

Algebra

Second Edition

Volume 2

P. M. COHN, FRS

University College London

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto · Singapore

Copyright © 1989 by John Wiley & Sons Ltd.

All rights reserved.

No part of this book may be reproduced by any means,
or transmitted, or translated into a machine language
without the written permission of the publisher.

Library of Congress Cataloging-in-Publication Data
(Revised for vol. 2)

Cohn, P. M. (Paul Moritz)

Algebra.

Includes indexes.

Bibliography: v. 2, p.

1. Algebra. I. Title.

QA154.2.C63 1982 512.9 81-21932

ISBN 0 471 10168 0 (v. 1)

ISBN 0 471 10169 9 (pbk.: v. 1)

ISBN 0 471 92234 X (v. 2)

ISBN 0 471 92235 8 (pbk.: v. 2)

British Library Cataloguing in Publication Data

Cohn, P. M. (Paul Moritz)

Algebra.—2nd ed.

Vol. 2

1. Abstract algebra

I. Title

512'.02

ISBN 0 471 92234 X

ISBN 0 471 92235 8

Phototypesetting by Thomson Press (India) Ltd., New Delhi

Printed in Great Britain at The Bath Press, Avon

Contents

Preface to the Second Edition	ix
From the Preface to the First Edition	xi
Conventions on terminology	xiii
Table of interdependence of chapters (Leitfaden)	xv
1 Sets	
1.1 Finite, countable and uncountable sets	1
1.2 Zorn's lemma and well-ordered sets.	8
1.3 Categories	16
1.4 Graphs.	21
Further exercises	27
2 Lattices	
2.1 Definitions; modular and distributive lattices	30
2.2 Chain conditions	38
2.3 Boolean algebras	45
2.4 Möbius functions	54
Further exercises	58
3 Field theory	
3.1 Fields and their extensions	62
3.2 Splitting fields	69
3.3 The algebraic closure of a field	74
3.4 Separability	77
3.5 Automorphisms of field extensions	80
3.6 The fundamental theorem of Galois theory	85
3.7 Roots of unity	91
3.8 Finite fields	97
3.9 Primitive elements; norm and trace	102
3.10 Galois theory of equations	107

3.11 The solution of equations by radicals	113
Further exercises	121

4 Modules

4.1 The category of modules over a ring	124
4.2 Semisimple modules	130
4.3 Matrix rings	135
4.4 Free modules	140
4.5 Projective and injective modules	146
4.6 Duality of finite abelian groups	152
4.7 The tensor product of modules	155
Further exercises	163

5 Rings and algebras

5.1 Algebras: definition and examples	165
5.2 Direct products of rings	170
5.3 The Wedderburn structure theorems	174
5.4 The radical	178
5.5 The tensor product of algebras	183
5.6 The regular representation; norm and trace	187
5.7 Composites of fields	191
Further exercises	195

6 Quadratic forms and ordered fields

6.1 Inner product spaces	197
6.2 Orthogonal sums and diagonalization	200
6.3 The orthogonal group of a space	204
6.4 Witt's cancellation theorem and the Witt group of a field	208
6.5 Ordered fields	212
6.6 The field of real numbers	215
Further exercises	220

7 Representation theory of finite groups

7.1 Basic definitions	221
7.2 The averaging lemma and Maschke's theorem	226
7.3 Orthogonality and completeness	229
7.4 Characters	233
7.5 Complex representations	241
7.6 Representations of the symmetric group	247
7.7 Induced representations	253
7.8 Applications: the theorems of Burnside and Frobenius	258
Further exercises	262

8 Valuation theory		
8.1 Divisibility and valuations	264	
8.2 Absolute values	269	
8.3 The p -adic numbers	280	
8.4 Integral elements	289	
8.5 Extension of valuations	294	
Further exercises	302	
9 Commutative rings		
9.1 Operations on ideals	305	
9.2 Prime ideals and factorization	307	
9.3 Localization	310	
9.4 Noetherian rings	317	
9.5 Dedekind domains	319	
9.6 Modules over Dedekind domains	329	
9.7 Algebraic equations	334	
9.8 The primary decomposition	338	
9.9 Dimension	345	
9.10 The Hilbert Nullstellensatz	350	
Further exercises	353	
10 Coding theory		
10.1 The transmission of information	356	
10.2 Block codes	358	
10.3 Linear codes	361	
10.4 Cyclic codes	369	
10.5 Other codes	374	
Further exercises	378	
11 Languages and automata		
11.1 Monoids and monoid actions	379	
11.2 Languages and grammars	383	
11.3 Automata	387	
11.4 Variable-length codes	395	
11.5 Free algebras and formal power series rings	404	
Further exercises	413	
Bibliography	415	
List of notations	418	
Index	421	

Preface to the Second Edition

Ever since the publication in 1977 of the First Edition of *Algebra* Vol. 2, I have been conscious of the many quite basic parts of algebra that had been omitted or inadequately treated. Besides the traditional topics this also included such parts of applied algebra as coding theory, which not only provides an excellent illustration of the way algebraic methods and results are used, but also in its turn has influenced the development of field theory. So when the chance to prepare a revised edition presented itself, I decided to rewrite Vol. 2 completely, to take account of these additions. The resulting increase in material necessitated a division into two further volumes, making three in all. This arrangement should suit the user, who in the new Vol. 2 will find topics for third-year undergraduate courses, while the projected Vol. 3 will be devoted to postgraduate work.

