Finite-Dimensional Spaces

Algebra, Geometry, and Analysis
Volume I

By

Walter Noll

Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213 USA

This book was published originally by Martinus Nijhoff Publishers in 1987. This is a corrected reprint, posted in 2006 on my website math.cmu.edu/ wn0g/noll.

Introduction

A. **Audience.** This treatise (consisting of the present Vol.I and of Vol.II, to be published) is primarily intended to be a textbook for a core course in mathematics at the advanced undergraduate or the beginning graduate level. The treatise should also be useful as a textbook for selected students in honors programs at the sophomore and junior level. Finally, it should be of use to theoretically inclined scientists and engineers who wish to gain a better understanding of those parts of mathematics that are most likely to help them gain insight into the conceptual foundations of the scientific discipline of their interest.

B. **Prerequisites.** Before studying this treatise, a student should be familiar with the material summarized in Chapters 0 and 1 of Vol.I. Three one-semester courses in serious mathematics should be sufficient to gain such familiarity. The first should be an introduction to contemporary mathematics and should cover sets, families, mappings, relations, number systems, and basic algebraic structures. The second should be an introduction to rigorous real analysis, dealing with real numbers and real sequences, and with limits, continuity, differentiation, and integration of real functions of one real variable. The third should be an introduction to linear algebra, with emphasis on concepts rather than on computational procedures.

C. **Organization.** There are ten chapters in Vol.I, numbered from 0 to 9. A chapter contains from 4 to 12 sections. The first digit of a section number indicates the chapter to which the section belongs; for example, Sect.611 is the 11th section of Chap.6. The appropriate section title and number are printed on the top of each odd-numbered page.

A descriptive name is used for each important theorem. Less important results are called *Propositions* and are enumerated in each section; for example, Prop.5 of Sect.83 refers to the 5th proposition of the 3rd section of Chap.8. Similar enumerations are used, if needed, for formal Definitions, Remarks, and Pitfalls. The term Pitfall is used for comments designed to prevent possible misconceptions.

At the end of most sections there are notes in small print. Their purpose is to relate the notations and terms used in the text to other notations and terms in the mathematical literature, and to comment on symbols, terms, and procedures that appear in print here for the first time (to the best of my knowledge).

A list of problems is presented at the end of each chapter. Some of the problems deal with examples that should help the students to better understand the concepts explained in the text. Some of the problems present additional easy results. A few problems present harder results and may require ingenuity to solve, although hints are often given.

Theorems, Propositions, and Definitions are printed in italic type. Important terms are printed in boldface type when they are defined, so that their definitions can be easily located. The end of a Proof, Remark, or Pitfall is indicated by a block: $\blacksquare$. Words enclosed in brackets indicate substitutes that can be used without invalidating the statement. For example, "The maximum [minimum] of the set S is denoted by $\max S[\min S]$" is a shorthand for two statements. The first is obtained by omitting the words in brackets and the second is "The minimum of the set S is denoted by $\min S$".

D. **Style.** I do not believe that it is a virtue to present merely the logical skeleton of an argument and to hide carefully all motivations from students by refusing to draw pictures and analogies. Unfortunately, it is much easier to write down formal definitions, theorems, and proofs than to describe motivations. I wish I had been able to do more of the latter. In this regard, a large contribution must be made by the instructor who uses this treatise as a textbook.

I have tried to be careful and honest with wording. For example, the phrases "it is evident" and "clearly" mean that a student who has understood the terms used in the statement should be able to agree immediately. The phrases "it is easily seen" and "an easy calculation shows" mean that it should take a good student no more than fifteen minutes with pencil and paper to verify the result. The phrase "it can be shown" is for information only and is not intended to imply anything about the difficulty of the proof.

A mathematical symbol should be used only in one of three ways. It should denote a specific object (e.g. $0, 1, \mathbb{N}, \mathbb{R}$), or it should denote a definite but unspecified object (e.g. n in "Let $n \in \mathbb{N}$ be given"), or it should be a dummy (e.g. n in "$2n$ is even for all $n \in \mathbb{N}$"). I have tried to avoid any other uses. Thus, there should be no "dangling dummies" or expressions such as $\frac{dy}{dx}$, to which I cannot assign a reasonable, precise meaning, although they appear very often in the literature.

E. **Genesis.** About 25 years ago I started to write notes for a course for seniors and beginning graduate students at Carnegie Institute of Technology (renamed Carnegie-Mellon University in 1968). At first, the course was entitled "Tensor Analysis". I soon realized that what usually passes for "Tensor Analysis" is really an undigested mishmash of linear and multilinear algebra, differential calculus in finite-dimensional spaces, manipulation of curvilinear coordinates, and differential geometry on manifolds, all treated with mindless formalisms and without real insight. As a result, I omitted the abstract differential geometry, which is too difficult to be treated properly at this level, and renamed the course "Multidimensional Algebra, Geometry, and Analysis", and later "Finite-Dimensional Spaces". The notes were rewritten several times. They were widely distributed and they served as the basis for appendices to the books *Viscometric Flows of Non-Newtonian Fluids* by B. D. Coleman, H. Markovitz, and W. Noll (Springer-Verlag 1966) and *A First Course in Rational Continuum Mechanics* by C. Truesdell (Academic Press 1977).

Since 1973 my notes have also been used by J. J. Schäffer and me in an undergraduate honors program entitled "Mathematical Studies". One of the purposes of the program has been to present mathematics as an integrated whole and to avoid its traditional division into separate and seemingly unrelated courses. In this connection, Schäffer and I gradually developed a system of notation and terminology that we believe is useful for all branches of mathematics. My involvement in the Mathematical Studies Program has had a profound influence on my thinking; it has led to radical revisions of my notes and finally to this treatise.

Chapter 9 of Vol. I is an adaptation of notes entitled "On the Structure of Linear Transformations", which were written for a course in "Modern Algebra". (They were issued as Report 70–12 of the Department of Mathematics, Carnegie-Mellon University, in March 1970.)

F. **Apologia.** I wish to list certain features which make this treatise different from much, and in some cases most, of the existing literature.

Much of the substance of this treatise is covered in textbooks with titles such as "Linear Algebra", "Analytic Geometry", "Finite-Dimensional Vector Spaces", "Modern Algebra", "Vector and Tensor Analysis", "Advanced Calculus", "Functions of Several Variables", or " Elementary Differential Geometry". However, I believe this treatise to be the first that deals with finite-dimensional spaces in a unified way and that emphasizes the interplay between algebra, geometry, and analysis.

The approach of this treatise is conceptual, geometric, and uncompromisingly "coordinate-free". In some of the literature, "tensors" are still defined in terms of coordinates and their transformations. To me, this is like looking at shadows dancing on the wall rather than at reality itself. Coordinates have no place in the definition of concepts. Of course, when it comes to dealing with specific problems, coordinates are sometimes useful. For this reason, I have included a chapter in which I show how to handle coordinates efficiently.

The space $\mathbb{R}^n$, with $n \in \mathbb{N}$, is very rarely mentioned in this treatise. It is misused nearly every time it appears in the literature, because it is only a special model for the structure that is appropriate in most situations, and as a special model $\mathbb{R}^n$ contains extraneous features that impede geometric insight. Thus any textbook on finite-dimensional calculus with a title like "Functions of Several Variables" must be defective. I consider it a travesty to call $\mathbb{R}^n$ "the Euclidean n-space", as so many do. To quote N. D. Goodman: "Obviously, this is not what Euclid meant" (in "Mathematics as an objective science", Am. Math. Monthly, Vol. 86, p. 549, 1979).

In this treatise, I have tried to present every mathematical topic in a setting that fits the topic naturally and hence leads to a maximum of insight. For example, the structure of a flat (a.k.a. affine) space is the most natural setting for the differential and integral calculus. Most treatments use $\mathbb{R}^n$, a linear space, a normed linear space, or a Euclidean space as the setting. Each of them has extraneous structure which conceals the true nature of the calculus. On the other hand, the structure of a differentiable manifold is too impoverished to be a setting for many aspects of calculus.

In this treatise, a very careful distinction is made between a set and a family (see Sect.02). Almost all the literature is very sloppy on this point. I have found it liberating to resist the compulsion to think of finite sets always in enumerated form and thus to confuse them with lists. Also, I have found it very useful to be able to use a single symbol for a family and to use the *same* symbol, with an index, for the terms of the family. For example, I use $M_{i,j}$ for the (i,j)-term of the matrix M. It seems nonsensical to me to change from an upper case letter M to the lower case letter m when changing from the matrix to its terms. A notation such as $(m_{i,j})$ for a matrix, often seen in textbooks, is poison to me because it contains the dangling dummies i and j (dangling dummies are like cigarettes: both are poison, but I used both when young.)

In this treatise, I have paid more careful attention than is usual to the specification of domains and codomains of mappings. In particular, the adjustment of domains and codomains is done explicitly (see Sect.03). Most authors are either ambiguous on this matter or talk around it clumsily.

The terminology and notation used in this treatise have been very carefully selected. They often do not conform with what a particular reader might consider “standard”, especially since what is “standard” for an expert in category theory may differ from what is “standard” for an expert in engineering mechanics. When the more common terminology or notation is obscure, clumsy, or leads to clashes, I have introduced new terminology or notation. Perhaps the most conspicuous example is the term “lineon” for the too cumbersome “linear transformation”. Some terms, such as the noun “tensor”, have been used with so many different meanings in the past that I found it wise to avoid them altogether.

I have avoided naming concepts and theorems after their purported inventors or discoverers. For one thing, there is often much doubt about who the originators were and frequently credit is given to the wrong people. Secondly, the use of descriptive names makes it easier to learn the material. I have introduced such descriptive names in almost all cases. The experienced mathematician will rarely have any difficulty in understanding my meaning. For example, I have yet to find a mathematician who could not tell immediately that my “Inner-Product Inequality” is what is commonly called the “Cauchy-Schwarz Inequality”. The notes at the end of each section list the names that have been used elsewhere for the concepts and theorems introduced.

Chapter 0

Basic Mathematics

In this chapter, we introduce the notation and terminology used throughout the book. Also, we give a brief explanation of the basic concepts of contemporary mathematics to the extent needed in this book. Finally, we give a summary of those topics of elementary algebra and analysis that are a prerequisite for the remainder of the book.

00 Notations

The equality sign $=$ is used to express the assertion that on either side of $=$ are symbolic names (possibly very complicated) for one and the same object. Thus, $a = b$ means that a and b are names for the same object; $a \neq b$ means that a and b are names for distinct objects. The symbol $:=$ is used to mean that the left side is *defined* by the right side, that the left side is an abbreviation of the right side, or that the right side is to be substituted for the left side. The symbol $=:$ has an analogous meaning.

The logical equivalence sign $\Leftrightarrow$ is used to indicate logical equivalence of statements. The symbol $:\Leftrightarrow$ is used to *define* a phrase or property; it may be read as "means by definition that" or "is equivalent by definition to".

Given a set S and a property p that any given member of S may or may not have, we use the shorthand

$$?\, x \in S, \quad x \text{ has the property } p \tag{00.1}$$

to describe the problem "Find all $x \in S$, if any, such that x has the property p". An element of S having the property p is called a **solution** *of the problem.* Often, the property p involves an equality; then the problem is called an **equation**.

On occasion we use one and the same symbol for two different objects. Of course, this is permissible only if the two objects are related by some natural correspondence and if the context makes it clear which of the meanings of the symbol is appropriate in each instance. Suppose that S and T are two sets and that there is a natural one-to-one correspondence between them. We say that we **identify** S and T and we write $S \cong T$ if we wish to use the same symbol for an element of S and the corresponding element of T. Identification must be handled with great care to avoid notational clashes, i.e. instances in which the meaning of a symbol becomes truly ambiguous, even in context. On the other hand, total avoidance of identifications would lead to staggering notational complexity.

We consider 0 to be a real number that is both positive and negative. If we wish to exclude 0, we use the terms "*strictly positive*" and "*strictly negative*". The following is a list of notations for specific number sets.

Set	Description
$\mathbb{N}$	set of all natural numbers 0, 1, 2, ..., including 0.
$\mathbb{Z}$	set of all integers $\ldots - 2, -1, 0, 1, 2 \ldots$
$\mathbb{Q}$	set of all rational numbers.
$\mathbb{R}$	set of all real numbers.
$\mathbb{P}$	set of all positive real numbers, including 0.
$\mathbb{C}$	set of all complex numbers.
$n^{[}$	given $n \in \mathbb{N}$, the set of the first n natural numbers $0, 1, \ldots, n-1$, starting with 0.
$n^{]}$	given $n \in \mathbb{N}$, the set of the first n natural numbers $1, 2, \ldots, n$, starting with 1.

It is useful to read $n^{[}$ as "n out" and $n^{]}$ as "n in".

We use a superscript cross to indicate that 0 has been taken out of a set. For example, $\mathbb{P}^{\times}$ denotes the set of all strictly positive numbers and $\mathbb{N}^{\times}$ the set of all non-zero natural numbers. This cross notation may be used on any set that has a special "zero-element", not necessarily the number zero.

Notes 00

(1) The "quantifier" ? as in (00.1) was introduced by J. J. Schäffer in about 1973.

(2) Many people say "positive" when we say "strictly positive" and use the awkward "non-negative" when we say "positive".

