds by definition of '$\equiv \pmod{I}$'.

By 4.1.8 we know that '$\equiv \pmod{I}$ is an equivalence relation. Also $a + I$ is the equivalence class of a and so Theorem 1.5.5 implies that (d)-(i) are equivalent. □

Theorem 4.1.11. *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 4.1.10(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 4.1.10(f),(g) shows that $a + I = b + I$. So two cosets of I in R are either disjoint or equal. □

Exercises 4.1:

#1. Show that:

(a) Let R be a ring, then both $\{0_R\}$ and R are ideals in R.

(b) $\{3n \mid n \in \mathbb{Z}^+\}$ is an ideal in $\mathbb{Z}$.

(c) $\mathbb{Z}$ is not an ideal in $\mathbb{Q}$.

(d) Let F be a field and $a \in F$. Then $\{f \in F[x] \mid f^*(a) = 0_F\}$ is an ideal in $F[x]$.

(e) 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.

(f) Let R and S be rings. Let I be an ideal in R and J an ideal in S. Then $I \times J$ is an ideal in R. In particular, both $R \times \{0_S\}$ and $\{0_R\} \times S$ are ideals 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. 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 $I = F[x]d$.

#5. 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.

4.2 Quotient Rings

Theorem 4.2.1. *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 $a + I = \tilde{a} + I$ 4.1.10 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 4.1.10 $(a + b) + I = (\tilde{a} + \tilde{b}) + I$.
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 4.1.10. □

Definition 4.2.2. *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 here that these operations are well defined by 4.2.1/

Remark 4.2.3. (a) *Let R be a commutative ring and $n \in R$. Then $R_n = R/nR$.*

(b) *Let F be a field and $p \in F[x]$. Then $F[x]/(p) = F[x]/pF[x]$.*

Proof. (a) By Remark 4.1.7 the relations '$\equiv \pmod n$' and '$\equiv \pmod{nR}$' are the same. So also their sets of equivalence classes R_n and R/nR are the same.

(b) Since $F[x]/(p) = F[x]_p$ this is a special case of (a). □

Theorem 4.2.4. *Let R be a ring and I an ideal in R*

(a) *The function $\pi : R \to R/I$, $a \mapsto a + I$ is a surjective 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 surjective.

(b), (c) and (d): By (a) π is a surjective homomorphism. Thus we can apply E.0.3 and conclude that (b),(c) and (d) hold.

(e): By (a) π is a surjective homomorphism. Thus (e) follows from 2.11.8(a) □

Theorem 4.2.5. *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 4.1.10 $r \in 0_R + I$ if and only if $r + I = 0_R + I$. By 4.2.4(c) $0_R + I = I$ and so (a) and (b) are equivalent.

(b) $\iff$ (c): By 4.2.4(c) $0_{R/I} = I$ and so (b) and (c) are equivalent. □

Definition 4.2.6. (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*

$$\pi: \quad R \to R/I, \quad r \mapsto r + I$$

is called the natural homomorphism *from R to R/I.*

Theorem 4.2.7. *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 4.1.2. Let $r \in R$. By definition of $\ker f$ we have

$$(*) \qquad r \in \ker f \quad \Longleftrightarrow \quad f(r) = 0_S.$$

Let $a, b \in \ker f$. By $(*)$

$$(**) \qquad f(a) = 0_S \quad \text{and} \quad f(b) = 0_S.$$

(i) $f(0_R) \overset{2.11.7(a)}{=} 0_S$ and so $0_R \in \ker f$ by $(*)$.

(ii) $f(a+b) \overset{\text{f hom}}{=} f(a) + f(b) \overset{(**)}{=} 0_S + 0_S \overset{\textbf{Ax 4}}{=} 0_S$ and so $a + b \in \ker f$ by $(*)$.

(iii) $f(ra) \overset{\text{f hom}}{=} f(r)f(a) \overset{(**)}{=} f(r)0_S \overset{2.2.9(c)}{=} 0_S$ and so $ra \in \ker f$ by $(*)$. Similarly, $ar \in \ker f$.

(iv) $f(-a) \overset{2.11.7(b)}{=} -f(a) \overset{(**)}{=} -0_S \overset{2.2.9(a)}{=} 0_S$ and so $-a \in \ker f$ by $(*)$. □

Example 4.2.8. Define

$$\Phi: \quad \mathbb{R}[x] \to \mathbb{C}, \qquad f \mapsto f(i)$$

Verify that Φ is a surjective homomorphism and compute $\ker \Phi$.