The present volume includes the whole of four chapters from the old Vol. 2, about half of another four, while the rest are represented by smaller proportions. The introductory chapters, on sets and lattices, omit the Peano axioms, but sections on graphs and categories have been added. This is followed by a chapter on field theory, covering Galois theory as well as the notion of algebraic closure. Next come modules, rings and algebras; here many examples and constructions are given, but the theory is only taken as far as the Wedderburn theorems and the radical. A chapter on quadratic forms and ordered fields deals with the more elementary parts from the old Vol. 2, while the chapters on valuations and commutative rings are included in their entirety. But even where the organization has remained the same, there have been many changes of detail; these consist mainly in amplifying and where possible simplifying the proofs, but also in further examples and illustrations. Some results of intrinsic interest or importance have been added, among them Hensel's lemma, Ramsey's theorem and the continuity criterion for p -adic functions.

In addition there are three new chapters. Ch. 7 on representation theory replaces the old seven-page sketch. It allows a more leisurely pace, as well as a wider range, including the symmetric group and induced representations, with the theorems of Frobenius and Burnside as applications. Ch. 10 is an introduction to block codes; of course only a small selection could be presented, but, it is

hoped, enough to whet the reader's appetite. Finally Ch. 11 deals with algebraic language theory and the related topics of variable-length codes, automata and power series rings. Again it is only possible to take the first steps in the subject, but we go far enough to show how techniques from coding theory are used in the study of free algebras.

Throughout, an effort has been made to base the development on Vol. 1. Definitions and key properties are usually recalled in some detail, but not necessarily on their first occurrence; the reader can easily trace explanations through the index. As before, there are numerous exercises (though no solutions), and some historical references.

A number of colleagues have read and commented on parts of the manuscript: M. P. Drazin, W. A. Hodges, F. C. Piper, M. L. Roberts, B. A. F. Wehrfritz; I should like to thank them for their help. Mark Roberts in particular read the manuscript as well as the proofs, and saved me from a number of errors. I am also indebted to numerous correspondents who have pointed out errors or suggested improvements. It goes without saying that any such correspondence relating to the present volume will always be welcome.

My final thanks go to the staff of John Wiley and Sons, who have, as always, been helpful and efficient in getting the volume published.

University College London
February 1989

P. M. COHN

From the Preface to the First Edition

Vol. 1 of this work was intended to cover the first two years of an undergraduate course, and there is fairly general agreement on the topics to be included. At the third-year level and beyond there is a much wider range of optional topics, and the element of personal choice enters in a greater measure. The present volume includes a variety of algebraic topics that are useful elsewhere, important in their own right, or that fairly quickly lead to interesting results (preferably all three). To a large extent the selection follows the tradition set by v.d. Waerden's classic treatise, but some departures are clearly necessary, the main ones here being homological algebra, lattices and non-commutative Noetherian rings.

The material presented may be grouped roughly into three parts of four chapters each*. The first part deals with basic topics: numbers (natural and transfinite), modular and distributive lattices, tensor products and graded rings (with the Golod–Shafarevich theorem as an application) and an introduction to homological algebra. Here ‘introduction’ means an account which goes far enough to describe Ext and Tor for the category of modules, show their use in group theory and prove the syzygy theorem (*à la* Roganov), but stopping short of Künneth-type theorems or direct and inverse limits. On the other hand, injective and flat modules are studied more fully than is customary in elementary texts, to give the reader a feel for them. The basic notions of category theory are explained as they are needed; this seemed more appropriate than having a separate chapter on categories.

The next part, on fields, has a chapter on Galois theory and one on further field theory, mainly infinite field extensions, but also a little on skew fields. This is followed by real fields (Artin–Schreier theory) and quadratic forms, in an account which develops the Witt ring far enough to exhibit the link with orderings of a field. The third part, on rings, begins with valuations, a formal study of divisibility in fields. A chapter on Artinian rings presents the Jacobson theory (radical and semiprimitivity, density theorem) as well as semiperfect rings and a glimpse of

*As explained in the new Preface, the present volume covers about half the material in the old Vol. 2; the remainder will be included in the projected Vol. 3.

group representations, by way of application. It also includes the basic facts on central simple algebras (the Brauer group) and Hochschild cohomology. The chapter on commutative rings emphasizes the links with algebraic geometry and goes as far as the Nullstellensatz. In a slightly different direction, earlier chapters are put to use in a discussion of Dedekind domains and finitely generated modules over them. The final chapter, on Noetherian and PI-rings, brings Goldie's and Posner's theorems and some of the recent work of Amitsur, Razmyslov, Rowen and M. Artin on rings with polynomial identities.