(3) The use of the special letters $\mathbb{N}, \ldots, \mathbb{C}$ for the various number sets has become fairly standard in recent years. However, some people use $\mathbb{N}$ for what we denote by $\mathbb{N}^{\times}$,

the set of non-zero natural numbers; they do not consider zero to be a natural number. Sometimes $\mathbb{P}$ is used for what we call $\mathbb{P}^\times$, the set of strictly positive real numbers. The notation $\mathbb{R}^+$ for what we denote by $\mathbb{P}$ is often used. Older textbooks often use boldface letters or script letters instead of the special letters now common.

(4) In most of the literature, $S \cong T$ is used to indicate that S is isomorphic to T. I prefer to use $\cong$ only when the isomorphism is natural and used for identification.

(5) The notations $n^{[}$ and $n^{]}$ were invented by J. J. Schäffer in about 1973. I cannot understand any more how I ever got along without them. Some textbooks use the boldface **n** for what we call $n^{]}$. I do not consider the change from lightface to boldface a legitimate notation for a functorial process, quite apart from the fact that it is impossible to produce on a blackboard.

(6) The use of the superscript $\times$ to indicate the removal of 0 was introduced by J. J. Schäffer and me in about 1973. It has turned out to be a very effective notation. It is consistent with the commonly found notation $\mathbb{F}^\times$ for the multiplicative group of a field $\mathbb{F}$ (see Sect.06).

01 Sets, Partitions

To specify a set S, one must have a criterion for deciding whether any given object x belongs to S. If it does, we write $x \in S$ and say that x is a **member** or an **element** of S, that S **contains** x, or that x is **in** S or **contained in** S. If x does not belong to S we write $x \notin S$. We use abbreviations such as "$x, y \in S$" for "$x \in S$ and $y \in S$".

Let S and T be sets. If every member of S is also a member of T we write $S \subset T$ or $T \supset S$ and say that S is a **subset** of T, that S is **included** in T, or that T **includes** S. We have

$$S = T \iff (S \subset T \text{ and } T \subset S). \tag{01.1}$$

If $S \subset T$ but $S \neq T$ we write $S \subsetneqq T$ and say that S is a **proper subset** of T or that S is **properly included** in T.

There is exactly one set having no members at all; it is denoted by $\emptyset$ and called the **empty set**. The empty set is a subset of every set. A set having exactly one member is called a **singleton**; it is denoted by $\{a\}$ if a denotes its only member. If the set S is known to be a singleton, we write $a :\in S$ to indicate that we wish to denote the only member of S by a. A set having exactly two members is called a **doubleton**; it is denoted by $\{a, b\}$ if a and b denote its two (distinct) members.

A set $\mathcal{C}$ whose members are themselves sets is often called a **collection** of sets. The collection of all subsets of a given set S is denoted by $\operatorname{Sub} S$.

Hence, if T is a set, then

$$T \subset S \iff T \in \operatorname{Sub} S. \tag{01.2}$$

Many sets in mathematics are specified by naming an encompassing set A and a property p that any given member of A may or may not have. The set S of all members of A that have this property p is denoted by

$$S := \{x \in A \mid x \text{ has the property } p\}, \tag{01.3}$$

which is read as "S is the set of all x in A such that x has the property p". Occasionally, one has no encompassing set and p is a property that any object may or may not have. In this case, (01.3) is replaced by

$$S := \{x \mid x \text{ has the property } p\}. \tag{01.4}$$

Remark: Definitions of sets of the type (01.4) must be treated with caution. Indiscriminate use of (01.4) can lead to difficulties known as "paradoxes". ■

Sometimes, a set with only few members is specified by an explicit listing of its members and by enclosing the list in braces $\{\}$. Thus, $\{a, b, c, d\}$ denotes the set whose members are a, b, c and d.

Given any sets S and T, one can form their **union** $S \cup T$, consisting of all objects that belong either to S or to T (or to both), and one can form their **intersection** $S \cap T$, consisting of all objects that belong to both S and T. We say that S and T are **disjoint** if they have no elements in common; i.e. if $S \cap T = \emptyset$. The following rules (01.5)–(01.10) are valid for any sets S, T, U.

$$S \cup S = S \cap S = S \cup \emptyset = S, \quad S \cap \emptyset = \emptyset, \tag{01.5}$$

$$T \subset S \iff T \cup S = S \iff T \cap S = T. \tag{01.6}$$

The following rules remain valid if $\cap$ and $\cup$ are interchanged.

$$S \cup T = T \cup S, \tag{01.7}$$

$$(S \cup T) \cup U = S \cup (T \cup U), \tag{01.8}$$

$$(S \cup T) \cap U = (S \cap U) \cup (T \cap U), \tag{01.9}$$

$$T \subset S \implies T \cup U \subset S \cup U. \tag{01.10}$$

Given any sets S and T, the set of all members of S that do not belong to T is called the **set-difference** of S and T and is denoted by $S \setminus T$, so that

$$S \setminus T := \{x \in S \mid x \notin T\}. \tag{01.11}$$

We read $S \setminus T$ as "S without T". The following rules (01.12)–(01.16) are valid for any sets S, T, U.

$$S \setminus S = \emptyset, \quad S \setminus \emptyset = S, \tag{01.12}$$
$$S \setminus (T \setminus U) = (S \setminus T) \cup (S \cap U), \tag{01.13}$$
$$(S \setminus T) \setminus U = (S \setminus T) \cap (S \setminus U). \tag{01.14}$$

The following rules remain valid if $\cap$ and $\cup$ are interchanged.

$$(S \cup T) \setminus U = (S \setminus U) \cup (T \setminus U), \tag{01.15}$$

$$S \setminus (T \cup U) = (S \setminus T) \cap (S \setminus U). \tag{01.16}$$

If T is a subset of S, then $S \setminus T$ is also called the **complement** of T in S. The complement $S \setminus T$ is the largest (with respect to inclusion) among all the subsets of S that are disjoint from T.

The **union** $\bigcup \mathcal{C}$ of a collection $\mathcal{C}$ of sets is defined to be the set of all objects that belong to at least one member-set of the collection $\mathcal{C}$. For any sets S and T we have

$$\bigcup \emptyset = \emptyset, \ \bigcup \{S\} = S, \ \bigcup \{S, T\} = S \cup T. \tag{01.17}$$

If $\mathcal{C}$ and $\mathcal{D}$ are collections of sets then

$$\bigcup (\mathcal{C} \cup \mathcal{D}) = (\bigcup \mathcal{C}) \cup (\bigcup \mathcal{D}). \tag{01.18}$$

The **intersection** $\bigcap \mathcal{C}$ of a non-empty collection $\mathcal{C}$ of sets is defined to be the set of all objects that belong to each of the member-sets of the collection $\mathcal{C}$. For any sets S and T we have

$$\bigcap \{S\} = S, \ \bigcap \{S, T\} = S \cap T. \tag{01.19}$$

If $\mathcal{C}$ and $\mathcal{D}$ are non-empty collections, then

$$\bigcap (\mathcal{C} \cup \mathcal{D}) = (\bigcap \mathcal{C}) \cap (\bigcap \mathcal{D}). \tag{01.20}$$

We say that a collection $\mathcal{C}$ of sets is **disjoint** if any two distinct member-sets of $\mathcal{C}$ are disjoint. We say that $\mathcal{C}$ **covers** a given set S if $S \subset \bigcup \mathcal{C}$.

Let a set S be given. A disjoint collection $\mathcal{P}$ of non-empty subsets of S that covers S is called a **partition** of S. The member-sets of $\mathcal{P}$ are called the **pieces** of the partition $\mathcal{P}$. The empty collection is the only partition of the empty set. If S is any set then $\{E \in \operatorname{Sub} S \mid E \text{ is a singleton } \}$ is a partition of S, called the **singleton-partition** of S. If $S \neq \emptyset$, then $\{S\}$ is also a partition of S, called the **trivial partition**. If T is a non-empty proper subset of S, then $\{T, S \setminus T\}$ is a partition of S.

Let S be a set and let $\sim$ be a **relation** on S, i.e. a fragment that becomes a statement $x \sim y$, true or not, when $x, y \in S$. We say that $\sim$ is an **equivalence relation** if for all $x, y, z \in S$ we have

(i) $x \sim x$ (reflexivity),

(ii) $x \sim y \implies y \sim x$ (symmetry), and

(iii) $(x \sim y \text{ and } y \sim z) \implies x \sim z$ (transitivity).

If $\sim$ is an equivalence relation on S, then

$$\mathcal{P} := \{P \in \operatorname{Sub} S \mid P = \{x \in S \mid x \sim y\} \quad \text{for some} \quad y \in S\}$$

is a partition of S; its pieces are called the **equivalence classes** of the relation $\sim$, and we have $x \sim y$ if and only if x and y belong to the same piece of $\mathcal{P}$. Conversely, if $\mathcal{P}$ is a partition of S, then

$$x \sim y :\iff \quad (\text{for some } P \in \mathcal{P},\ x, y \in P)$$

defines an equivalence relation on S whose equivalence classes are the pieces of $\mathcal{P}$.

Notes 01

(1) Some authors use $\subseteq$ when we use $\subset$, and $\subset$ when we use $\subsetneqq$. There is some confusion in the literature concerning the use of "contain" and "include". We carefully observe the distinction.

(2) The term"null set" is often used for what we call the "empty set". Also the phrase "S is void" instead of "S is empty" can often be found.

(3) The notation $\operatorname{Sub} S$ is used here for the first time. The notations $\mathfrak{P}(S)$ and 2^S are common. The collection $\operatorname{Sub} S$ is often called the "power set" of S.

(4) The notation S–T instead of $S \setminus T$ is used by some people. It clashes with the member-wise difference notation (06.16). If T is a subset of S, the notations $\complement_S T$ or T^c are used by some people for the complement $S \setminus T$ of T in S.

02 Families, Lists, Matrices

A **family** a is specified by a procedure by which one associates with each member i of a given set I an object a_i. The given set I is called the **index set** of the family and the object a_i is called the **term** of index i or simply the i-*term* of the family a. If a and b are families with the same index set I and if $a_i = b_i$ for all $i \in I$, then a and b are considered to be the same, i.e. $a = b$. The notation $(a_i \mid i \in I)$ is often used to denote a family, especially if no name is available a priori.

The set of all terms of a family a is called the **range** of a and is denoted by $\operatorname{Rng} a$ or $\{a_i \mid i \in I\}$, so that

$$\operatorname{Rng} a = \operatorname{Rng}(a_i \mid i \in I) = \{a_i \mid i \in I\}. \tag{02.1}$$

Many sets in mathematics are specified by naming a family a and by letting the set be the range (02.1) of a. We say that a family $a = (a_i \mid i \in I)$ is **injective** if, for all $i, j \in I$,

$$a_i = a_j \implies i = j.$$

Roughly, a family is injective if there is no repetition of terms.

The concept of a family may be viewed as a generalization of the concept of a set. With each set S one can associate a family by letting the index set be S itself and by letting the term corresponding to any given $x \in S$ be x itself. Thus, the family corresponding to S is $(x \mid x \in S)$. We identify this family with S and refer to it as "S **self-indexed**." In this manner, every assertion involving arbitrary families includes, as a special case, an assertion involving sets. The empty set $\emptyset$ is identified with the empty family, which is the only family whose index set is empty.

If all the terms of a given family a belong to a given set S; i.e. if $\operatorname{Rng} a \subset S$, we say that a is a **family in** S. The set of all families in S with a given index set I is denoted by S^I and called the I-**set-power** of S. If $T \subset S$, then $T^I \subset S^I$. We have $S^\emptyset = \{\emptyset\}$.

Let $n \in \mathbb{N}$ be given. A family whose index set is $n^]$ or $n^[$ is called a **list of length** n. If n is small, a list a indexed on $n^]$ can often be specified by a bookkeeping scheme of the form

$$(a_1, a_2, \ldots, a_n) := (a_i \mid i \in n^]). \tag{02.2}$$

where $a_1, a_2, \ldots, a_n$ should be replaced by specific names of objects, to be filled in an actual use. For each $i \in n^]$, we call a_i the i'*th term* of a. The only list of length 0 is the empty family $\emptyset$. A list of length 1 is called a **singlet**, a list of length 2 is called a **pair**, a list of length 3 is called a **triple**, and a list of length 4 a **quadruple**.

If S is a set and $n \in \mathbb{N}$, we use the abbreviation $S^n := S^{n^\rceil}$ for the set of all lists of length n in S. We call S^n the n'*th set-power* of S.

Let S and T be sets. The set of all pairs whose first term is in S and whose second term is in T is called the **set-product** of S and T and is denoted by $S \times T$; i.e.

$$S \times T := \{(x, y) \mid x \in S,\ \ y \in T\}. \tag{02.3}$$

We have $S \times T = \emptyset$ if and only if $S = \emptyset$ or $T = \emptyset$. We have $S \times S = S^{2^\rceil} = S^2$; we call it the **set-square** of S.