Define $\rho : \mathbb{R} \to \mathbb{C}, r \mapsto r$. Then ρ is a homomorphism and Φ is the function ρ_i from Theorem 3.4.1. So Φ is a homomorphism. Alternatively, note that for $f, g \in \mathbb{R}[x]$:

$$\Phi(f+g) = (f+g)(i) = f(i) + g(i) = \Phi(f) + \Phi(g) \text{ and } \Phi(fg) = (fg)(i) = f(i)g(i) = \Phi(f)\Phi(g).$$

To show that f is surjective, let $c \in \mathbb{C}$. Then $c = a + bi$ for some $a, b \in \mathbb{R}$. Thus $\Phi(a+bx) = a+bi = c$ and so Φ is surjective.

To compute $\ker f$ let $f \in \mathbb{R}[x]$. 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

$$\begin{array}{lll}
 & f \in \ker \Phi & \\
\iff & \Phi(f) = 0 & \text{– definition of } \ker \Phi \\
\iff & f(i) = 0 & \text{– definition of } \Phi \\
\iff & a + bi = 0 & \text{– } (*) \\
\iff & a = 0 \text{ and } b = 0 & \text{– Property of } \mathbb{C} \\
\iff & a + bx = 0 & \text{– definition of polynomial ring} \\
\iff & r = 0 & \text{– } r = a + bx \\
\iff & x^2 + 1 \mid f & \text{– 3.2.1} \\
\iff & f = (x^2+1) \cdot q \text{ for some } q \in \mathbb{R}[x] & \text{– Definition of `divide'} \\
\iff & f \in (x^2+1)\mathbb{R}[x] & \text{– Definition of } (x^2+1)\mathbb{R}[x]
\end{array}$$

Thus $\ker \Phi = (x^2+1)\mathbb{R}[x]$.

Theorem 4.2.9. *Let R be a ring.*

(a) *Let I an ideal in R and*

$$\pi : \quad R \to R/I, \quad a \mapsto a + I.$$

the natural homomorphism from R to I. Then $\ker \pi = I$.

(b) *Let I be subset of R. Then I is an ideal in R if and only if $I = \ker f$ for some homomorphism of rings $f : R \to S$.*

Proof. (a): Let $r \in R$. Then

$$\begin{array}{lcl}
 & r \in \ker \pi & \\
\iff & \pi(r) = 0_{R/I} & \text{– definition of } \ker \pi \\
\iff & r + I = 0_{R/I} & \text{– definition of } \pi \\
\iff & r \in I & \text{– 4.2.5}
\end{array}$$

Thus $\ker \pi = I$.

(b) The forward direction follows from (a) and the backwards direction from 4.2.7. □

Theorem 4.2.10. *Let $f : R \to S$ be a ring homomorphism.*

(a) *Let $a, b \in R$. Then*

$$\begin{array}{lc}
 & f(a) = f(b) \\
\iff & a - b \in \ker f \\
\iff & a + \ker f = b + \ker f
\end{array}$$

(b) *f is injective if and only if* $\ker f = \{0_R\}$.

Proof. (a)

$$\begin{array}{lrcll} & f(a) & = & f(b) & \\ \iff & f(a)-f(b) & = & 0_S & -\ 2.2.9(\text{f}) \\ \iff & f(a-b) & = & 0_S & -\ 2.11.7(\text{c}) \\ \iff & & a-b\in\ker f & & -\ \text{Definition of } \ker f \\ \iff & a+\ker f & = & b+\ker f & -\ 4.1.10 \end{array}$$

(b) $\Longrightarrow$: Suppose f is injective and let $a \in R$. Then

$$\begin{array}{lcl} & a\in\ker f & \\ \iff & f(a)=0_S & -\ \text{Definition of } \ker f \\ \iff & f(a)=f(0_R) & -\ 2.11.7(\text{a}) \\ \iff & a=0_R & -\ f \text{ is injective} \end{array}$$

Thus $\ker f = \{0_R\}$.

$\Longleftarrow$: Suppose $\ker f = \{0_R\}$ and let $a, b \in R$ with $f(a) = f(b)$. Then by (a) $a - b \in \ker f$. As $\ker f = \{0_R\}$ this gives $a - b = 0_R$, so $a = b$ by 2.2.9(f). Hence f is injective. □

Theorem 4.2.11 (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*

$$\overline{f}: \quad R/\ker f \mapsto \operatorname{Im} f, \quad a+\ker f \ \mapsto\ f(a)$$

is a well-defined ring isomorphism. In particular $R/\ker f$ and $\operatorname{Im} f$ are isomorphic rings

Proof. Let $a, b \in R$. By 4.2.10 $f(a) = f(b)$ if and only if $a + \ker f = b + \ker f$. The forward direction shows that $\overline{f}$ is injective 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 surjective.