My aim throughout has been to present the basic techniques of the subject in their simplest form, but to develop them far enough to obtain significant results. Many topics, such as Morita theory, duality, representation theory, were merely touched on. Other large fields, such as group theory, have been excluded entirely; to go beyond what was done in Vol. 1 would have given the book a rather specialized character. Inevitably the choice of topics is influenced by the author's interests but it should not be (and in this case has not been) dictated by them. The book was fun to write (at least for the first five years) and, I hope, will be fun to read.

Bedford College
August 1976

P. M. COHN

Conventions on terminology

We recall that all rings (and monoids) have a unit-element or one, i.e. a neutral element for multiplication, usually denoted by 1. By contrast an algebra (over a coefficient ring) need not have a 1, although it usually does in the present volume. A ring is *trivial* if $1 = 0$; this means that it consists of 0 alone. An element a of a ring R is called a *zerodivisor* if $a \neq 0$ and $ab = 0$ or $ba = 0$ for some $b \neq 0$; a *non-zerodivisor* is an element $a \neq 0$ such that $ab \neq 0$, $ba \neq 0$ for all $b \neq 0$. Thus each element of a ring is either a zerodivisor, a non-zerodivisor or 0, and these three possibilities are mutually exclusive. A non-trivial ring without zerodivisors is called an *integral domain*; this term is not taken to imply commutativity. A ring in which the non-zero elements form a group under multiplication is called a *skew field* or *division ring*; in the commutative case this reduces to a *field*. In any ring R the set of non-zero elements is denoted by R^\times ; this notation is mainly used for integral domains, where R^\times is a monoid. A skew field finite-dimensional over its centre is called a *division algebra*; this use of the term was agreed at the Amitsur Conference, Ramat Gan, 1989. The term ‘algebra’ by itself is *not* taken to imply finite dimensionality.

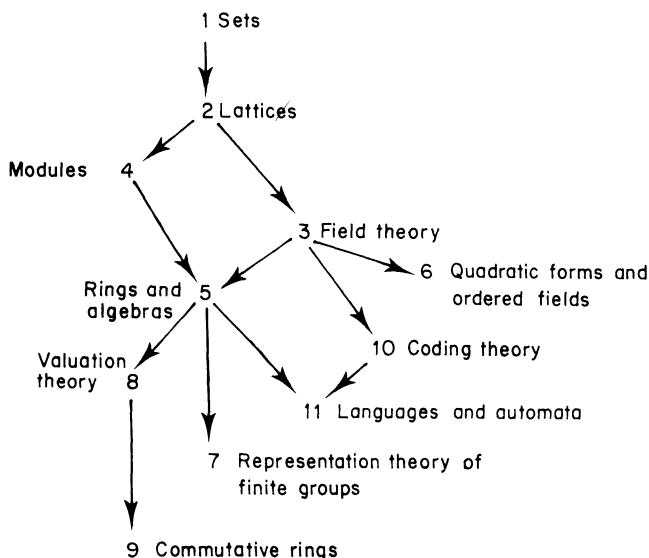
If G is a group and H a subgroup, then the subsets xH of G are called the *right cosets* of H in G , and a subset T of G containing one element from each right coset is called a *left transversal* of H in G ; this is also known as a *complete set of coset representatives* of H in G . We write $H \triangleleft G$ to indicate that H is a normal subgroup of G , and for a subset X of G , $\langle X \rangle$ is the subgroup generated by X ; similarly for submodules.

If S is a set, a property is said to hold for *almost all* members of S if it holds for all but a finite number of members of S . If T is a subset of S , its complement in S is written $S \setminus T$.

As a rule mappings are written on the right; in particular this is done when mappings have to be composed, so that $\alpha\beta$ means: first α , then β . If α is a mapping from a set S and T is a subset of S , then the restriction of α to T is denoted by $\alpha|T$.

References to the bibliography are by name of author and date. All results in a given section are numbered consecutively, e.g. in 4.2 we have Prop. 2.1, Lemma 2.2, Theorem 2.3, which are so referred to in Ch. 4, but as Prop. 4.2.1, etc., elsewhere. As in Vol. 1, we use iff as an abbreviation for ‘if and only if’ and ■ indicates the end (or absence) of a proof.