A family whose index set is the set-product $I \times J$ of given sets I and J is called an $(I \times J)$-**matrix**. The (i, j)-term of an $(I \times J)$-matrix M is usually written $M_{i,j}$ instead of $M_{(i,j)}$. For each $i \in I$, the family $(M_{i,j} \mid j \in J)$ is called the i-**row** of M, and for each $j \in J$, the family $(M_{i,j} \mid i \in I)$ is called the j-**column** of M. The $(J \times I)$-matrix $M^\top$ defined by $(M^\top)_{j,i} := M_{i,j}$ for all $(j, i) \in J \times I$ is called the **transpose** of M. If $m,\ n \in \mathbb{N}$, an $(m^\rceil \times n^\rceil)$-matrix M is called an *m-by-n-matrix.* Its rows are lists of length n and its columns are lists of length m. If m and n are small, an m-by-n-matrix can often be specified by a bookkeeping scheme of the form

$$\begin{bmatrix} M_{1,1} & M_{1,2} & \cdots & M_{1,n} \\ M_{2,1} & M_{2,2} & \cdots & M_{2,n} \\ \vdots & \vdots & & \vdots \\ M_{m,1} & M_{m,2} & \cdots & M_{m,n} \end{bmatrix} := M. \tag{02.4}$$

A family M whose index set is the set-square I^2 of a given set I is called a **square matrix**. We then say that M is **symmetric** if $M = M^\top$, i.e. if $M_{i,j} = M_{j,i}$ for all $i, j \in I$. The family $(M_{i,i} \mid i \in I)$ is called the **diagonal** of M. We say that a square matrix M in $\mathbb{R}$ (or in any set containing a zero element 0) is a **diagonal matrix** if $M_{i,j} = 0$ for all $i, j \in I$ with $i \neq j$.

Let a set S be given. For every $U \in \operatorname{Sub} S$ we define the **characteristic family** $\mathrm{ch}_{U \subset S}$ of U in S by

$$(\mathrm{ch}_{U \subset S})_x := \left\{ \begin{array}{lll} 1 & \text{if} & x \in U \\ 0 & \text{if} & x \in S \backslash U \end{array} \right\}. \tag{02.5}$$

If the context makes it clear what S is, we abbreviate $\mathrm{ch}_{U \subset S}$ to ch_U.

A family whose index set is $\mathbb{N}$ or $\mathbb{N}^\times$ is called a **sequence**.

Notes 02

(1) A family is sometimes called an "indexed set". The trouble with this term is that an "indexed set" is not a set. The notation $\{a_i\}_{i\in I}$ is used by some people for what we denote by $(a_i \mid i \in I)$. Some people even use just $\{a_i\}$, which is poison because of the dangling dummy i and because it also denotes a singleton with member a_i.

(2) The terms "member" or "entry of a family" are often used instead of "term of a family". The use of "member" can lead to confusion with "member of a set".

(3) One often finds the barbarism "n-tuple" for what we call "list of length n". The term "finite sequence" is sometimes used for what we call "list". A list of numbers is very often called a "vector". I prefer to use the term "vector" only when it has its original geometric meaning (see Def.1 of Sect.32).

(4) The terms "Cartesian product" and "direct product" are often used for what we call "set product".

(5) In most of the literature, the use of the term "matrix" is confined to the case when the index set is of the form $n^{]} \times m^{]}$ and when the terms are numbers of some kind. The generalization used here turns out to be very useful.

03 Mappings

In order to specify a **mapping** f, one first has to prescribe two sets, say D and C, and then some kind of prescription, called the **assignment rule** of f, by which one can assign to every element $x \in D$ an element $f(x) \in C$. We call $f(x)$ the **value** of f **at** x. It is very important to distinguish very carefully between the mapping f and its values $f(x),\ x \in D$. The set D of objects to which the prescription embodied in f can be applied is called the **domain** of the mapping f and is denoted by $\operatorname{Dom} f := D$. The set C to which the values of f must belong is called the **codomain** of f and is denoted by $\operatorname{Cod} f := C$. In order to put C and D into full view, we often write

$$f : D \to C \qquad \text{or} \qquad D \xrightarrow{f} C$$

instead of just f and we say that f *maps* D *to* C or that f is a mapping *from* D *to* C. The phrase "f *is defined on* D" expresses the assertion that D is the domain of f. If f and g are mappings with $\operatorname{Dom} f = \operatorname{Dom} g$, $\operatorname{Cod} f = \operatorname{Cod} g$, and $f(x) = g(x)$ for all $x \in \operatorname{Dom} f$, then f and g are considered to coincide, i.e. $f = g$.

Terms such as "function", "map", "functional", "transformation", and "operator" are often used to mean the same thing as "mapping". The term "function" is preferred when the codomain is the set of real or complex numbers or a subset thereof. A still greater variety of names is used for mappings having special properties. Also, in some contexts, the value of f at x is not written $f(x)$ but fx, xf, f_x or x^f.

In order to specify a mapping f explicitly without introducing unnecessary symbols, it is often useful to employ the notation $(x \mapsto f(x)) : \operatorname{Dom} f \to \operatorname{Cod} f$ instead of just f. (Note the use of $\mapsto$ instead of $\to$.) For example, $(x \mapsto \frac{1}{x}) : \mathbb{R}^\times \to \mathbb{R}$ denotes the function f with $\operatorname{Dom} f := \mathbb{R}^\times, \operatorname{Cod} f := \mathbb{R}$ and evaluation rule

$$f(x) := \frac{1}{x} \quad \text{for all} \quad x \in \mathbb{R}^\times.$$

The **graph** of a mapping $f : D \to C$ is the subset $\operatorname{Gr}(f)$ of the set-product $D \times C$ defined by

$$\operatorname{Gr}(f) := \{(x, y) \in D \times C \mid y = f(x)\}. \tag{03.1}$$

The mappings f and g coincide if and only if they have the same domain, codomain, and graph.

Remark: Very often, a mapping is specified by two sets D and C and a statement scheme $F(x, y)$, which may be become valid or not, depending on what elements of D and C are substituted for x and y, respectively. If, for every $x \in D$, there is exactly one $y \in C$ such that $F(x, y)$ is valid, then F defines a mapping $f : D \longrightarrow C$, namely by the prescription that assigns to $x \in D$ the unique $y \in C$ that makes $F(x, y)$ valid. Then

$$\operatorname{Gr}(f) = \{(x, y) \mid F(x, y) \ \text{ is valid } \} \ .$$

In some cases, given $x \in D$, one can define or obtain $f(x)$ by a formula, algorithm, or other procedure. Finding an efficient procedure to this end is often a difficult task.

With every mapping f we can associate the family $(f(x) \mid x \in \operatorname{Dom} f)$ of its values. Roughly, the family is obtained from the mapping by forgetting the codomain. Conversely, with every family $a := (a_i \mid i \in I)$ and every set C that includes $\operatorname{Rng} a$ we can associate the mapping $(i \mapsto a_i) : I \to C$. Roughly, the mapping is obtained from the family by specifying the codomain C. For example, if U is a subset of a given set S, we can obtain from the characteristic family of U in S, defined by (02.5), the **characteristic function** of U in S, also denoted by $\operatorname{ch}_{U \subset S}$ or simply ch_U, by specifying a codomain, usually $\mathbb{R}$.

The **range** of a mapping f is defined to be the range of the family of its values; i.e., with the notation (02.1), we have

$$\operatorname{Rng} f := \{f(x) \mid x \in \operatorname{Dom} f\}. \tag{03.2}$$

We say that f is **surjective** if its codomain and range coincide, i.e. if $\operatorname{Cod} f = \operatorname{Rng} f$.

We say that a mapping f is **injective** if the family of its values is injective; i.e. if for all $x, y \in D$

$$f(x) = f(y) \implies x = y.$$

A mapping $f : D \to C$ is said to be **invertible** if it is both injective and surjective. This is the case if and only if there is a mapping $g : C \to D$ such that for all $x \in D$ and $y \in C$

$$y = f(x) \iff x = g(y).$$

The mapping g is then uniquely determined by f; it is called the **inverse** of f and is denoted by $f^{\leftarrow}$. A given mapping $f : D \to C$ is invertible if and only if, for each $y \in C$, the problem

$$?\, x \in D, \quad f(x) = y$$

has exactly one solution. This solution is then given by $f^{\leftarrow}(y)$. Invertible mappings are also called **bijections** and injective mappings **injections**.

Let U be a subset of a given set S. The **inclusion mapping** *of* U *into* S is the mapping $1_{U \subset S} : U \to S$ defined by the rule

$$1_{U \subset S}(x) := x \qquad \text{for all } x \in U. \tag{03.3}$$

The inclusion mapping of the set S into itself is called the **identity mapping** of S and is denoted by $1_S := 1_{S \subset S}$. Inclusion mappings are injective. An identity mapping is invertible and equal to its own inverse.

A **constant** is a mapping whose range is a singleton. We sometimes use the symbol $c_{D \to C}$ to denote the constant with domain D, codomain C, and range $\{c\}$. In most cases, we can identify $c_{D \to C}$ with c itself, thus using the same symbol for the constant and its only value.

If f and g are mappings such that $\operatorname{Dom} g = \operatorname{Cod} f$, we define the **composite** $g \circ f : \operatorname{Dom} f \to \operatorname{Cod} g$ of f and g by the evaluation rule

$$(g \circ f)(x) := g(f(x)) \qquad \text{for all } x \in \operatorname{Dom} f. \tag{03.4}$$

If f, g and h are mappings with $\operatorname{Dom} g = \operatorname{Cod} f$ and $\operatorname{Cod} g = \operatorname{Dom} h$, then

$$(h \circ g) \circ f = h \circ (g \circ f). \tag{03.5}$$

Because of this rule, we may omit parentheses and write $h \circ g \circ f$.

Let f be a mapping from a set D to itself. For every $n \in \mathbb{N}$, the n'th **iterate** $f^{\circ n} : D \to D$ of f is defined recursively by $f^{\circ 0} = 1_D$, and $f^{\circ(k+1)} := f \circ f^{\circ k}$ for all $k \in \mathbb{N}$. We have $f^{\circ 1} = f$, $f^{\circ 2} = f \circ f$, $f^{\circ 3} = f \circ f \circ f$, etc. An element $z \in D$ is called a **fixed point** of f if $f(z) = z$. We say that given mappings f and g, both from D to itself, **commute** if $f \circ g = g \circ f$.

Let a mapping $f : D \to C$ be given. A mapping $g : C \to D$ is called a **right-inverse** of f if $f \circ g = 1_C$, a **left-inverse** of f if $g \circ f = 1_D$. If f has a right-inverse, it must be surjective; if f has a left-inverse, it must be injective. If g is a right-inverse of f and h a left-inverse of f, then f is invertible and $g = h = f^{\leftarrow}$. We always have

$$f \circ 1_{\operatorname{Dom} f} = 1_{\operatorname{Cod} f} \circ f = f. \tag{03.6}$$

Again, let a mapping $f : D \to C$ be given. We define the **image mapping** $f_> : \operatorname{Sub} D \to \operatorname{Sub} C$ of f by the evaluation rule

$$f_>(U) := \{f(x) \mid x \in U\} \quad \text{for all} \quad U \in \operatorname{Sub} D \tag{03.7}$$

and the **pre-image mapping** $f^< : \operatorname{Sub} C \to \operatorname{Sub} D$ of f by the rule

$$f^<(V) := \{x \in D \mid f(x) \in V\} \quad \text{for all} \quad V \in \operatorname{Sub} C. \tag{03.8}$$

The mappings $f_>$ and $f^<$ satisfy the following rules for all subsets U and U' of D and all subsets V and V' of C:

$$U \subset U' \implies f_>(U) \subset f_>(U'), \tag{03.9}$$
$$V \subset V' \implies f^<(V) \subset f^<(V'), \tag{03.10}$$
$$U \subset f^<(f_>(U)), \quad f_>(f^<(V)) = V \cap \operatorname{Rng} f, \tag{03.11}$$
$$f_>(U \cup U') = f_>(U) \cup f_>(U'), \quad f_>(U \cap U') \subset f_>(U) \cap f_>(U'), \tag{03.12}$$
$$f^<(V \cup V') = f^<(V) \cup f^<(V'), \quad f^<(V \cap V') = f^<(V) \cap f^<(V'), \tag{03.13}$$
$$f^<(C \setminus V) = D \setminus f^<(V). \tag{03.14}$$

The inclusions $\subset$ in (03.11) and (03.12) become equalities if f is injective. If f is injective, so is $f_>$, and $f^<$ is a left-inverse of $f_>$. If f is surjective, so is $f_>$, and $f^<$ is a right-inverse of $f_>$. If f is invertible, then $(f_>)^{\leftarrow} = f^<$. If f and g are mappings such that $\operatorname{Dom} g = \operatorname{Cod} f$, then

$$(g \circ f)_> = g_> \circ f_>, \quad (g \circ f)^< = f^< \circ g^<, \tag{03.15}$$

$$\operatorname{Rng}(g \circ f) = g_{>}(\operatorname{Rng} f). \tag{03.16}$$

If $\mathrm{ch}_{V \subset C}$ is the characteristic function of a subset V of a given set C and if $f : D \to C$ is given, we have

$$\mathrm{ch}_{V \subset C} \circ f = \mathrm{ch}_{f^{<}(V) \subset D}. \tag{03.17}$$