It remains to verify that $\overline{f}$ is a ring homomorphism. We compute

$$\begin{array}{rclcl} \overline{f}\Big((a+\ker f)+(b+\ker f)\Big) & \overset{\text{Def }+}{=} & \overline{f}\Big((a+b)+\ker f\Big) & \overset{\text{Def }\overline{f}}{=} & f(a+b) \\ & \overset{f \text{ hom}}{=} & f(a)+f(b) & \overset{\text{Def }\overline{f}}{=} & \overline{f}(a+\ker f)+\overline{f}(b+\ker f) \end{array}$$

and

$$\begin{array}{rclcl} \overline{f}\Big((a+\ker f)\cdot(b+\ker f)\Big) & \overset{\text{Def }\cdot}{=} & \overline{f}\Big(ab+\ker f\Big) & \overset{\text{Def }\overline{f}}{=} & f(ab) \\ & \overset{f \text{ hom}}{=} & f(a)\cdot f(b) & \overset{\text{Def }\overline{f}}{=} & \overline{f}(a+\ker f)\cdot\overline{f}(b+\ker f) \end{array}$$

and so $\overline{f}$ is a homomorphism. □

Example 4.2.12. Let n and m be non-zero integers with $\gcd(n,m) = 1$. Apply the isomorphism theorem to the homomorphism

$$f : \mathbb{Z} \to \mathbb{Z}_n \times \mathbb{Z}_m, \quad a \mapsto \big([a]_n, [b]_m \big)$$

We first compute $\ker f$

$$\begin{array}{lcl}
 & a \in \ker f & \\
\iff & f(a) = 0_{\mathbb{Z}_n \times \mathbb{Z}_m} & \text{– definition of } \ker \pi \\
\iff & f(a) = ([0]_n, [0]_m) & \text{– 2.1.7(b), 2.6.4(b)} \\
\iff & \big([a]_n, [b]_m\big) = ([0]_n, [0]_m) & \text{– definition of } f \\
\iff & [a]_n = [0]_n \quad \text{and} \quad [b]_m = [0]_m & \text{– 1.3.2} \\
\iff & n \,|\, a - 0 \quad \text{and} \quad m \,|\, a - 0 & \text{– 2.4.9} \\
\iff & n \,|\, a \quad \text{and} \quad m \,|\, a & \text{– 2.2.9}(b) \\
\iff & nm \,|\, a & \text{– } \gcd(n,m) = 1, \text{Exercise 2.9\#2} \\
\iff & a = nmk \quad \text{for some} \quad k \in \mathbb{Z} & \text{– definition of 'divide'} \\
\iff & a \in nm\mathbb{Z} & \text{– definition of } nm\mathbb{Z}
\end{array}$$

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 4.2.3(a).

By the First Isomorphism Theorem $\mathbb{Z}/\ker f$ is isomorphic to $\operatorname{Im} f$ and so

$$(*) \qquad \mathbb{Z}_{nm} \quad \text{is isomorphic to} \quad \operatorname{Im} f.$$

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 $\operatorname{Im} f = \mathbb{Z}_n \times \mathbb{Z}_m$. Hence $(*)$ implies

$$\mathbb{Z}_{nm} \quad \text{is isomorphic to} \quad \mathbb{Z}_n \times \mathbb{Z}_m.$$

Appendix A

Logic

A.1 Rules of Logic

In the following we collect a few statements which are always true.

Theorem 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) $(P \iff Q) \implies (P \implies Q)$.

(LR 21) $(P \implies Q) \iff (\text{not -}Q \implies \text{not -}P)$

(LR 22) $(P \iff Q) \iff (\text{not -}P \iff \text{not -}Q)$.

(LR 23) $\text{not -}(P \text{ and } Q) \iff (\text{not -}P \text{ or not -}Q)$

(LR 24) $\text{not -}(P \text{ or } Q) \iff (\text{not -}P \text{ and not -}Q)$

(LR 25) $\big((P \text{ and } Q) \text{ and } R\big) \iff \big(P \text{ and } (Q \text{ and } R)\big)$.

(LR 26) $\big((P \text{ or } Q) \text{ or } R\big) \iff \big(P \text{ or } (Q \text{ or } R)\big)$.

(LR 27) $\big((P \text{ and } Q) \text{ or } R\big) \iff \big((P \text{ or } R) \text{ and } (Q \text{ or } R)\big)$.