Table of interdependence of chapters (Leitfaden)



Possible courses:

- (i) 1.2, 2.2, 3.1, 4.1–4, 5.1–4, 5.6.
- (ii) 3.1, 3.3, 4.1–2, 4.6, 5.1–4, 7.1–4.
- (iii) 1.1–2, 2.1–2, 3.1, 4.1, 4.5, 8.1–5, 9.1–6.
- (iv) 3.1, 3.7–8, 10.1–5, 11.1–5.

1

Sets

Much of algebra can be done using only very little set theory; all that is needed is a means of comparing infinite sets, and the axiom of choice in the form of Zorn's lemma. These topics occupy the first two sections. This is followed by an outline of the notions of categories needed later; this adds little to what was said in Vol. 1 but it forms a base on which later chapters and Vol. 3 can build. The chapter ends with an introduction to graph theory. This is an extensive theory and all that can be done here is to present a few basic results which convey the flavour of the topic, some of which will be used later.

1.1 Finite, countable and uncountable sets

The primary purpose of the natural numbers is to count. In essence counting serves to compare the ‘numerousness’ or ‘number of elements’ of two sets. When Man Friday wanted to tell Robinson Crusoe that he had seen a boat with 17 men in it, he did this by exhibiting another 17-element set, and he could do this without being able to count up to 17. Even for a fully numerate person it may be easier to compare two sets rather than to count each; e.g. in a full lecture room a brief glance may suffice to convince us that there are as many people as seats. This suggests that it may be easier to determine when two sets have the same ‘number of elements’ than to find that number. Let us call two sets *equipotent* if there is a bijection between them. This relation of equipotence between sets is an equivalence relation on any given collection of sets*, so in order to compare two sets, we may compare each to a set of natural numbers. A set S is said to be *finite*, of *cardinal* n , if S is equipotent to the set $\{1, 2, \dots, n\}$ consisting of the natural numbers from 1 to n . By convention the empty set, having no elements, is reckoned among the finite sets; its cardinal is 0 and it is denoted by \emptyset .

It is clear that two finite sets are equipotent if they have the same cardinal, and this may be regarded as the basis of counting. It is also true that sets of different finite cardinals are not equipotent. This may seem intuitively obvious; we shall assume it here and defer to Vol. 3 its derivation from the axioms for the natural

*We avoid speaking about the collection of *all* sets, as that would bring us dangerously close to the paradoxes mentioned in Vol. 1, Ch. 1. In any case we shall have no need to do so.

numbers. More generally, we shall assume that for any natural numbers m, n , if there is an injective mapping from $\{1, 2, \dots, m\}$ to $\{1, 2, \dots, n\}$, then $m \leq n$. Let us abbreviate $\{1, 2, \dots, v\}$ by $[v]$, for any $v \in \mathbb{N}$. It follows that if there is a bijection between $[m]$ and $[n]$, then $m \leq n$ and $n \leq m$, hence $m = n$. Thus for any finite set, the natural number which indicates its cardinal is uniquely determined. The contrapositive form of the above assertion states that if $m > n$, then there can be no injective mapping from $[m]$ to $[n]$. A more illuminating way of expressing this observation is Dirichlet's celebrated

BOX PRINCIPLE (Schubfachprinzip) *If $n + 1$ objects are distributed over n boxes, then some box must contain more than one of the objects.*

Although intuitively obvious, this principle is of great use in number theory and elsewhere.

Having given a formal definition of finite sets, we now define a set to be *infinite* if it is not finite. Until relatively recent times the notion of 'infinity' was surrounded by a good deal of mystery and uncertainty, even in mathematics. Thus towards the middle of the 19th century, Bolzano propounded as a paradox the fact that (in modern terms) an infinite set might be equipotent to a proper subset of itself. A closer study reveals that every infinite set has this property, and this has even been taken as the basis of a definition of infinite sets; it certainly no longer seems a paradox. The work of Cantor, Dedekind and others from 1870 onwards has dispelled most of the uncertainties, and though mysteries remain, they will not hamper us in the relatively straightforward use we shall make of the theory.

In order to extend the notion of counting to infinite sets, we associate with every set X , finite or not, an object $|X|$ called its *cardinal* or *cardinal number*, defined in such a way that two sets have the same cardinal iff they are equipotent. Such a definition is possible because, as we have seen, equipotence is an equivalence relation on any collection of sets.

A non-empty finite set has a natural number as its cardinal; the empty set has cardinal 0. All other sets are infinite; their cardinals are said to be *transfinite*, or *infinite*. In particular, the set \mathbb{N} of all natural numbers is infinite; its cardinal is denoted by \aleph_0 . The letter aleph, \aleph , the first of the Hebrew alphabet, is customarily used for infinite cardinal numbers. A set of cardinal \aleph_0 is also said to be *countable* (or *enumerable*); thus A is countable iff there is a bijection from \mathbb{N} to A . If a set A is countable, we can write it in the form

$$A = \{a_1, a_2, a_3, \dots\}, \quad (1)$$

where the a_i are distinct. Such a representation of A is called an *enumeration* of A , and a proof that a set is countable will often consist in giving an enumeration. Sometimes the term 'enumeration' is used for a set written as in (1) even if the a_i are not distinct; in that case we can always produce a strict enumeration by going

through the sequence and omitting all repetitions. The set so obtained is finite or countable.