Let a mapping f and sets A and B be given. We define the **restriction** $f|_A$ *of* f *to* A by $\operatorname{Dom} f|_A := A \cap \operatorname{Dom} f$, $\operatorname{Cod} f|_A := \operatorname{Cod} f$, and the evaluation rule

$$f|_A(x) := f(x) \quad \text{for all} \quad x \in A \cap \operatorname{Dom} f. \tag{03.18}$$

We define the mapping

$$f|_A^B : \ A \cap f^{<}(B \cap \operatorname{Cod} f) \to B \tag{03.19}$$

by the rule

$$f|_A^B(x) := f(x) \quad \text{for all} \quad x \in A \cap f^{<}(B \cap \operatorname{Cod} f). \tag{03.20}$$

We say that $f\,|_A^B$ is an **adjustment** of f. We have $f\,|_A = f\,|_A^{\operatorname{Cod} f}$. We use the abbreviations

$$f\,|^B := f\,|_{\operatorname{Dom} f}^B, \qquad f\,|^{\operatorname{Rng}} := f\,|_{\operatorname{Dom} f}^{\operatorname{Rng} f}. \tag{03.21}$$

We have $\operatorname{Dom}(f\,|^B) = \operatorname{Dom} f$ if and only if $\operatorname{Rng} f \subset B$. We note that

$$f\,|_A^B = (f\,|_A)\,|^B = (f\,|^B)\,|_A. \tag{03.22}$$

Let f be a mapping from a set D to itself. We say that a subset A of D is f-**invariant** if $f_{>}(A) \subset A$. If this is the case, we define the A-**adjustment** $f_{|A} : A \to A$ of f by

$$f_{|A} := f|_A^A. \tag{03.23}$$

Let a set A and a collection $\mathcal{C}$ of subsets of A be given such that $A \in \mathcal{C}$. For every $S \in \operatorname{Sub} A$, the subcollection $\{U \in \mathcal{C} \mid S \subset U\}$ of $\mathcal{C}$ then contains A and hence is not empty. We define $\operatorname{Sp} : \operatorname{Sub} A \to \operatorname{Sub} A$, the **span-mapping** corresponding to $\mathcal{C}$, by the rule

$$\operatorname{Sp}(S) := \bigcap \{U \in \mathcal{C} \mid S \subset U\} \quad \text{for all} \quad S \in \operatorname{Sub} A. \tag{03.24}$$

The following rules hold for all $S, T \in \operatorname{Sub} A$:

$$S \subset \operatorname{Sp}(S), \tag{03.25}$$

$$\mathrm{Sp}\,(\mathrm{Sp}\,(S)) = \mathrm{Sp}\,(S), \tag{03.26}$$

$$S \subset T \implies \mathrm{Sp}\,(S) \subset \mathrm{Sp}\,(T). \tag{03.27}$$

If the collection $\mathcal{C}$ is **intersection-stable** i.e. if for every non-empty subcollection $\mathcal{D}$ of $\mathcal{C}$ we have $\bigcap \mathcal{D} \in \mathcal{C}$, then, for all $U \in \mathrm{Sub}\, A$,

$$\mathrm{Sp}\,(U) = U \iff U \in \mathcal{C}, \tag{03.28}$$

and we have $\mathrm{Rng}\ \mathrm{Sp} = \mathcal{C}$. Also, for every $S \in \mathrm{Sub}\, A$, $\mathrm{Sp}\,(S)$ is the smallest (with respect to inclusion) member of $\mathcal{C}$ that includes S.

Many of the definitions and rules of this section can be applied if one or more of the mappings involved are replaced by a family. For example, if $a = (a_i \mid i \in I)$ is a family in a set S and if f is a mapping with $\mathrm{Dom}\, f = S$, then $f \circ a$ denotes the family given by $f \circ a := (f(a_i) \mid i \in I)$.

Notes 03

(1) Some of the literature still confuses a mapping f with its value $f(x)$, and one still finds phrases such as "Consider the function $f(x)$", which contains the dangling dummy x and hence is poison.

(2) There is some confusion in the literature concerning the term "image of a mapping f". Some people use it for what we call "value of f" and others for what we call "range of f". Occasionally, the term "range" is used for what we call "codomain".

(3) Many people do not consider the specification of a codomain as part of the specification of a mapping. If we did likewise, we would have no formal distinction between a family and a mapping. I believe it is very useful to have such a distinction.

(4) The terms "onto" for "surjective" and "one-to-one" for "injective" are very often used. Also "one-to-one correspondence" is a common term for what we call "bijection" or "invertible mapping". We sometimes use "one-to-one correspondence" informally.

(5) The notation $f^{\leftarrow}$ for the inverse of the mapping was introduced in about 1973 by J. J. Schäffer and me. The notation f^{-1} is more common, but it clashes with notations for value-wise reciprocals. However, for linear mappings, it is useful to revert to the more traditional notation (see Sect.13).

(6) The notation id_S is often used for the identity mapping 1_S of the set S.

(7) The notations $1_{U \subset S}$ for an inclusion mapping and $c_{D \to C}$ for a constant were introduced by J. J. Schäffer and me in about 1973. Also, we started to write $f^{\circ n}$ instead of the more common f^n for the n'th iterate of f to avoid a clash with notations for value-wise powers.

(8) The notations $f_>$ and $f^<$ for the image mapping and pre-image mapping of a mapping f were introduced by J. J. Schäffer and me in about 1973. Many years before, I had introduced the notations f_* and f^*, which were also introduced, independently, by S. MacLane and G. Birkhoff in their book "Algebra" (MacMillan, 1967). In most of the literature, the same symbol f is used for the image mapping as for the mapping f itself. The pre-image mapping of f is often denoted by $\overset{-1}{f}$ or f^{-1}, the latter leading to confusion with the inverse (if there is one). A distinctive notation for image and pre-image mappings avoids a lot of confusion and leads to great economy when expressing relations (see, for example, (56.3)).

(9) The definition of the adjustment $f|_A^B$ of a mapping f, in the generality given here, is new. For the case when $f_>(A) \subset B$ it has been used by J. J. Schäffer and me since about 1973. I used the notation ${}^B|f|_A$ instead of $f|_A^B$ in a paper published in 1971. Most people "talk around" such adjustments and do not use an explicit notation. The notation $f|_A$ for the restriction is in fairly common use. The notation $f_{|A}$ for $f|_A^A$ was introduced recently by J. J. Schäffer.

04 Families of Sets; Families and Sets of Mappings

Let $(A_i \mid i \in I)$ be a family of sets, i.e. a family whose terms A_i are all sets. The range of this family is the collection $\{A_i \mid i \in I\}$ of sets; it is non-empty if and only if I is not empty. We define the **union** and, if $I \neq \emptyset$, the **intersection** of the family to be, respectively, the union and intersection of the range. We use the notations

$$\bigcup_{i\in I} A_i = \bigcup(A_i \mid i \in I) := \bigcup\{A_i \mid i \in I\}, \tag{04.1}$$

$$\bigcap_{i\in I} A_i = \bigcap(A_i \mid i \in I) := \bigcap\{A_i \mid i \in I\} \quad \text{if} \quad I \neq \emptyset. \tag{04.2}$$

Given any set S, we have the following generalizations of (01.9) and (01.16):

$$S \cap \bigcup_{i\in I} A_i = \bigcup_{i\in I}(S \cap A_i), \qquad S \cup \bigcap_{i\in I} A_i = \bigcap_{i\in I}(S \cup A_i), \tag{04.3}$$

$$S \setminus \bigcup_{i\in I} A_i = \bigcap_{i\in I}(S \setminus A_i), \qquad S \setminus \bigcap_{i\in I} A_i = \bigcup_{i\in I}(S \setminus A_i), \tag{04.4}$$

Given any mapping f with $\bigcup(A_i \mid i \in I) \subset \operatorname{Dom} f$ we have the following generalizations of (03.12):

$$f_>(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} f_>(A_i), \qquad f_>(\bigcap_{i\in I} A_i) \subset \bigcap_{i\in I} f_>(A_i), \tag{04.5}$$

and the inclusion becomes an equality if f is injective. If f is a mapping with $\bigcup(A_i \mid i \in I) \subset \operatorname{Cod} f$, we have the following generalizations of (03.13):

$$f^{<}(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} f^{<}(A_i), \qquad f^{<}(\bigcap_{i\in I} A_i) = \bigcap_{i\in I} f^{<}(A_i). \tag{04.6}$$

The **set-product** of the family $(A_i \mid i \in I)$ of sets is the set of all families a with index set I such that $a_i \in A_i$ for all $i \in I$. This set product is denoted by

$$\bigtimes_{i\in I} A_i = \bigtimes (A_i \mid i \in I)$$

It generalizes the set-product $A_1 \times A_2$ of sets A_1, A_2 as defined in Sect.02 because $A_1 \times A_2$ is the set-product of the pair (A_1, A_2) in the sense just defined. If the terms of a family $(A_i \mid i \in I)$ are all the same, i.e. if there is a set S such that $A_i = S$ for all $i \in I$, then the set-product reduces to the set-power defined in Sect.02, i.e. $\bigtimes (A_i \mid i \in I) = S^I$.

Let a family $(A_i \mid i \in I)$ of sets and $j \in I$ be given. Then the mapping

$$(a \mapsto (a_j, a\,|_{I\setminus\{j\}})) : \bigtimes_{i\in I} A_i \;\to\; A_j \times \bigtimes_{i\in I\setminus\{j\}} A_i \tag{04.7}$$

is a natural bijection.

Let A_1 and A_2 be sets. We define the two **evaluations**

$$\mathrm{ev}_1 : A_1 \times A_2 \to A_1 \quad \text{and} \quad \mathrm{ev}_2 : A_1 \times A_2 \to A_2$$

by the rules

$$\mathrm{ev}_1(a) := a_1,\ \mathrm{ev}_2(a) := a_2 \quad \text{for all} \quad a \in A_1 \times A_2. \tag{04.8}$$

We have

$$a = (\mathrm{ev}_1(a),\ \mathrm{ev}_2(a)) \quad \text{for all } a \in A_1 \times A_2.$$

More generally, given a family $(A_i \mid i \in I)$ of sets, we can define, for each $j \in I$, the evaluation

$$\mathrm{ev}_j : \bigtimes_{i\in I} A_i \to A_j$$

by the rule

$$\mathrm{ev}_j(a) := a_j \quad \text{for all} \quad a \in \bigtimes_{i\in I} A_i, \tag{04.9}$$

and we have

$$a = (\mathrm{ev}_i(a) \mid i \in I) \quad \text{for all} \quad a \in \bigtimes_{i\in I} A_i.$$

Given any sets S and T, we denote the set of all mappings f with $\mathrm{Dom}\, f = S$ and $\mathrm{Cod}\, f = T$ by $\mathrm{Map}\,(S,T)$. The set of all injective mappings from S to T will be denoted by $\mathrm{Inj}\,(S,T)$. The set of all invertible mappings from a set S to itself is denoted by $\mathrm{Perm}\; S$ and its members are called **permutations** of S.

Let S and T be sets. For each $x \in S$, we define the **evaluation at** x, $\mathrm{ev}_x : \mathrm{Map}\,(S,T) \to T$, by the rule

$$\mathrm{ev}_x(f) := f(x) \quad \text{for all} \quad f \in \mathrm{Map}\,(S,T). \tag{04.10}$$

Assume now that a subset F of $\mathrm{Map}\,(S,T)$ is given. Then the restriction $\mathrm{ev}_x |_F$ is also called the evaluation at x and is simply denoted by ev_x if the context makes clear what F is. We define the mapping $\mathrm{ev}^F : S \to \mathrm{Map}\,(F,T)$ by

$$(\mathrm{ev}^F(x))(f) := \mathrm{ev}_x(f) = f(x) \quad \text{for all} \quad x \in S, f \in F. \tag{04.11}$$

This mapping ev^F, or a suitable adjustment of it, is called an **evaluation mapping**; it is simply denoted by ev if the context makes clear what F and the adjustments are.

There is a natural bijection from the set $\mathrm{Map}\,(S,T)$ of all mappings from S to T onto the set T^S of all families in T with index set S, as described at the end of Sect.03. If U is a subset of T we have $U^S \subset T^S$, but $\mathrm{Map}\,(S,U)$ is not a subset of $\mathrm{Map}\,(S,T)$. However, we have the natural injection

$$(f \mapsto f\,|^T) : \ \mathrm{Map}\,(S,U) \to \mathrm{Map}\,(S,T).$$

Let f and g be mappings having the same domain D. The **value-wise pair formation**

$$(x \mapsto (f(x), g(x))) : \ D \to \mathrm{Cod}\, f \times \mathrm{Cod}\, g \tag{04.12}$$

will be identified with the pair (f,g), so that

$$(f,g)(x) := (f(x), g(x)) \quad \text{for all } x \in D. \tag{04.13}$$

Thus, given any sets D, C_1, C_2, we obtain the identification

$$\mathrm{Map}\,(D,C_1) \times \mathrm{Map}\,(D,C_2) \cong \mathrm{Map}\,(D, C_1 \times C_2). \tag{04.14}$$

More generally, given any family $(f_i \mid i \in I)$ of mappings, all having the same domain D, we identify the **term-wise evaluation**

$$(x \mapsto (f_i(x) \mid i \in I)) : \ D \to \bigtimes_{i \in I} \operatorname{Cod} f_i \tag{04.15}$$

with the family itself, so that

$$(f_i \mid i \in I)(x) := (f_i(x) \mid i \in I) \quad \text{for all} \quad x \in D. \tag{04.16}$$

Thus, given any set D and any family $(C_i \mid i \in I)$ of sets, we obtain the identification

$$\bigtimes_{i \in I} \operatorname{Map}(D, C_i) \cong \operatorname{Map}(D, \bigtimes_{i \in I} C_i). \tag{04.17}$$

Let f and g be any mappings. We call the mapping

$$f \times g : \ \operatorname{Dom} f \times \operatorname{Dom} g \ \to \ \operatorname{Cod} f \times \operatorname{Cod} g$$

defined by

$$(f \times g)(x, y) := (f(x), g(y)) \quad \text{for all } x \in \operatorname{Dom} f, \ y \in \operatorname{Dom} g \tag{04.18}$$

the **cross-product** of f and g.