(LR 28) $\big((P \text{ or } Q) \text{ and } R\big) \iff \big((P \text{ and } R) \text{ or } (Q \text{ and } R)\big)$.

(LR 29) $\big((P \implies Q) \text{ and } (Q \implies R)\big) \implies (P \implies R)$

(LR 30) $\big((P \iff Q) \text{ and } (Q \iff R)\big) \implies (P \iff R)$

Proof. If any of these statements are not evident to you, you should use a truth table to verify it. □

Theorem A.1.2. *Let $P(x)$ be a 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

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

holds. By definition of "There exists:" we conclude that there exists an object a such that

$$(*) \qquad P(a) \text{ is true}$$

q Also by definition of "There exists at most one":

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

From $(**)$ and the definition of "for all x :" we get

$$(***) \qquad \text{for all } y : \qquad P(a) \text{ and } P(y) \qquad \Longrightarrow \qquad a = y$$

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

$$\text{for all } y : \qquad P(y) \qquad \iff \qquad P(a) \text{ and } P(y).$$

By A.1.1(LR 20) $P \equiv Q$ implies $P \Longrightarrow Q$ and so we conclude that

$$(+) \qquad \text{for all } y : \qquad P(y) \qquad \Longrightarrow \qquad P(a) \text{ and } P(y)$$

By A.1.1(LR 29)

$$\big((P \Longrightarrow Q) \text{ and} \,(Q \Longrightarrow T)\big) \quad \Longrightarrow \quad \big(P \Longrightarrow Q\big)$$

Together with $(+)$ and $(***)$ this gives

$$(++) \qquad \text{for all } y : \qquad P(y) \qquad \Longrightarrow \qquad a = y.$$

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

$$(+++) \qquad \text{for all } y : \qquad a = y \qquad \Longrightarrow \qquad P(y)$$

By A.1.1(LR 20) $P \equiv Q$ if and only if $P \Longrightarrow Q$ and $Q \Longrightarrow P$. Together with $(++)$ and $(+++)$ we get

$$\text{for all } y : \qquad P(y) \qquad \iff \qquad a = y.$$

Thus by definition of 'there exists x:' this gives

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

Hence the definition of "There exists a unique" gives

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

$\Longleftarrow$: Suppose next that

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

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

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

and so there exists an object a such that

(#) $$\text{for all } y : P(y) \qquad \Longleftrightarrow \qquad a = y.$$

In particular, by definition of "for all y":

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

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

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

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

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

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

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

From (##) and (###) we have

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

□

Theorem A.1.3. *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* $\Big\{\Phi(s) \;\Big|\; s \in S \text{ and } P(s)\Big\}$ *such that*

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

Proof. Define

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

Then

$$\begin{array}{rcl}
& t \in \Big\{\Phi(s) \mid s \in S \text{ and } P(s)\Big\} & \\
\Longleftrightarrow & t \in \Big\{\Phi(s) \mid s \in \{r \in S \mid \Phi(r)\}\Big\} & \text{By } (*) \\
\Longleftrightarrow & \text{there exists } s \in \{r \in S \mid P(r)\} \text{ with } t = \Phi(s) & 1.2.9 \\
\Longleftrightarrow & \text{there exists } s \text{ with } \Big(s \in \{r \in S \mid P(r)\} \text{ and } t = \Phi(s)\Big) & \text{definition of `there exists } s \in\text{' see 1.2.7} \\
\Longleftrightarrow & \text{there exists } s \text{ with } \Big(\big(s \in S \text{ and } P(s)\big) \text{ and } t = \Phi(s)\Big) & 1.2.5 \\
\Longleftrightarrow & \text{there exists } s \text{ with } \Big(s \in S \text{ and } \big(P(s) \text{ and } t = \Phi(s)\big)\Big) & \text{Rule of Logic: A.1.1(LR 25) :} \\
& & \Big(P \text{ and } (Q \text{ and } R)\Big) \Longleftrightarrow \Big((P \text{ and } Q) \text{ and } R\Big) \\
\Longleftrightarrow & \text{there exists } s \in S \text{ with } \Big(P(s) \text{ and } t = \Phi(s)\Big) & \text{definition of `there exists } s \in\text{' see 1.2.7}
\end{array}$$

□

Appendix B

Relations, Functions and Partitions

B.1 Equality of functions

Theorem B.1.1. *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 1.3.2(b) :

$(*)$ $\quad$ *$f = g$ 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 $b \in B$.