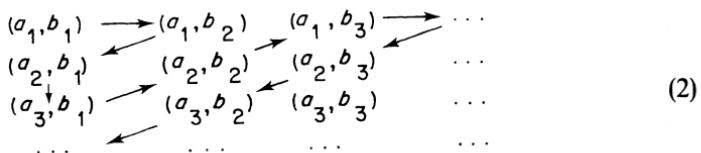
Many sets formed from countable sets are again countable, as our first result shows.

THEOREM 1.1 *Any subset and any quotient set of a countable set is finite or countable. If A and B are countable sets, then the union $A \cup B$ and Cartesian product $A \times B$ are again countable; more generally, the Cartesian product of any finite number of countable sets is countable. Further, a countable union of countable sets is countable and the collection of all finite subsets of a countable set is countable.*

We recall that a *quotient set* of A is the set of all blocks, i.e. equivalence classes, of some equivalence on A.

Proof. Any countable set A may be taken in the form (1); if A' is a subset, we go through the sequence a_1, a_2, \dots of elements of A and omit all terms not in A' to obtain an enumeration of A' . If A'' is a quotient set, and $x \mapsto \bar{x}$ is the natural mapping from A to A'' , then $\{\bar{a}_1, \bar{a}_2, \dots\}$ is an enumeration of A'' , possibly with repetitions; hence A'' is countable (or finite).

Next let A be given by (1) and let $B = \{b_1, b_2, \dots\}$; then $A \cup B$ may be enumerated as $\{a_1, b_1, a_2, b_2, \dots\}$, where repetitions (which may occur if $A \cap B \neq \emptyset$) may be discarded. To prove that $A \times B$ is countable, we arrange its elements as an (infinite) matrix:



which we enumerate by going along successive diagonals as indicated; we have an enumeration because any pair (a_i, b_j) is reached in a finite number of steps. Now the result for a product of r countable sets follows by induction on r. If we have a countable family $\{A_n\}$ of countable sets, say $A_n = \{a_{ni}\}$, then we can enumerate the union $\bigcup A_n = \{a_{ni} | n, i \in \mathbb{N}\}$ by writing the elements a_{ni} as a matrix and using the pattern (2).

Finally let A be any countable set and denote by A_r for $r = 1, 2, \dots$ the set of all r-element subsets of A. Clearly A_r is countable, for it may be mapped into the Cartesian power A^r by the rule

$$\{a_{i_1}, \dots, a_{i_r}\} \mapsto (a_{j_1}, \dots, a_{j_r}),$$

where j_1, \dots, j_r is the sequence i_1, \dots, i_r arranged in ascending order. This provides a bijection of A_r with a subset of A^r , and it follows that A_r is countable. Hence A_r can be enumerated as $A_r = \{c_{1r}, c_{2r}, \dots\}$, and now $c_{ij} \mapsto (i, j)$ provides a bijection

between $A_1 \cup A_2 \cup \dots$ and \mathbb{N}^2 . Thus the collection of all non-empty finite subsets of A is countable, and adding \emptyset as a further member we still have a countable set. ■

With the help of this result many sets can be proved to be countable which do not at first sight appear to be so. Thus the set \mathbf{Z} of all integers can be written as a union of $\mathbf{N} = \{1, 2, 3, \dots\}$ and $\mathbf{N}' = \{0, -1, -2, \dots\}$; both \mathbf{N} and \mathbf{N}' are countable, hence so is \mathbf{Z} . The set \mathbf{Q}_+ of all positive rational numbers is countable, as image of \mathbf{N}^2 under the mapping $(a, b) \mapsto ab^{-1}$. Now \mathbf{Q} itself can be written as the union of the set of positive rational numbers, the negative rational numbers and 0; therefore \mathbf{Q} is countable. The set of all algebraic numbers (cf. Ch. 3 below) is countable: for a given degree n , the set of all monic equations of degree n over \mathbf{Q} is equipotent to \mathbf{Q}^n , if we map

$$(a_1, \dots, a_n) \mapsto x^n + a_1 x^{n-1} + \dots + a_n = 0.$$

Each equation has at most n complex roots, so the set S_n of all roots of equations of degree n is countable, and now the set of all algebraic numbers is $S_1 \cup S_2 \cup \dots$, which is again countable.

At this point a newcomer might be forgiven for thinking that perhaps every infinite set is countable. If that were so, there would of course be no need for an elaborate theory of cardinal numbers. In fact the existence of uncountable sets is one of the key results of Cantor's theory, and we shall soon meet examples of such sets.