More generally, given any family $(f_i \mid i \in I)$ of mappings, we call the mapping

$$\bigtimes_{i \in I} f_i = \bigtimes (f_i \mid i \in I) : \ \bigtimes_{i \in I} \operatorname{Dom} f_i \ \to \ \bigtimes_{i \in I} \operatorname{Cod} f_i$$

defined by

$$(\bigtimes_{i \in I} f_i)(x) := (f_i(x_i) \mid i \in I) \quad \text{for all } x \in \bigtimes_{i \in I} \operatorname{Dom} f_i \tag{04.19}$$

the **cross-product** of the family $(f_i \mid i \in I)$. If the terms of the family are all the same, i.e., if there is a mapping g such that $f_i = g$ for all $i \in I$, we write $g^{\times I} := \bigtimes (f_i \mid i \in I)$ and call it the I-**cross-power** of g. If no confusion can arise, we write simply $g(c) := g^{\times I}(c) = (g(c_i) \mid i \in I)$ when $c \in (\operatorname{Dom} g)^I$.

Given any families $(D_i \mid i \in I)$ and $(C_i \mid i \in I)$ of sets, the mapping

$$((f_i \mid i \in I) \mapsto \bigtimes_{i \in I} f_i) : \ \bigtimes_{i \in I} \operatorname{Map}(D_i, C_i) \to \operatorname{Map}(\bigtimes_{i \in I} D_i, \bigtimes_{i \in I} C_i) \tag{04.20}$$

is a natural injection.

Let sets S and T be given. For every $a \in S$ and every $b \in T$ we define mappings $(a,\cdot) : T \to S \times T$ and $(\cdot,b) : S \to S \times T$ by the rules

$$(a,\cdot)(y) := (a,y) \quad \text{and} \quad (\cdot,b)(x) := (x,b) \tag{04.21}$$

for all $x \in S, y \in T$. A mapping f whose domain $\operatorname{Dom} f$ is a subset of a set product $S \times T$ is often called a "function of two variables". Its value at $(x,y) \in S \times T$ is written simply $f(x,y)$ rather than $f((x,y))$. Given such a mapping and any $a \in S$, $b \in T$, we define

$$f(a,\cdot) := f \circ (a,\cdot) |^{\operatorname{Dom} f}, \qquad f(\cdot,b) := f \circ (\cdot,b) |^{\operatorname{Dom} f}. \tag{04.22}$$

Hence we have

$$f(\cdot,y)(x) = f(x,y) = f(x,\cdot)(y) \quad \text{for all} \quad (x,y) \in \operatorname{Dom} f. \tag{04.23}$$

More generally, if $(A_i \mid i \in I)$ is a family of sets, we define, for each $j \in I$ and each $c \in \times (A_i \mid i \in I \setminus \{j\})$, the mapping

$$(c.j) : \ A_j \to \mathop{\times}_{i \in I} A_i$$

by the rule

$$((c.j)(z))_i := \left\{ \begin{array}{lll} c_i & \text{if} & i \in I \setminus \{j\} \\ z & \text{if} & i = j \end{array} \right\}. \tag{04.24}$$

If $a \in \times (A_i \mid i \in I)$, we abbreviate $(a.j) := (a\,|_{I \setminus \{j\}}.j)$. If $I = \{1,2\}$, we have $(a.1) = (\cdot, a_2)$, $(a.2) = (a_1, \cdot)$. Given a mapping f whose domain is a subset of $\times (A_i \mid i \in I)$ and given $j \in I$ and $c \in \times (A_i \mid i \in I \setminus \{j\})$ or $c \in \times (A_i \mid i \in I)$, we define

$$f(c.j) := f \circ (c.j) |^{\operatorname{Dom} f}. \tag{04.25}$$

We have

$$f(x.j)(x_j) = f(x) \quad \text{for all } x \in \mathop{\times}_{i \in I} A_i. \tag{04.26}$$

Let S, T and C be sets. Given any mapping $f : S \times T \to C$, we identify the mapping

$$(x \mapsto f(x,\cdot)) : S \to \operatorname{Map}(T,C) \tag{04.27}$$

with f itself, so that

$$(f(x,\cdot))(y) := f(x,y) \quad \text{for all } x \in S, y \in T. \tag{04.28}$$

Thus, we obtain the identification

$$\text{Map}\,(S \times T, C) \cong \text{Map}\,(S, \text{Map}\,(T, C)). \tag{04.29}$$

Notes 04

(1) The notation $\bigcup(A_i \,|\, i \in I)$ for the union of the family $(A_i \,|\, i \in I)$ is used when it occurs in the text rather than in a displayed formula. For typographical reasons, the notation on the left of (04.1) can be used only in displayed formulas. A similar remark applies to intersections, set-products, sums, products, and other operations on families (see (04.2) and (07.1)).

(2) The term "projection" is often used for what we call "evaluation", in particular when the domain is a set-product rather than a set of mappings.

(3) The notation $\prod(A_i \,|\, i \in I)$ is sometimes used for the set-product $\bigtimes (A_i \,|\, i \in I)$, which is often called "Cartesian product" or "direct product" (see Note (4) to Sect.02).

(4) The notations $(a,\cdot)$ and $(\cdot,b)$ as defined by (04.21) and the notation $(c.j)$ as defined by (04.24) were introduced by me a few years ago. The notations $f(\cdot,y)$ and $f(x,\cdot)$, using (04.23) as definitions rather than propositions, are fairly common in the literature.

(5) Many people use the term "permutation" only if the domain is a finite set. I cannot see any advantage in such a limitation.

05 Finite Sets

The **cardinal** of a finite set S, i.e. the number of elements in S, will be denoted by $\sharp S \in \mathbb{N}$. Given $n \in \mathbb{N}$, the range $S := \{a_i \mid i \in n^{]}\}$ of an injective list $(a_i \mid i \in n^{]})$ is a finite set with cardinal $\sharp S = n$. By an **enumeration** of a finite set S we mean a list $(a_i \mid i \in n^{]})$ of length $n := \sharp S$ whose range is S, or the corresponding mapping $(i \mapsto a_i) : n^{]} \to S$. There are $n!$ such enumerations. We have $\sharp\,(n^{]}) = \sharp\,(n^{[}) = n$ for all $n \in \mathbb{N}$.

Let S and T be finite sets. Then $S \cap T$, $S \cup T$, $S \setminus T$, and $S \times T$ are again finite and their cardinals satisfy the rules

$$\begin{aligned} \sharp\,(S \cup T) + \sharp\,(S \cap T) &= \sharp\,S + \sharp\,T, && (05.1)\\ \sharp\,(S \setminus T) &= \sharp\,S - \sharp\,(S \cap T) = \sharp\,(S \cup T) - \sharp\,T, && (05.2)\\ \sharp\,(S \times T) &= (\sharp\,S)(\sharp\,T). && (05.3) \end{aligned}$$

We say that a family is *finite* if its index set is finite. If $(A_i \mid i \in I)$ is a finite family of finite sets, then $\times(A_i \mid i \in I)$ is finite and

$$\sharp\,(\mathop{\times}_{i\in I} A_i) = \prod_{i\in I}(\sharp\, A_i). \tag{05.4}$$

(See Sect.07 concerning the product $\prod(n_i \mid i \in I)$ of a finite family $(n_i \mid i \in I)$ in $\mathbb{N}$.)

Let $\mathcal{P}$ be a partition of the given finite set S. Then $\mathcal{P}$ and the members of $\mathcal{P}$ are all finite and

$$\sharp\, S = \sum_{P\in\mathcal{P}}(\sharp\, P). \tag{05.5}$$

(See Sect.07 concerning the sum $\sum(n_i \mid i \in I)$ of a finite family $(n_i \mid i \in I)$ in $\mathbb{N}$.)

Pigeonhole Principle: *Let f be a mapping with finite domain and codomain. If f is injective, then $\sharp\,\mathrm{Dom}\, f \leq \sharp\,\mathrm{Cod}\, f$. If f is surjective, then $\sharp\,\mathrm{Dom}\, f \geq \sharp\,\mathrm{Cod}\, f$. If $\sharp\,\mathrm{Dom}\, f = \sharp\,\mathrm{Cod}\, f$ then the following are equivalent:*

(i) *f is injective;*

(ii) *f is surjective;*

(iii) *f is invertible.*

If S and T are finite sets, then the set T^S of all families in T indexed on S and the set $\mathrm{Map}\,(S,T)$ of all mappings from S to T are finite and

$$\sharp\,(T^S) = \sharp\,\mathrm{Map}\,(S,T) = (\sharp\, T)^{\sharp\, S} \tag{05.6}$$

If S is finite, so is $\mathrm{Sub}\, S$ and

$$\sharp\,\mathrm{Sub}\, S = 2^{(\sharp\, S)}. \tag{05.7}$$

Let S be any set. We denote the set of all finite subsets of S by $\mathrm{Fin}\, S$, and the set of all subsets of S having m elements by $\mathrm{Fin}_m S$, so that

$$\mathrm{Fin}_m S := \{U \in \mathrm{Fin}\, S \mid \sharp\, U = m\}. \tag{05.8}$$

If S is finite, so is $\mathrm{Fin}_m S$. The notation

$$\begin{pmatrix} n \\ m \end{pmatrix} := \sharp\,\mathrm{Fin}_m(n^{]})$$

(read "n choose m") is customary, and we have

$$\sharp\,\mathrm{Fin}_m(S) = \binom{\sharp\,S}{m} \quad \text{for all } m \in \mathbb{N}. \tag{05.9}$$

We have the following rules, valid for all n, $m \in \mathbb{N}$:

$$\binom{n}{m} = 0 \quad \text{if} \quad m > n, \qquad \binom{n}{0} = \binom{n}{n} = 1, \tag{05.10}$$

$$\binom{n+m}{m} = \binom{n+m}{n}, \tag{05.11}$$

$$\binom{n+1}{m+1} = \binom{n}{m} + \binom{n}{m+1}, \tag{05.12}$$

$$\binom{m+n}{m} m! = \prod_{k \in m]} (n+k) = \frac{(m+n)!}{n!}, \tag{05.13}$$

$$\sum_{k \in (n+1)[} \binom{n}{k} = 2^n. \tag{05.14}$$

If S is a finite set, then the set $\mathrm{Perm}\,S$ of all permutations of S is again finite and

$$\sharp\,\mathrm{Perm}\,S = (\sharp\,S)!. \tag{05.15}$$

If S and T are finite sets, then the set $\mathrm{Inj}\,(S,T)$ of all injections from S to T is again finite and we have $\mathrm{Inj}\,(S,T) = \emptyset$ if $\sharp\,S > \sharp\,T$ while

$$\sharp\,\mathrm{Inj}\,(S,T) = \frac{(\sharp\,T)!}{(\sharp\,T - \sharp\,S)!} \quad \text{if } \sharp\,S \leqq \sharp\,T. \tag{05.16}$$

Notes 05

(1) Many people use $|S|$ for the cardinal $\sharp S$ of a finite set.

(2) The notations $\mathrm{Fin}\,S$ and $\mathrm{Fin}_m S$ are used here for the first time. J. J. Schäffer and I have used $\mathfrak{F}(S)$ and $\mathfrak{F}_m(S)$, respectively, since about 1973.

06 Basic Algebra

A **pre-monoid** is a set M endowed with structure by the prescription of a mapping $\text{cmb} : M \times M \to M$ called **combination**, which satisfies the *associative law*

$$\text{cmb}(\text{cmb}(a, b), c) = \text{cmb}(a, \text{cmb}(b, c)) \quad \text{for all} \quad a, b, c \in M. \tag{06.1}$$

A **monoid** M is a pre-monoid, with combination cmb, endowed with additional structure by the prescription of a **neutral** $n \in M$ which satisfies the *neutrality law*

$$\text{cmb}(a, n) = \text{cmb}(n, a) = a \quad \text{for all} \quad a \in M. \tag{06.2}$$

A **group** G is a monoid, with combination cmb and neutral n, endowed with additional structure by the prescription of a mapping $\text{rev} : G \to G$, called **reversion** , which satisfies the *reversion law*

$$\text{cmb}(a, \text{rev}(a)) = \text{cmb}(\text{rev}(a), a) = n \quad \text{for all} \quad a \in G. \tag{06.3}$$

Let H be a subset of a pre-monoid M. We say that H is a **sub-pre-monoid** if it is stable under combination, i.e. if

$$\text{cmb}_{>}(H \times H) \subset H. \tag{06.4}$$

If this is the case, then the adjustment $\text{cmb}\,|_{H\times H}^{H}$ endows H with the natural structure of a pre-monoid.