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

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 1.2.1. ☐

B.2 The inverse of a function

Definition B.2.1. *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 $f \circ g = \mathrm{id}_B$.*

Theorem B.2.2. *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 $a \in A, b \in B$:

(2)) $$f(a) = b \Longrightarrow a = g(b)$$

Let $a \in A$ and put

(3) $$b = f(a)$$

Then by (2)

(4) $$a = g(b)$$

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 1.3.14 $g \circ f = \mathrm{id}_A$. Hence g is a left inverse of f. □

Theorem B.2.3. *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 $b \in B$:*

$$fa = b \iff a = gb$$

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 theorem follows from B.2.2 □

Theorem B.2.4. *Let $f : A \to B$ be a function and suppose $A \neq \varnothing$.*

(a) *f is injective if and only if f has a right inverse.*

(b) *f is surjective if and only if f has left inverse.*

(c) *f is a injective 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 injective. 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 injective, we conclude that $a_b = b$ and so $ga = a_b = b$. Thus by B.2.2, g is right inverse of f.

(b) Suppose f is surjective. Let $a \in A$ and $b \in B$ with $gb = a$. Then $a = a_b$. Since f is surjective, $B = \operatorname{Im} f$ and so $a \in \operatorname{Im} f$ and $f(a_b) = b$. Hence $fa = b$ and so by B.2.2 (with the roles of f and f interchanged), g is left inverse of f.

(c) Suppose f is a injective correspondence. Then f is injective and surjective 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.2.2 $g(fa) = a$ and $g(fb) = b$. So $a = b$ and f -s injective.

(b) Suppose g is a right inverse of f and let $b \in B$. Then by B.2.2, $f(gb) = b$ and so f is surjective.

(c) Suppose f has an inverse. Then f has a left and a right inverse and so by (a) and (b), f is injective and surjective. So f is a injective correspondence. □

B.3 Partitions

Definition B.3.1. *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) $\sim_\Delta = \left(A, A, \left\{(a,b) \in A \times A \mid \{a,b\} \subseteq D \text{ for some } D \in \Delta\right\}\right)$.

Example B.3.2. The relation corresponding to a partition $\Delta = \left\{\{1,3\},\{2\}\right\}$ 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.

Theorem B.3.3. *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 $a \in [b]$. 1.5.5 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 1.5.5 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. □

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*

$$a + b \in \mathbb{R} \qquad \textit{(}\text{Closure of addition}\textit{)}$$

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*

$$a \cdot b \in \mathbb{R} \qquad \textit{(}\text{Closure of multiplication}\textit{)}$$

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

Theorem 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

Theorem 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

Theorem 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 $a_1, a_2, \dots a_n \in G$. 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 $\sum_{i=1}^{n} a_i = \Bigg(\bigg(\dots \big((a_1 + a_2) + a_3 \big) + \dots + a_{n-2} \bigg) + a_{n-1} \Bigg) + a_n.$

Inductively, we say that z is a sum of $(a_1, \dots, a_n)$ 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 $(a_1, \dots, a_k)$, y is a sum of $(a_{k+1}, a_{k+2}, \dots, a_n)$ 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 $(a_1, a_2, \dots, a_n)$ 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 $a_1, a_2, \dots a_n$ are elements of G and z is a sum of $(a_1, a_2, \dots, a_n)$, 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 $1 \le k < n$. Let $a_1, a_2, \dots a_n$ be elements of G and z is a sum of $(a_1, a_2, \dots, a_n)$. 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 $(a_1, \dots, a_{n-1})$, 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 $1 \le k < n$, a sum x of $(a_1, \dots, a_k)$ and a sum y of $(a_{k+1}, \dots, a_n)$ such that $z = x + y$.

Case 1: $k = n - 1$.

In this case x is a sum of $(a_1, \dots, a_{n-1})$ 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: $1 \le k < n - 1$.

Observe that $n - k \le n - 1 < n$ and so by the induction assumption $P(n-k)$ holds. Since y is a sum of $a_{k+1}, \dots, a_n)$ 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 $(a_1, \dots, a_k)$ and $\sum_{i=1}^{n-k-1} a_{k+i}$ is a sum of $(a_k, \dots, a_{n-1})$. So by definition of a sum, u is a sum of $(a_1, \dots, a_{n-1})$. 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 $a_1, a_2, \dots, a_n \in G$ and $f : [1 \dots n] \to [1 \dots n]$ 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 surjective there exists a unique integer k with $f(k) = n$.