Our next task is to extend the natural order on \mathbf{N} to cardinal numbers. If α, β are any cardinals, let A, B be sets such that $|A| = \alpha, |B| = \beta$. We shall write $\alpha \leq \beta$ whenever there is an injective mapping from A to B . Whether such a mapping exists clearly depends only on α, β and not on A, B themselves, so the notation is justified. Further, $\alpha \leq \alpha$ holds for all α , because the identity mapping on A is injective, and since the composition of two injections is an injection, it follows that $\alpha \leq \beta, \beta \leq \gamma$ implies $\alpha \leq \gamma$. Thus we have a preordering; this will in fact turn out to be a total ordering (i.e. for any cardinals α, β , either $\alpha \leq \beta$ or $\beta \leq \alpha$), but for the moment we content ourselves with proving that it is an ordering, i.e. that ' \leq ' is antisymmetric. In terms of sets we must establish

THEOREM 1.2 (Schröder–Bernstein theorem) *Let A, B be any sets and $f: A \rightarrow B, g: B \rightarrow A$ any injective mappings. Then there is a bijection $h: A \rightarrow B$.*

Proof. By alternating applications of f and g we produce an infinite sequence of successive images starting from $a \in A: a, af, ag, agf, \dots$. Further, each element $a \in A$ is the image of at most one element of B under g , which may be written ag^{-1} , and each $b \in B$ is the image of at most one element bf^{-1} of A under f , so from $a \in A$ we obtain a sequence of inverse images which may or may not break off: $ag^{-1}, ag^{-1}f^{-1}, \dots$. If we trace a given element $a \in A$ as far back as possible we find one

of three cases: (i) there is a first ‘ancestor’ in A , i.e. $a_0 \in A \setminus Bg$, such that $a = a_0(fg)^n$ for some $n \geq 0$; (ii) there is a first ancestor in B , i.e. $b_0 \in B \setminus Af$, such that $a = b_0(gf)^n$ for some $n \geq 0$; (iii) the sequence of inverse images continues indefinitely.

Each element of A comes under exactly one of these headings, and likewise each element of B . Thus A is partitioned into three subsets A_1, A_2, A_3 ; similarly B is partitioned into $B_1 = A_1f, B_2 = A_2g^{-1}$ and $B_3 = A_3f = A_3g^{-1}$. It is clear that the restriction of f to A_1 is a bijection between A_1 and B_1 , for each element of B_1 comes from one element in A_1 . For the same reason the restriction of g to B_2 provides a bijection between B_2 and A_2 , and we can use either f restricted to A_3 , or g restricted to B_3 to obtain a bijection between A_3 and B_3 . Thus we have found a bijection between A_i and B_i ($i = 1, 2, 3$) and putting these together we obtain a bijection between A and B . ■

This proof is essentially due to J. König (in 1906).

The sum and product of cardinals may be defined as follows. Let α, β be any cardinals, say $\alpha = |A|, \beta = |B|$, and assume that $A \cap B = \emptyset$. Then $|A \cup B|$ depends only on α, β , not on A, B and we define

$$\alpha + \beta = |A \cup B|.$$

Similarly we put

$$\alpha\beta = |A \times B|.$$

It is easy to verify that these operations satisfy the commutative and associative laws, and a distributive law, as in the case of natural numbers. Moreover, for finite cardinals these operations agree with the usual operations of addition and multiplication. On the other hand, the cancellation law does not hold, thus we may have $\alpha + \beta = \alpha' + \beta$ or $\alpha\beta = \alpha'\beta$ for $\alpha \neq \alpha'$, and there is nothing corresponding to subtraction or division. In fact, it can be shown that if $\alpha, \beta \neq 0$ and at least one of α, β is infinite, then

$$\alpha + \beta = \alpha\beta = \max \{\alpha, \beta\}. \quad (3)$$

For any cardinals α, β we define β^α as $|B^A|$, where A, B are sets such that $|A| = \alpha, |B| = \beta$ and B^A denotes the set of all mappings from A to B . It is again clear that β^α is independent of the choice of A, B , and we note that for finite cardinals, β^α has its usual meaning: if A has m elements and B has n elements, then there is a choice of n elements to which to map each element of A , and these choices are independent, so in all there are $n.n \dots n$ (m factors) choices. Of course this interpretation applies only to finite sets.