Let H be a subset of a monoid M. We say that H is a **submonoid** of M if it is a sub-pre-monoid of M and if it contains the neutral n of M. If this is the case, then the designation of n as the neutral of H endows H with the natural structure of a monoid.

Let H be a subset of a group G. We say that H is a **subgroup** of G if it is a submonoid of G and if H is reversion-invariant, i.e. if

$$\text{rev}_{>}(H) \subset H.$$

If this is the case, then the adjustment $\text{rev}_{|H}$ (see (03.23)) endows H with the natural structure of a group. The singleton $\{n\}$ is a subgroup of G, called the **neutral-subgroup** of G.

Let M be a pre-monoid. Then M contains at most one element n which satisfies the neutrality law (06.2). If such an element exists, we say that M is **monoidable**, because we can use n to endow M with the natural structure of a monoid.

Let M be a monoid. Then there exists at most one mapping $\mathrm{rev} : M \to M$ which satisfies the reversion law (06.3). If such a mapping exists, we say that M is **groupable** because we can use rev to endow M with the natural structure of a group. We say that a pre-monoid M is groupable if it is monoidable and if the resulting monoid is groupable. If G is a group, then every groupable sub-pre-monoid of G is in fact a subgroup.

Pitfall: A monoidable sub-pre-monoid of a monoid need not be a submonoid. For example, the set of natural numbers $\mathbb{N}$ with multiplication as combination and 1 as neutral is a monoid and the singleton $\{0\}$ a sub-pre-monoid of $\mathbb{N}$. Now, $\{0\}$ is monoidable, in fact groupable, but it does not contain 1 and hence is not a submonoid of $\mathbb{N}$. ■

Let E be a set. The set $\mathrm{Map}\,(E, E)$ of all mappings from E to itself becomes a monoid, if we define its combination by composition, i.e. by $\mathrm{cmb}(f, g) := f \circ g$ for all $f, g \in \mathrm{Map}\,(E, E)$, and if we designate the identity 1_E to be its neutral. We call this monoid the **transformation-monoid** of E. The set $\mathrm{Perm}\, E$ of all permutations of E is a groupable submonoid of $\mathrm{Map}\,(E, E)$. It becomes a group if we designate the reversion to be the inversion of mappings, i.e. if $\mathrm{rev}(f) := f^{\leftarrow}$ for all $f \in \mathrm{Perm}\, E$. We call this group $\mathrm{Perm}\, E$ the **permutation-group** of E.

Let G and G' be groups with combinations $\mathrm{cmb}, \mathrm{cmb}'$, neutrals n, n' and reversions rev and rev$'$, respectively. We say that $\tau : G \to G'$ is a **homomorphism** if it preserves combinations, i.e. if

$$\mathrm{cmb}'(\tau(a), \tau(b)) = \tau(\mathrm{cmb}(a, b)) \quad \text{for all } a, b \in G. \tag{06.5}$$

It then also preserves neutrals and reversions, i.e.

$$n' = \tau(n) \quad \text{and} \quad \mathrm{rev}'(\tau(a)) = \tau(\mathrm{rev}(a)) \quad \text{for all } a \in G. \tag{06.6}$$

If τ is invertible, then $\tau^{\leftarrow} : G' \to G$ is again a homomorphism; we then say that τ is a **group-isomorphism**.

Let $\tau : G \to G'$ be a homomorphism. Then the image $\tau_{>}(H)$ under τ of every subgroup H of G is a subgroup of G' and the pre-image $\tau^{<}(H')$ under τ of every subgroup H' of G' is a subgroup of G. The pre-image of the neutral-subgroup $\{n'\}$ of G' is called the **kernel** of τ and is denoted by

$$\mathrm{Ker}\, \tau := \tau^{<}(\{n'\}). \tag{06.7}$$

The homomorphism τ is injective if and only if its kernel is the neutral-subgroup of G, i.e. $\mathrm{Ker}\, \tau = \{n\}$.

We say that a pre-monoid or monoid M is **cancellative** if it satisfies the following *cancellation laws* for all $b, c \in M$:

$$(\text{cmb}(a,b) = \text{cmb}(a,c) \quad \text{for some} \quad a \in M) \implies b = c, \tag{06.8}$$

$$(\text{cmb}(b,a) = \text{cmb}(c,a) \quad \text{for some} \quad a \in M) \implies b = c. \tag{06.9}$$

A group is always cancellative.

One often uses multiplicative terminology and notation when dealing with a pre-monoid, monoid, or group. This means that one uses the term "**multiplication**" for the combination, one writes $ab := \text{cmb}(a,b)$ and calls it the "**product**" of a and b, one calls the neutral "**unity**" and denotes it by 1, and one writes $a^{-1} := \text{rev}(a)$ and calls it the "**reciprocal**" of a. The number sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{P}$, and $\mathbb{C}$ (see Sect.00) are all multiplicative monoids and $\mathbb{Q}^\times$, $\mathbb{R}^\times$, $\mathbb{P}^\times$, and $\mathbb{C}^\times$ are multiplicative groups. If M is a multiplicative pre-monoid and if S and T are subsets of M, we write

$$ST := \text{cmb}_>(S \times T) = \{st \mid s \in S, t \in T\}. \tag{06.10}$$

and call it the **member-wise product** of S and T.

If $t \in M$, we abbreviate $St := S\{t\}$ and $tS := \{t\}S$. If G is a multiplicative group and S a subset of G, we write

$$S^{-1} := \text{rev}_>(S) = \{s^{-1} \mid s \in S\}. \tag{06.11}$$

and call it the **member-wise reciprocal** of S.

A pre-monoid, monoid, or group M is said to be **commutative** if

$$\text{cmb}(a,b) = \text{cmb}(b,a) \quad \text{for all } a, b \in M. \tag{06.12}$$

If this is the case, one often uses additive terminology and notation. This means that one uses the term "**addition**" for the combination, one writes $a + b := \text{cmb}(a,b)$ and calls it the "**sum**" of a and b, one calls the neutral "**zero**" and denotes it by 0, and one writes $-a := \text{rev}(a)$ and calls it the "**opposite**" of a. The abbreviation $a - b := a + (-b)$ is customary. The number sets $\mathbb{N}$ and $\mathbb{P}$ are additive monoids while $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are additive groups. If M is an additive pre-monoid and if S and T are subsets of M, we write

$$S + T := \text{cmb}_>(S \times T) = \{s + t \mid s \in S,\ t \in T\}. \tag{06.13}$$

and call it the **member-wise sum** of S and T. If $t \in M$, we abbreviate

$$S + t := S + \{t\} = \{t\} + S =: t + S. \tag{06.14}$$

If G is an additive group and if S is a subset of G, we write

$$-S := \mathrm{rev}_{>}(S) = \{-s \mid s \in S\}. \tag{06.15}$$

and call it the **member-wise opposite** of S. If S and T are subsets of M, we write

$$T - S := \{t - s \mid t \in T,\ \ s \in S\}. \tag{06.16}$$

and call it the **member-wise difference** of T and S.

A **ring** is a set R endowed with the structure of a commutative group, called the *additive structure* of R and described with additive terminology and notation, and the structure of a monoid, called the *multiplicative structure* of R and described with multiplicative terminology and notation, provided that the following *distributive laws* hold

$$\left.\begin{array}{l}(a+b)c = ac+bc\\ c(a+b) = ca+cb\end{array}\right\}\quad \text{for all } a,b,c \in R. \tag{06.17}$$

A ring then has a zero, denoted by 0, and a unity, denoted by 1. We have $0 \neq 1$ unless R is a singleton. The following rules hold for all $a, b, c, \in R$:

$$\begin{aligned}(a-b)c &= ac-bc, \quad c(a-b) = ca-cb, && (06.18)\\ (-a)b &= a(-b) = -(ab), \quad (-a)(-b) = ab, && (06.19)\\ a0 &= 0a = 0. && (06.20)\end{aligned}$$

A ring R is called a **commutative ring** if its multiplicative monoid-structure is commutative. A ring R is called an **integral ring** if $R^{\times}$ is a submonoid of R with respect to its multiplicative structure. This multiplicative monoid $R^{\times}$ is necessarily cancellative. A ring F is called a **field** if it is commutative and integral and if the multiplicative submonoid $F^{\times}$ of F is groupable. Endowed with its natural group structure, $F^{\times}$ is then called the **multiplicative group** of the field F.

A subset S of a ring R is called **subring** if it is a subgroup of the additive group R and a submonoid of the multiplicative monoid R. If this is the case, then S has the natural structure of a ring. If a ring R is commutative [integral], so is every subring of it. A subring of a field is called a subfield if its natural ring-structure is that of a field.

The set $\mathbb{C}$ of complex numbers is a field, $\mathbb{R}$ is a subfield of $\mathbb{C}$, and $\mathbb{Q}$ is a subfield of $\mathbb{R}$. The set $\mathbb{Z}$ is an (integral) subring of $\mathbb{Q}$. The set $\mathbb{P}$ is a submonoid of $\mathbb{R}$ both with respect to its additive structure and its multiplicative structure, but it is not a subring of $\mathbb{R}$. A similar statement applies to $\mathbb{N}$ and $\mathbb{Z}$ instead of $\mathbb{P}$ and $\mathbb{R}$.

Notes 06

(1) I am introducing the term "pre-monoid" here for the first time. In some textbooks, the term "semigroup" is used for it. However, "semigroup" is often used for what we call "monoid" and "monoid" is sometimes used for what we call "pre-monoid".

(2) In much of the literature, no clear distinction is made between a group and a groupable pre-monoid or between a monoid and a monoidable pre-monoid. Some textbooks refer to an element satisfying the neutrality law (06.2) with the definite article before its uniqueness has been proved; such sloppiness should be avoided.

(3) In most textbooks, monoids and groups are introduced with multiplicative terminology and notation. I believe it is useful to at least start with an impartial terminology and notation. It is for this purpose that I introduce the terms "combination", "neutral", and "reversion".

(4) Some people use the term "unit" or "identity" for what we call "unity". This can lead to confusion because "unit" and "identity" also have other meanings.

(5) Some people use the term "inverse" for what we call "reciprocal" or the awkward "additive inverse" for what we call "opposite". The use of "negative" for "opposite" and the reading of the minus-sign as "negative" are barbarisms that should be stamped out.

(6) The term "abelian" is often used instead of "commutative". The latter term is more descriptive and hence preferable.

(7) Some people use the term "ring" for a slightly different structure: they assume that the multiplicative structure is not that of a monoid but only that of a pre-monoid. I call such a structure a *pre-ring*. Another author calls it "rng" (I assume as the result of a fit of humor).

(8) Most textbooks use the term "integral domain", or even just "domain", for what we call an "integral ring". Often these terms are used only in the commutative case. I see good reason not to use the separate term "domain" for a ring, especially since "domain" has several other meanings.

07 Summations

Let M be a commutative monoid, described with additive notation. Also, let a family $c = (c_i \mid i \in I)$ in M be given. One can define, by recursion, for every finite subset J of I, the sum

$$\sum_J c = \sum (c_i \mid i \in J) = \sum_{i \in J} c_i \tag{07.1}$$

of c over J. This summation in M over finite index sets is characterized by the requirement that

$$\sum_{\emptyset} c = 0 \quad \text{and} \quad \sum_{J} c = c_j + \sum_{J\setminus\{j\}} c. \tag{07.2}$$

for all $J \in \operatorname{Fin}(I)$ and all $j \in J$. If I itself is finite, we write $\sum c := \sum_I c$. If the index set I is arbitrary, we then have $\sum_J c = \sum c|_J$ for all $J \in \operatorname{Fin}(I)$.

If K is a finite set and if $\phi : K \to I$ is an injective mapping, we have

$$\sum_{K} c \circ \phi = \sum_{k\in K} c_{\phi(k)} = \sum_{i\in\phi_>(K)} c_i = \sum_{\phi_>(K)} c. \tag{07.3}$$

If J and K are disjoint finite subsets of I, we have

$$\sum_{J\cup K} c = \sum_{J} c + \sum_{K} c. \tag{07.4}$$

More generally, if J is a finite subset of I and if $\mathcal{P}$ is a partition of J, then

$$\sum_{J} c = \sum_{P\in\mathcal{P}} (\sum_{P} c). \tag{07.5}$$

If $J := \{j\}$ is a singleton or if $J := \{j, k\}$ is a doubleton we have

$$\sum_{\{j\}} c = c_j, \qquad \sum_{\{j,k\}} c = c_j + c_k. \tag{07.6}$$

If $c, d \in M^I$, we have

$$\sum_{i\in J} (c_i + d_i) = (\sum_{i\in J} c_i) + (\sum_{i\in J} d_i) \tag{07.7}$$

for all $J \in \operatorname{Fin} I$. More generally, if $m \in M^{I\times K}$ is a matrix in M we have

$$\sum_{i\in J}\sum_{k\in L} m_{i,k} = \sum_{(i,k)\in J\times L} m_{i,k} = \sum_{k\in L}\sum_{i\in J} m_{i,k} \tag{07.8}$$

for all $J \in \operatorname{Fin} I$ and all $L \in \operatorname{Fin} K$.