Define $g : \{1, \dots n-1\} \to \{1, \dots, n-1\}$ 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 injective, $m \neq k$. If $m < k$, then $g(m) = f(m) = l$ and if $m > k$, then $g(m-1) = f(m) = l$. Thus g is surjective and by G.1.7(b) g is also injective. 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

$$\begin{array}{rcl}
 & & \sum_{i=1}^{n} a_{f(i)} \\
\text{(GAL)} & = & \left(\sum_{i=1}^{k-1} a_{f(i)}\right) + \left(a_{f(k)} + \sum_{i=k+1}^{n} a_{f(i)}\right) \\
(n = f(k)) & = & \left(\sum_{i=1}^{k-1} a_{f(i)}\right) + \left(a_n + \sum_{i=k+1}^{n} a_{f(i)}\right) \\
(`+' \text{ commutative }) & = & \left(\sum_{i=1}^{k-1} a_{f(i)}\right) + \left(\sum_{i=k+1}^{n} a_{f(i)} + a_n\right) \\
('+'\text{associative }) & = & \left(\left(\sum_{i=1}^{k-1} a_{f(i)}\right) + \left(\sum_{i=k+1}^{n} a_{f(i)}\right)\right) + a_n \\
(\text{Substitution } j = i+1) & = & \left(\left(\sum_{i=1}^{k-1} a_{f(i)}\right) + \left(\sum_{j=k}^{n-1} a_{f(j+1)}\right)\right) + a_n \\
(\text{definition of } g) & = & \left(\left(\sum_{i=1}^{k-1} a_{g(i)}\right) + \left(\sum_{j=k}^{n-1} a_{g(j)}\right)\right) + a_n \\
\text{(GAL)} & = & \left(\sum_{i=1}^{n-1} a_{g(i)}\right) + a_n \\
(*) & = & \left(\sum_{i=1}^{n-1} a_i\right) + a_n \\
(\text{definition of } \sum) & = & \sum_{i=1}^{n} a_i
\end{array}$$

So the Theorem holds for n and thus by the Principal of Mathematical induction for all positive integers. □

Theorem 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 $g, h : \{1, \dots, n\} \to I$ 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 $f := \{1, \dots, n\}$ 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 $f_i : \{1, \dots, n_i\} \to I_{j_i}$ be a bijection and define $f : \{1 \dots, n_1 + n_2\} \to I$ by $f(i) = f_1(i)$ if $1 \le i \le n_1$ and $f(i) = f_2(i - n_1)$ if $n_1 + 1 \le i \le n_1 + n_2$. Then clearly f is a surjective and so by G.1.7(b), f is injective. We compute

$$\begin{aligned}
\textstyle\sum_{i\in I} a_i &= \textstyle\sum_{i=1}^{n_1+n_2} a_{f(i)} \\
&\overset{\text{GAL}}{=} \left(\textstyle\sum_{i=1}^{n_1} a_{f(i)}\right) + \left(\textstyle\sum_{i=n_1+1}^{n_1+n_2} a_{f(i)}\right) \\
&= \left(\textstyle\sum_{i=1}^{n_1} a_{f_1(i)}\right) + \left(\textstyle\sum_{i=1}^{n_2} a_{f_2(i)}\right) \\
&= \left(\textstyle\sum_{i\in I_{j_1}} a_i\right) + \left(\textstyle\sum_{i\in I_{j_2}} a_i\right) \\
&= \textstyle\sum_{j\in J}\left(\textstyle\sum_{i\in I_j} a_i\right)
\end{aligned}$$

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 $Y = I \smallsetminus J_j$. 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

$$\begin{aligned}
\textstyle\sum_{i\in I} a_i &= \left(\textstyle\sum_{i\in I_j} a_i\right) + \left(\textstyle\sum_{i\in Y} a_i\right) \\
&= \left(\textstyle\sum_{i\in I_j} a_i\right) + \left(\textstyle\sum_{j\neq k\in J}\left(\textstyle\sum_{i\in I_k} a_i\right)\right) \\
&= \textstyle\sum_{j\in J}\left(\textstyle\sum_{i\in I_J} a_i\right)
\end{aligned}$$

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)

$$\left(\sum_{i=1}^{m} a_i \right) \cdot \left(\sum_{i=1}^{k} b_i \right) \overset{(b)}{=} \sum_{i=1}^{n} \left(a_i \sum_{j=1}^{m} b_j \right) \overset{(a)}{=} \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_i b_j \right)$$