If B is a 1-element set, then so is B^A , for any set A : each element of A is mapped to the unique element of B , and this applies even if A is empty, for a mapping $A \rightarrow B$ is defined as soon as we have specified the images of the elements of A ; so when $A = \emptyset$, nothing needs to be done. When B is empty, then so is B^A , unless

also $A = \emptyset$, for there is nowhere for the elements of A to map to. Hence we have

$$1^\alpha = 1, \quad 0^\alpha = \begin{cases} 0 & \text{if } \alpha \neq 0, \\ 1 & \text{if } \alpha = 0. \end{cases} \quad (4)$$

Let us now assume that B has more than one element. Then we necessarily have

$$|B^A| \geq |A|. \quad (5)$$

For let b, b' be distinct elements of B ; we can map A to B^A by the rule $a \mapsto \delta_a$, where

$$x\delta_a = \begin{cases} b & \text{if } x = a, \\ b' & \text{if } x \neq a. \end{cases}$$

This mapping is injective because for $a \neq a'$, δ_a differs from $\delta_{a'}$ at a . It is a remarkable fact that the inequality (5) is always strict. As usual we write $\alpha < \beta$ or $\beta > \alpha$ to mean ' $\alpha \leq \beta$ and $\alpha \neq \beta$ '.

THEOREM 1.3 *For any cardinals α, β , if $\beta > 1$, then $\alpha < \beta^\alpha$. In particular,*

$$\alpha < 2^\alpha \quad (6)$$

for any cardinal α .

Proof. We have just seen that $\alpha \leq \beta^\alpha$ and it only remains to show that equality cannot hold. Taking sets A, B such that $|A| = \alpha, |B| = \beta$, we shall show that there is no surjective mapping from A to B^A ; it then follows that these sets are not equipotent. Thus let $f: A \rightarrow B^A$ be given; in detail, f associates with each $a \in A$ a mapping from A to B , which may be denoted by f_a . We must show that f is not surjective, i.e. we must find $g: A \rightarrow B$ such that $g \neq f_a$ for all $a \in A$. This may be done very simply by constructing a mapping g to differ from f_a at a . By hypothesis, B has at least two elements, say b, b' , where $b \neq b'$. We put

$$ag = \begin{cases} b' & \text{if } af_a = b, \\ b & \text{otherwise.} \end{cases}$$

Then g is well-defined and for each $a \in A$, $g \neq f_a$ because $ag \neq af_a$. ■

If in this theorem we take A to be countable and B a 2-element set, simply denoted by 2, then 2^A is again infinite, but uncountable. Moreover, we can in this way obtain arbitrarily large cardinals by starting from any infinite cardinal α and forming in succession $2^\alpha, 2^{2^\alpha}, \dots$.

Theorem 1.3 again illustrates the dangers of operating with the ‘set of all sets’. If we could form the union of all sets, U say, then U would contain 2^U as a subset, hence $|2^U| \leq |U|$, in contradiction to Th. 1.3. This paradox was discussed by Burali-Forti and others in the closing years of the 19th century, and it provided the impetus for much of the axiomatic development that followed. Any axiomatic system now in use is designed to avoid the possibility of such paradoxes. For our

purposes it is sufficient to note that we can avoid the paradoxes by not admitting constructions involving ‘all sets’ without further qualification.

We conclude with some applications of Th. 1.3. Given any set A , we denote by $\mathcal{P}(A)$ the set whose members are all the subsets of A ; e.g. $\mathcal{P}(\emptyset) = \{\emptyset\}$, $\mathcal{P}(\{x\}) = \{\emptyset, \{x\}\}$. This set $\mathcal{P}(A)$ is often called the *power set* of A ; it is equipotent with 2^A . To obtain a bijection we associate with each subset C of A its *characteristic function* $\chi_C \in 2^A$; taking $2 = \{0, 1\}$, we have

$$\chi_C(x) = \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{if } x \notin C. \end{cases}$$

It is easily seen that the mapping $C \mapsto \chi_C$ provides a bijection between $\mathcal{P}(A)$ and 2^A . The inverse mapping is obtained by associating with each $f \in 2^A$ the inverse image of 1: $f^{-1} = \{x \in A \mid xf = 1\}$. Now Th. 1.3 shows the truth of the following:

COROLLARY 1.4 *No set A is equipotent with its power set. More precisely, given any set A , there is no surjection from A to $\mathcal{P}(A)$.* ■

As a further application we determine the cardinal of the set \mathbf{R} of all real numbers. This cardinal is usually denoted by c and is called the *cardinal (or power) of the continuum*.

PROPOSITION 1.5 $c = 2^{\aleph_0}$.