The **support** of a family $c = (c_i \mid i \in I)$ in M is defined to be the set of all indices $i \in I$ with $c_i \neq 0$ and is denoted by

$$\operatorname{Supt} c := \{i \in I \mid c_i \neq 0\}. \tag{07.9}$$

We use the notation

$$M^{(I)} := \{c \in M^I \mid \operatorname{Supt} c \text{ is finite}\} \tag{07.10}$$

for the set of all families in M with index set I and finite support. For every $c \in M^{(I)}$ we define

$$\sum_J c := \sum_{(J \cap \operatorname{Supt} c)} c \tag{07.11}$$

for all $J \in \operatorname{Sub} I$. With this definition, the rules (07.1)–(07.5), (07.7), and (07.8) extend to the case when J and L are not necessarily finite, provided that c, d and m have finite support.

If $n \in \mathbb{N}$ and $a \in M$ then $(a \mid i \in n^{]})$ is the constant list of length n and range $\{a\}$ if $n \neq 0$, and the empty list if $n = 0$. We write

$$na := \sum (a \mid i \in n^{]}) = \sum_{i \in n^{]}} a. \tag{07.12}$$

and call it the nth **multiple** of a. If $(a \mid i \in I)$ is any constant finite family with range $\{a\}$ and if $I \neq \emptyset$, we have

$$\sum_{i \in I} a = (\sharp I)a. \tag{07.13}$$

If $(A_i \mid i \in I)$ is a finite family of subsets of M we write

$$\sum_{i \in I} A_i := \{\sum_I a \mid a \in \bigtimes_{i \in I} A_i\} \tag{07.14}$$

and call it the **member-wise sum** of the family.

Let G be a commutative group and let $c = (c_i \mid i \in I)$ be a family in G. We then have

$$\sum_{i \in J} (-c_i) = -\sum_{i \in J} c_i \tag{07.15}$$

if $J \in \operatorname{Sub} I$ is finite or if $c|_J$ has finite support.

If M is a commutative pre-monoid, described with additive notation, and if $c = (c_i \mid i \in I)$ is a family in M, one still can define the sum $\sum_J c$ provided that J is a *non-empty* finite set. This summation is characterized by $(07.6)_1$ and $(07.1)_2$, restricted by the condition that $J \setminus \{j\} \neq \emptyset$. All rules stated before remain valid for summations over non-empty finite index sets.

If M is a commutative monoid described with multiplicative rather than additive notation, the symbol $\sum$ is replaced by $\prod$ and

$$\prod_J c = \prod_{i \in J} c_i = \prod c|_J \tag{07.16}$$

is called the **product** of the family $c = (c_i \mid i \in I)$ in M over $J \in \operatorname{Fin} I$. Of course, 0 must be replaced by 1 wherever it occurs. Also, if $n \in \mathbb{N}$ and $a \in \mathbb{M}$, the n'th multiple na is replaced by the n'th **power** a^n of a.

Let R be commutative ring. If $c = (c_i \mid i \in I)$ is a family in R and if $J \in \operatorname{Fin} I$, it makes sense to form the product $\prod_J c$ as well as the sum $\sum_J c$ of c over J. We list some generalized versions of the distributive law. If $c = (c_i \mid i \in I)$ and $d = (d_k \mid k \in K)$ are families in R then

$$(\sum_J c)(\sum_L d) = \sum_{(i,k)\in J\times L} c_i d_k \tag{07.17}$$

for all $J \in \operatorname{Fin} I$ and $L \in \operatorname{Fin} K$. If $m \in M^{I\times K}$ is a matrix in R then

$$\prod_{i\in J}(\sum_{k\in L} m_{i,k}) = \sum_{s\in L^J}(\prod_{i\in J} m_{i,s_i}) \tag{07.18}$$

for all $J \in \operatorname{Fin} I$ and $L \in \operatorname{Fin} K$. If I is finite and if $c, d \in R^I$, then

$$\prod_{i\in I}(c_i + d_i) = \sum_{J\in\operatorname{Sub} I}(\prod_{j\in J} c_j)(\prod_{k\in I\setminus J} d_k). \tag{07.19}$$

If $a, b \in R$ and $n \in \mathbb{N}$, then

$$(a+b)^n = \sum_{k\in(n+1)^{[}} \binom{n}{k} a^k b^{n-k}. \tag{07.20}$$

If I is finite and $c \in R^I$, then

$$(\sharp I)! \prod_I c = (-1)^{\sharp I} \sum_{K\in\operatorname{Sub} I} (-1)^{\sharp K} (\sum_K c)^{\sharp I}. \tag{07.21}$$

Notes 07

(1) In much of the literature, summations are limited to those over sets of the form $\{i \in \mathbb{Z} \mid n \le i \le m\}$ with $n, m \in \mathbb{Z},\ n \le m$. The notation

$$\sum_{i=n}^{m} c_i := \sum(c_i \,|\, i \in \mathbb{Z},\ n \le i \le m)$$

is often used. Much flexiblity is gained by allowing summations over arbitrary finite index sets.

08 Real Analysis

As mentioned in Sect.06, the set $\mathbb{R}$ of real numbers is a field and the set $\mathbb{P}$ of positive real numbers (which contains 0) is a submonoid of $\mathbb{R}$ both with respect to addition and multiplication. The set $\mathbb{P}^\times = \mathbb{P} \setminus \{0\}$ of strictly positive real numbers is a subgroup of $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ with respect to multiplication and a sub-pre-monoid of $\mathbb{R}$ with respect to addition. The set of all negative [strictly negative] real numbers is $-\mathbb{P}$ $[-\mathbb{P}^\times]$. The collection $\{-\mathbb{P}^\times, \{0\}, \mathbb{P}^\times\}$ is a partition of $\mathbb{R}$. The relation $<$ (read as "strictly less than") on $\mathbb{R}$ is defined by

$$x < y \;:\Longleftrightarrow\; y - x \in \mathbb{P}^\times \tag{08.1}$$

for all $x, y \in \mathbb{R}$. For every $x, y \in \mathbb{R}$, exactly one of the three statements $x < y$, $y < x$, $x = y$ is valid. The relation $\leq$ (read as "less than") on $\mathbb{R}$ is defined by

$$x \leq y \;:\Longleftrightarrow\; y - x \in \mathbb{P}. \tag{08.2}$$

We have $x \leq y \iff (x < y \text{ or } x = y)$ and

$$(x \leq y \text{ and } y \leq x) \quad\Longleftrightarrow\quad x = y. \tag{08.3}$$

The following rules are valid for all x, y, $z \in \mathbb{R}$ and remain valid if $<$ is replaced by $\leq$:

$$(x < y \quad\text{and}\quad y < z) \quad\Longrightarrow\quad x < z, \tag{08.4}$$

$$x < y \quad\Longleftrightarrow\quad x + z < y + z, \tag{08.5}$$

$$x < y \quad\Longleftrightarrow\quad -y < -x, \tag{08.6}$$

$$(x < y \iff ux < uy) \qquad \text{for all } u \in \mathbb{P}^\times. \tag{08.7}$$

The *absolute value* $|x| \in \mathbb{P}$ of a number $x \in \mathbb{R}$ is defined by

$$|x| := \left\{ \begin{array}{rl} x & \text{if} \quad x \in \mathbb{P}^\times \\ 0 & \text{if} \quad x = 0 \\ -x & \text{if} \quad x \in -\mathbb{P}^\times \end{array} \right\}. \tag{08.8}$$

The following rules are valid for all $x, y \in \mathbb{R}$:

$$|x| = 0 \iff x = 0, \tag{08.9}$$

$$|-x| = |x|, \tag{08.10}$$

$$||x| - |y|| \leq |x + y| \leq |x| + |y|, \tag{08.11}$$

$$|xy| = |x||y|. \tag{08.12}$$

The **sign** $\operatorname{sgn} x \in \{-1, 0, 1\}$ of a number $x \in \mathbb{R}$ is defined by

$$\operatorname{sgn} x := \left\{ \begin{array}{rcl} 1 & \text{if} & x \in \mathbb{P}^\times \\ 0 & \text{if} & x = 0 \\ -1 & \text{if} & x \in -\mathbb{P}^\times \end{array} \right\}. \tag{08.13}$$

We have, for all $x, y \in \mathbb{R}$,

$$\operatorname{sgn}(xy) = (\operatorname{sgn} x)(\operatorname{sgn} y), \tag{08.14}$$

$$x = (\operatorname{sgn} x)|x|. \tag{08.15}$$

Let $a, b \in \mathbb{R}$ be given. If $a < b$, we use the following notations:

$$]a, b[:= \{t \in \mathbb{R} \mid a < t < b\}, \tag{08.16}$$

$$[a, b] := \{t \in \mathbb{R} \mid a \leq t \leq b\}, \tag{08.17}$$

$$]a, b] := \{t \in \mathbb{R} \mid a < t \leq b\}, \tag{08.18}$$

$$[a, b[:= \{t \in \mathbb{R} \mid a \leq t < b\}. \tag{08.19}$$

If $a > b$, we sometimes write

$$]a, b[:=]b, a[, \;\; [a, b] := [b, a], \;\; \text{etc.} \tag{08.20}$$

A subset I of $\mathbb{R}$ is called an **interval** if for all $a, b \in I$ with $a < b$ we have $[a, b] \subset I$. The empty set $\emptyset$ and singleton subsets of $\mathbb{R}$ are intervals. All other intevals are infinite sets and will be called **genuine intervals**. The whole of $\mathbb{R}$ is a genuine interval. All other genuine intervals can be classified into eight types. The first four are described by (08.16)–(08.19). The term **open interval** is used if it is of the form (08.16), **closed interval** if it is of the form (08.17), and **half-open interval** if it is of the form (08.18) or (08.19). Intervals of the remaining four types have the form $a + \mathbb{P}$, $a + \mathbb{P}^\times$, $a - \mathbb{P}$, or $a - \mathbb{P}^\times$ for some $a \in \mathbb{R}$; they are called **half-lines**.

Let S be a subset of $\mathbb{R}$. A number $a \in \mathbb{R}$ is called an **upper bound** of S if $S \subset a - \mathbb{P}$ and a **lower bound** of S if $S \subset a + \mathbb{P}$. There is at most one $a \in S$ which is also an upper bound [lower bound] of S. If it exists, it is called the **maximum [minimum]** of S and is denoted by $\max S$ $[\min S]$. If S is non-empty and finite, then both $\max S$ and $\min S$ exist. We say that S is **bounded above [bounded below]** if the set of its upper bounds [lower bounds] is not empty. We say that S is **bounded** if it is both bounded above and below. This is the case if and only if there is a $b \in \mathbb{P}$ such that $|s| \leq b$ for all $s \in S$. Every finite subset of $\mathbb{R}$ is bounded. If S is bounded, then every subset of S is bounded. The union of a finite collection of bounded subsets of $\mathbb{R}$ is again bounded. A genuine interval is bounded if and only if it is of one of the types described by (08.16)–(08.19).

Let S be a non-empty subset of $\mathbb{R}$. If S is bounded above [bounded below] then the set of all its upper bounds [lower bounds] has a minimum [maximum], which is called the **supremum [infimum]** of S and is denoted by $\sup S$ [$\inf S$]. If S has a maximum [minimum] then $\sup S = \max S$ [$\inf S = \min S$]. It is often useful to consider the **extended-real-number set** $\bar{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$ and then to express the assertion that S fails to be bounded above [bounded below] by writing $\sup S = \infty$ [$\inf S = -\infty$]. In this sense $\sup S \in \bar{\mathbb{R}}$ and $\inf S \in \bar{\mathbb{R}}$ are well-defined if S is an arbitrary non-empty subset of $\mathbb{R}$. We also use the notation $\bar{\mathbb{P}} := \mathbb{P} \cup \{\infty\}$. The relation $<$ on $\mathbb{R}$ is extended to $\bar{\mathbb{R}}$ in such a way that $-\infty < \infty$ and that $t \in \bar{\mathbb{R}}$ satisfies $-\infty < t < \infty$ if and only if $t \in \mathbb{R}$.

We say that a family in $\mathbb{R}$, or a function whose codomain is included in $\mathbb{R}$, is **bounded [bounded above, bounded below]** if its range is bounded [bounded above, bounded below]. Let I be a subset of $\mathbb{R}$. We say that the family $a \in \mathbb{R}^I$ is **isotone [antitone]** if

$$i \leq j \implies a_i \leq a_j \ [a_j \leq a_i] \quad \text{for all} \quad i, j \in I. \tag{08.21}$$

We say that the family is **strictly isotone [strictly antitone]** if (08.21) still holds after $\leq$ has been replaced by $<$. This terminology applies, in particular, to lists and sequences (see Sect. 02). We say that a mapping f with $\mathrm{Dom}\, f, \mathrm{Cod}\, f \in \mathrm{Sub}\, \mathbb{R}$ is *isotone, antitone, strictly isotone,* or *strictly antitone* if the family $(f(t) \mid t \in \mathrm{Dom}\, f)$ of its values has the corresponding property.