□

Appendix E

Verifying Ring Axioms

Theorem E.0.3. *Let* $(R, +, \cdot)$ *be ring and* $(S, \oplus, \odot)$ *a set with binary operations* $\oplus$ *and* $\odot$*. Suppose there exists an surjective homomorphism* $\Phi : R \to S$ *(that is an surjective 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 surjective 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

$$\begin{aligned} a \odot (b \odot c) &= \Phi(x) \odot (\Phi(y) \odot \Phi(z)) = \Phi(x) \odot \Phi(yz) = \Phi(x(yz)) \\ &= \Phi((xy)z)) = \Phi(xy) \odot \Phi(z) = (\Phi(x) \odot \Phi(y)) \odot \Phi(z) = (a \odot b) \odot c \end{aligned}$$

Ax 8

$$\begin{aligned} a \odot (b \oplus c) &= \Phi(x) \odot (\Phi(y) \oplus \Phi(z)) = \Phi(x) \odot \Phi(y+z) = \Phi(x(y+z)) \\ &= \Phi(xy+xz) = \Phi(xy) + \Phi(xz) = (\Phi(x) \odot \Phi(y)) + (\Phi(x) \odot \Phi(z)) = (a \odot b) \oplus (a \odot c) \end{aligned}$$

Similarly $(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c)$.
(b) Suppose R is commutative then
2.1.2 $a \odot b = \Phi(x) \odot \Phi(y) = \Phi(xy) = \Phi(yx) = \Phi(y) \odot \Phi(x) = b \odot a$ □

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* $rs = (r_i s_i)$.

Theorem 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

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{mp} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+\ldots+a_{1n}b_{n1} & a_{11}b_{12}+a_{12}b_{22}+\ldots+a_{1n}b_{n2} & \dots & a_{11}b_{1p}+a_{12}b_{2p}+\ldots+a_{1n}b_{np} \\ a_{21}b_{11}+a_{22}b_{21}+\ldots+a_{2n}b_{n1} & a_{21}b_{12}+a_{22}b_{22}+\ldots+a_{2n}b_{n2} & \dots & a_{21}b_{1p}+a_{22}b_{2p}+\ldots+a_{2n}b_{np} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1}b_{11}+a_{m2}b_{21}+\ldots+a_{mn}b_{n1} & a_{m1}b_{12}+a_{m2}b_{22}+\ldots+a_{mn}b_{n2} & \dots & a_{m1}b_{1p}+a_{m2}b_{2p}+\ldots+a_{mn}b_{np} \end{bmatrix}$$

Theorem 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*

$$\delta_{ij} = \begin{cases} 1_R & \text{if } i = j \\ 0_R & \text{if } i \neq j \end{cases}$$

Proof. Put $J = \{1, \dots, n\} \times \{1, \dots, m\}$ 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 $A(BC) = (AB)C$.

$$(A + B)C = [a_{ij} + b_{ij}]_{ij} \cdot [c_{jk}]_{ik} = \left[\sum_{j=1}^{n} (a_{ij} + b_{ij}) c_{jk}\right]_{ik}$$

$$= \left[\sum_{j=1}^{n} a_{ij} c_{jk}\right]_{ik} + \left[\sum_{j=1}^{n} b_{ij} c_{jk}\right]_{ik} = AC + BC.$$

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. □

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*

(1) $$f(i) = 0_R \text{ for all } i > m$$

We define an addition and multiplication on P by

(2) $$(f+g)(i) = f(i) + g(i) \quad \text{and} \quad (fg)(i) = \sum_{k=0}^{i} f(i)g(k-i)$$

(a) *P is a ring.*

(b) *For $r \in R$ define $r^\circ \in P$ by*

(3) $$r^\circ(i) := \begin{cases} r & \text{if } i = 0 \\ 0_R & \text{if } i \neq 0 \end{cases}$$

Then the map $R \to P, r \mapsto r^\circ$ is a injective 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

$$(r+s)^\circ(i) = \begin{cases} r+s & \text{if } i=0 \\ 0_R & \text{if } i \neq 0 \end{cases} = r^\circ(i) + s^\circ(i) = (r^\circ + s^\circ)(i) \tag{4}$$

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 $\rho : R \to P, r \mapsto r^\circ$. If $r, s \in R$ with $r^\circ = s^\circ$, then $r = r^\circ(1) = s^\circ(1) = s$ and so ρ is injective. 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

$$(f r^\circ)(i) = f(i) \cdot r \tag{8}$$

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

(10) $$(r^\circ\delta_k)(i) = (\delta_k r^\circ)(i) = \begin{cases} r & \text{if } i = k \\ 0_R & \text{if } i \neq k \end{cases}$$