Let a be a sequence in $\mathbb{R}$, i.e. $a \in \mathbb{R}^{\mathbb{N}^\times}$ or $a \in \mathbb{R}^{\mathbb{N}}$. We say that a **converges** to $t \in \mathbb{R}$ if for every $\varepsilon \in \mathbb{P}^\times$ there is $n \in \mathbb{N}^\times$ such that $a_>(n+\mathbb{N}) \subset\,]t-\varepsilon, t+\varepsilon[$. A sequence in $\mathbb{R}$ can converge to at most one $t \in \mathbb{R}$. If it does, we call t the **limit** of a and write $t = \lim a$ to express the assertion that a converges to t. Every isotone [antitone] sequence that is bounded above [below] converges.

Let a be any sequence. We say that the sequence b is a **subsequence** of a if $b = a \circ \sigma$ for some strictly isotone mapping σ whose domain, and whose codomain, is $\mathbb{N}$ or $\mathbb{N}^\times$, as appropriate. If a is a sequence in $\mathbb{R}$ that converges to $t \in \mathbb{R}$, then every subsequence of a also converges to t.

Let a be a sequence in $\mathbb{R}$. We say that $t \in \mathbb{R}$ is a **cluster point** of a, or that a **clusters** at t, if for every $\varepsilon \in \mathbb{P}^\times$ and every $n \in \mathbb{N}^\times$, we have $a_>(n+\mathbb{N}) \cap\,]t-\varepsilon, t+\varepsilon[\, \neq \emptyset$. The sequence a clusters at t if and only if some subsequence of a converges to t.

Cluster Point Theorem: *Every bounded sequence in $\mathbb{R}$ has at least one cluster point.*

Let a sequence $a \in \mathbb{R}^{\mathbb{N}}$ be given. The **sum-sequence** $\operatorname{ssq} a \in \mathbb{R}^{\mathbb{N}}$ of a is defined by

$$(\operatorname{ssq} a)_n := \sum_{k \in n^{[}} a_k \quad \text{for all} \quad n \in \mathbb{N}. \tag{08.22}$$

Let S be a subset of $\mathbb{R}$. Given $f, g \in \operatorname{Map}(S, \mathbb{R})$, we define the **value-wise sum** $f+g \in \operatorname{Map}(S, \mathbb{R})$ and the **value-wise product** $fg \in \operatorname{Map}(S, \mathbb{R})$ by

$$(f+g)(t) := f(t) + g(t), \;\; (fg)(t) := f(t)g(t) \quad \text{for all} \quad t \in S \tag{08.23}$$

The **value-wise quotient** $\frac{f}{g} \in \operatorname{Map}(g^{<}(\mathbb{R}^{\times}), \mathbb{R})$ is defined by

$$\left(\frac{f}{g}\right)(t) := \frac{f(t)}{g(t)} \quad \text{for all} \quad t \in g^{<}(\mathbb{R}^{\times}) = S \setminus g^{<}(\{0\}). \tag{08.24}$$

We write $f \leq g$ $[f < g]$ to express the assertion that $f(t) \leq g(t)$ $[f(t) < g(t)]$ for all $t \in S$. Given $f \in \operatorname{Map}(S, \mathbb{R})$ and $n \in \mathbb{N}$, we define the **value-wise opposite** $-f$, the **value-wise absolute value** $|f|$ and the **value-wise** n'**th power** f^n of f by

$$(-f)(t) := -f(t), \;\; |f|(t) := |f(t)|, \;\; f^n(t) = (f(t))^n \quad \text{for all } t \in S. \tag{08.25}$$

If $n \in -\mathbb{N}^{\times}$, we define $f^n \in \operatorname{Map}(f^{<}(\mathbb{R}^{\times}), \mathbb{R})$ by $f^n := \frac{1}{f^{|n|}}$.

The identity mapping of $\mathbb{R}$ will be abbreviated by

$$\iota := 1_{\mathbb{R}}. \tag{08.26}$$

The symbol ι will also be used for various adjustments of $1_{\mathbb{R}}$ if the context makes it clear which adjustments are appropriate.

For every $t \in \mathbb{R}$, the sum-sequence $\operatorname{ssq}(\frac{t^k}{k!} \mid k \in \mathbb{N})$ converges; its limit is called the **exponential** of t and is denoted by e^t. The **exponential function** $\exp : \mathbb{R} \to \mathbb{R}$ can be defined by the rule

$$\exp(t) := \mathrm{e}^t \quad \text{for all } t \in \mathbb{R}. \tag{08.27}$$

Let f be a function with $\operatorname{Dom} f, \operatorname{Cod} f \in \operatorname{Sub} \mathbb{R}$. Let $t \in \mathbb{R}$ be an **accumulation point** of $\operatorname{Dom} f$, which means that

$$(]t-\sigma, t+\sigma[\setminus \{t\}) \cap \operatorname{Dom} f \neq \emptyset \quad \text{for all} \quad \sigma \in \mathbb{P}^{\times}. \tag{08.28}$$

We say that $\lambda \in \mathbb{R}$ is a **limit** of f at t if for every $\varepsilon \in \mathbb{P}^\times$, there is $\sigma \in \mathbb{P}^\times$ such that

$$|f|_{]t-\sigma,t+\sigma[\setminus\{t\}} - \lambda \mid \ < \varepsilon. \tag{08.29}$$

The function f can have at most one limit at t. If it has, we write $\lim_t f = \lambda$ to express the assertion that f has the limit λ at t. We say that the function f is **continuous at** t if $t \in \operatorname{Dom} f$ and $\lim_t f = f(t)$.

Let I be a genuine interval. We say that $f : I \to \mathbb{R}$ is **continuous** if f is continuous at *every* $t \in I$. Let $t \in I$ be given. We say that f is **differentiable at** t if the limit

$$\partial_t f := \lim_t \left(\frac{f - f(t)}{\iota - t} \right) \tag{08.30}$$

exists. If it does, then $\partial_t f$ is called the **derivative** of f **at** t. We say that f is **differentiable** if it is differentiable at *every* $t \in I$. If this is the case, we define the **derivative** (-function) $\partial f : I \to \mathbb{R}$ of f by

$$(\partial f)(t) := \partial_t f \quad \text{for all} \quad t \in I. \tag{08.31}$$

If f and g are differentiable, so are $f + g$ and fg, and we have

$$\partial(f+g) = \partial f + \partial g, \qquad \partial(fg) = (\partial f)g + f(\partial g). \tag{08.32}$$

Let $n \in \mathbb{N}$ be given. We say that $f : I \to \mathbb{R}$ is n **times differentiable** if $\partial^n f : I \to \mathbb{R}$ can be defined by the recursion

$$\partial^0 f := f, \quad \partial^{k+1} f := \partial(\partial^k f) \quad \text{for all } k \in n^{[}. \tag{08.33}$$

We say that f is **of class** C^n if it is n times differentiable and if $\partial^n f$ is continuous. We say that f is **of class** C^∞ if it is n times differentiable for *all* $n \in \mathbb{N}$. We also use the notations

$$f^{(k)} := \partial^k f \quad \text{for all} \quad k \in n^{[} \tag{08.34}$$

and

$$f^{\bullet} := \partial f, \quad f^{\bullet\bullet} := \partial^2 f, \quad f^{\bullet\bullet\bullet} := \partial^3 f. \tag{08.35}$$

For each $n \in \mathbb{Z}$, the function ι^n is of class C^∞ and we have

$$\partial^k \iota^n = \left\{ \begin{array}{cl} (\prod_{j \in k^{[}} (n-j))\iota^{n-k} & \text{if} \quad n \notin k^{[} \\ 0 & \text{if} \quad n \in k^{[} \end{array} \right\} \quad \text{for all} \quad k \in \mathbb{N}. \tag{08.36}$$

The exponential function defined by (08.27) is of class C^∞ and we have $\partial^n \exp = \exp$ for all $n \in \mathbb{N}$.

Difference-Quotient Theorem: *Let a genuine interval I and $a, b \in I$ with $a < b$ and a function $f : I \to \mathbb{R}$ be given. If $f|_{[a,b]}$ is continuous and $f|_{]a,b[}$ differentiable, then*

$$\frac{f(b) - f(a)}{b - a} \in \{\partial_t f \mid t \in]a, b[\}. \tag{08.37}$$

Corollary: *If I is a genuine interval and if $f : I \to \mathbb{R}$ is differentiable with $\partial f = 0$, then f must be a constant.*

Let f be a function with $\operatorname{Cod} f \subset \mathbb{R}$. We say that f **attains a maximum [minimum]** (at $x \in \operatorname{Dom} f$) if $\operatorname{Rng} f$ has a maximum [minimum] (and if $f(x) = \max \operatorname{Rng} f$ $[\min \operatorname{Rng} f]$). The term **extremum** is used for "maximum or minimum".

Extremum Theorem: *Let I be an open interval. If $f : I \to \mathbb{R}$ attains an extremum at $t \in I$ and if f is differentiable at t then $\partial_t f = 0$.*

Theorem on Attainment of Extrema: *If I is a closed and bounded interval and if $f : I \to \mathbb{R}$ is continuous, then f attains a maximum and minimum.*

Let I be a genuine interval. Let $f : I \to \mathbb{R}$ be a continuous function. One can define, for every $a, b \in I$ the **integral** $\int_a^b f \in \mathbb{R}$. This integration is characterized by the requirement that

$$\int_a^b f = \int_a^c f + \int_c^b f \quad \text{for all} \quad a, b, c \in I \tag{08.38}$$

and

$$\min f_>([a, b]) \leq \frac{1}{b - a} \int_a^b f \leq \max f_>([a, b]) \quad \text{if} \quad a < b. \tag{08.39}$$

We have

$$\left| \int_a^b f \right| \leq \left| \int_a^b |f| \right| \quad \text{for all} \quad a, b \in I. \tag{08.40}$$

If $\lambda \in \mathbb{R}$, if $f, g \in \operatorname{Map}(I, \mathbb{R})$ are both continuous, and if $a, b \in I$, then

$$\lambda \int_a^b f = \int_a^b (\lambda f), \qquad \int_a^b (f + g) = \int_a^b f + \int_a^b g. \tag{08.41}$$

If $f, g \in \operatorname{Map}(I, \mathbb{R})$ are both continuous and if $a \leq b$, we have

$$f \leq g \implies \int_a^b f \leq \int_a^b g. \tag{08.42}$$

Fundamental Theorem of Calculus: *Let I be a genuine interval, let $a \in I$, and let $f : I \to \mathbb{R}$ be a continuous function. Then the function $g : I \to \mathbb{R}$ defined by*

$$g(t) := \int_a^t f \quad \text{for all} \quad t \in I \tag{08.43}$$

is differentiable and $\partial g = f$.

Let $f : I \to \mathbb{R}$ be a continuous function whose domain I is a genuine interval and let $g : I \to \mathbb{R}$ be an **antiderivative** of f, i.e. a differentiable function such that $\partial g = f$. For every $a, b \in I$, we then have

$$\int_a^b f = g(b) - g(a). \tag{08.44}$$

If no dummy-free symbol for the function f is available, it is customary to write

$$\int_a^b f(s)\mathrm{d}s := \int_a^b f. \tag{08.45}$$

Notes 08

(1) Many people use parentheses in notations such as (08.13) - (08.16) and write for example, $(a, b]$ for what we denote by $]a, b]$. This usage should be avoided because it leads to confusion between the pair (a, b) and an interval.

(2) Many people say "less than" when we say "strictly less than" and use the awkward "less than or equal to" when we say "less than".

(3) There are some disparities in the literature about the usage of "interval". Often the term is restricted to genuine intervals or even to bounded genuine intervals.

(4) Some people use the term "limit point" for "cluster point" and sometimes also for "accumulation point". This usage should be avoided because it can too easily be confused with "limit". There is no agreement in the literature about the terms "cluster point" and "accumulation point". Sometimes "cluster point" is used for what we call "accumulation point" and sometimes "accumulation point" for what we call "cluster point".

(5) In the literature, the term "infinite series" is often used in connection with sum-sequences. There is no agreement on what precisely a "series" should be. Some textbooks contain a "definition" that makes no sense. We avoid the problem by not using the term at all.

(6) The Cluster Point Theorem, or a variation thereof, is very often called the "Bolzano-Weierstrass Theorem".

(7) Most of the literature on real analysis suffers from the lack of a one-symbol notation for $1_{\mathbb{R}}$, as has been noted by several authors in the past. I introduced the use of ι for $1_{\mathbb{R}}$ in 1971 (class notes). Some of my colleagues and I cannot understand any more how we could ever have lived without it (as without a microwave oven).

(8) I introduced the notation $\partial_t f$ for the derivative of f at t a few years ago. Most of the more traditional notations, such as $df(t)/dt$, turn out, upon careful examination, to make no sense. The notations f' and $\dot{f}$ for the derivative-function of f are very common. The former clashes with the use of the prime as a mere distinction symbol (as in Def.1 of Sect.13), the latter makes it difficult to write derivatives of functions with compound symbols, such as $(\exp \circ f + \iota(\sin \circ h))^{\bullet}$.

(9) The Difference-Quotient Theorem is most often called the "Mean-Value Theorem of Differential Calculus". I believe that "Difference-Quotient" requires no explanation while "Mean-Value" does.