Let $m \in \mathbb{N}$ and $a_0, \dots a_m \in R$. Then (10) implies

(11) $$\left(\sum_{k=0}^{m} a_k^\circ \delta\right)(i) = \begin{cases} a_i & \text{if } i \leq m \\ 0_R & \text{if } i > m \end{cases}$$

We conclude that if $f \in P$ and $a_0, a_1, a_2, \dots a_m \in R$ then

(12) $$f = \sum_{k=0}^{m} a_k^\circ \delta_k \iff m \geq \deg f \text{ and } a_k = f(k) \text{ for all } 0 \leq k \leq m$$

We compute

(13) $$(\delta_k\delta_l)(i) = \sum_{j=0}^{i} \delta_k(j)\delta_l(i-j)$$

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

(14) $$(\delta_k\delta_l)(i) = \begin{cases} 1_R & \text{if } i = k + l \\ 0_R & \text{if } i \neq k + l \end{cases} = \delta_{k+l}(i)$$

and so

(15) $$\delta_k\delta_l = \delta_{k+l}$$

Note that $x = \delta_1$. We conclude that

(16) $$x^k = \delta_k$$

By (10)

(17) $$r^\circ x = x r^\circ \quad \text{for all } r \in R$$

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). □

Theorem F.3.2. *Let R and P be rings and $x \in P$. Suppose that Conditions (i)-(iv) in 3.1.1 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 $x \in P$, 3.1.1(iii) shows that $x = \sum_{i=0}^{m} e_i x^i$ for some $m \in \mathbb{N}$ and $e_0, e_1, \dots e_n \in \mathbb{R}$. Let $r \in R$. Then

$$rx = r \sum_{i=0}^{n} e_i x^i = \sum_{i=0}^{n} (re_i) x^i.$$

So 3.1.1(iv) shows that $re_1 = r$. Since $rx = xr$ by 3.1.1(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 $f_0, \dots, f_n \in R$. 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. □

Appendix G

Cardinalities

G.1 Cardinalities of Finite Sets

Notation G.1.1. *For $a, b \in \mathbb{Z}$ set* $[a \dots b] := \{c \in \mathbb{Z} \mid a \le c \le b\}$.

Theorem 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 $\alpha(A) \subseteq [1 \dots n] - 1$. □

Theorem G.1.3. *Let $n \in \mathbb{N}$ and let $\beta : [1 \dots n] \to [1 \dots n]$ be a function. If β is injective, then β is surjective.*

Proof. The proof is by induction on n. If $n = 1$, then $\beta(1) = 1$ and so β is surjective. 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 injective we conclude that $\alpha\beta(n) = n$. Thus $\alpha\beta$ is surjective 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 $\alpha : A \to [1 \dots n]$.*

Theorem 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 $n \le m$. Let $\gamma : [1 \dots n] \to [1 \dots m], i \to i$ and $\delta := \gamma \circ \alpha \circ \beta^{-1}$. Then γ is a injective function from $[1 \dots m]$ to $[1 \dots m]$ and so by G.1.3, δ is surjective. Thus also γ is surjective. 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 injective then $|A| \le |B|$, with equality if and only if α is surjective.*

(b) *If $\alpha : A \to B$ is surjective then $|A| \ge |B|$, with equality if and only if α is injective.*

(c) *If $A \subseteq B$ then $|A| \le |B|$, with equality if and only if $|A| = |B|$.*

Proof. (a) If α is surjective then α is a bijection and so $|A| = |B|$. So it suffices to show that if $|A| \ge |B|$, then α is surjective. Put $n = |A|$ and $m = |B|$ and let $\beta : A \to [1 \dots n]$ and $\gamma : B \to [1 \dots m]$ be bijection. Assume $n \ge m$ and let $\delta : [1 \dots m] \to [1 \dots n]$ be the inclusion map. Then $\delta\gamma\alpha\beta^{-1}$ is a injective function form $[1 \dots n]$ to $[1 \dots n]$ and so by G.1.3 its surjective. Hence δ is surjective, $n = m$ and δ is bijection. Since also γ is bijection, this forces $\alpha\beta^{-1}$ to be surjective and so also α is surjective.

(b) Since α is surjective there exists $\beta : B \to A$ with $\alpha\beta = \mathrm{id}_B$. Then β is injective and so by (a), $|B| \le |A|$ 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 $A \to B$. □

Theorem G.1.8. *Let A and be B be finite sets. Then*

(a) *If $A \cap B = \varnothing$, 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 $B = \varnothing$ and so also $A \times B = \varnothing$. 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 1.4.2. □

Bibliography

[1] T.W. Hungerford *Abstract Algebra, An Introduction* second edition, Brooks/Cole **1997**.