Proof. Firstly we can replace \mathbf{R} by the open interval $(0, 1) = \{x \in \mathbf{R} \mid 0 < x < 1\}$, for there is a bijection, e.g. $x \mapsto \frac{1}{2}[(1 + x^2)^{1/2} + x](1 + x^2)^{-1/2}$, or $x \mapsto \frac{1}{2}[1 + \sin(\tan^{-1} x)]$. If we express each number in the binary scale: $a = 0.a_1a_2\dots$ ($a_i = 0$ or 1), then $a \mapsto f_a$, where $f_a(n) = a_n$, is a mapping $(0, 1) \rightarrow 2^{\mathbb{N}}$ which is injective, for distinct numbers have distinct binary expansions. Indeed, some have more than one, e.g. $0.0111\dots = 0.1000\dots$, but we can achieve uniqueness by excluding representations in which only finitely many digits are 0. It follows that $c \leq 2^{\aleph_0}$. On the other hand, there is an injective mapping from $2^{\mathbb{N}}$ to $(0, 1)$, obtained by mapping f_a , defined as before, to $0.a_1a_2\dots$ in the decimal scale; thus the image consists of the real numbers between 0 and 1 whose decimal expansion contains only 0's and 1's. This shows that $2^{\aleph_0} \leq c$, and the desired equality follows. ■

It was conjectured by Cantor that c is the least cardinal greater than \aleph_0 ; this is known as *Cantor's continuum hypothesis (CH)*. In 1939 Gödel showed that it is consistent with the usual axioms of set theory; thus if the usual system of axioms (which we have not given explicitly) is consistent, then it remains consistent when CH is added. In 1963 P. J. Cohen showed CH to be independent of the usual axioms of set theory. Thus if the negation of CH is added to the axioms of set theory, we again get a consistent system. This means that within the usual system of set theory CH is undecidable.

Exercises

- (1) Show that the set of all intervals in \mathbf{R} with rational endpoints is countable.
- (2) Let A be an infinite set, A' a finite subset and B its complement in A . By picking a countable subset of B show that $|A| = |B|$ without using eqn. (3).
- (3) Let A be an uncountable set, A' a countable subset and B its complement in A . Show that $|A| = |B|$ without assuming eqn. (3).
- (4) Fill in the details of the following proof that the interval $(0, 1)$ is uncountable: If the real numbers in binary form (as in the proof of Prop. 1.4) could be enumerated $a^{(1)}, a^{(2)}, \dots$, we can find a number not included in the enumeration by putting $a = 0.b_1 b_2 \dots$, where $b_n = 0$ or 1 according as $a^{(n)}$ has 1 or 0 in the n th place. (This is Cantor's diagonal argument: b_n is chosen so as to differ from the diagonal term $a_n^{(n)}$.)
- (5) Show that the set of all real functions in the interval $[0, 1]$ has cardinal greater than the cardinal of the continuum. What about the subset of continuous functions?
- (6) Show that for any cardinals α, β, γ , if $\gamma \neq 0$ and $\alpha \leq \beta$, then $\alpha\gamma \leq \beta\gamma$.
- (7) Show that $\alpha^\gamma\beta^\gamma = (\alpha\beta)^\gamma$, $\alpha^\beta\alpha^\gamma = \alpha^{\beta+\gamma}$, $(\alpha^\beta)^\gamma = \alpha^{\beta\gamma}$.
- (8) Let $f: \mathbf{R} \rightarrow \mathbf{Q}$ be such that $x \leq y$ implies $xf \leq yf$. Show that there is an interval in which f is constant.

1.2 Zorn's lemma and well-ordered sets

In 1.1 we have already defined the relation $\alpha \leq \beta$ for cardinals and we have shown in Th.1.2 that it is a partial ordering. Let us recall that a set S is said to be *partially ordered* if there is a binary relation \leq , called a *partial ordering*, defined on S with the following properties:

- O.1** $x \leq x$ for all $x \in S$ (*reflexivity*),
- O.2** $x \leq y$ and $y \leq z$ imply $x \leq z$ for all $x, y, z \in S$ (*transitivity*),
- O.3** $x \leq y$ and $y \leq x$ imply $x = y$ for all $x, y \in S$ (*antisymmetry*).

If only O.1–2 hold, we speak of a *preordering*.

If ' \leq ' is a partial ordering on a set S , we shall write $x < y$ to mean ' $x \leq y$ and $x \neq y$ ', and we write $x \geq y$, $x > y$ for $y \leq x$, $y < x$ respectively. As is easily verified, the opposite ordering ' \geq ' again satisfies O.1–3 and so is again a partial ordering. Thus any general statement about partially ordered sets has a dual, which is obtained by interpreting the original statement for the oppositely ordered set. This principle of duality can often be used to shorten proofs.

Two elements x, y in a partially ordered set S are said to be *comparable* if $x \leq y$ or $y \leq x$. A subset of S in which any two elements are comparable is called a *chain*

or also *totally ordered*. A subset in which no two elements are comparable is called an *anti-chain*. For example, the set of natural numbers \mathbb{N} is totally ordered for the usual ordering by magnitude, and partially ordered with respect to divisibility: $a|b$ iff $b = ac$ for some $c \in \mathbb{N}$. For the divisibility ordering on \mathbb{N} the set of all prime numbers is an anti-chain.

In any partially ordered set S an element c is a *greatest element* if $x \leq c$ for all $x \in S$, while c is *maximal* if $c < x$ for no $x \in S$. Thus a greatest element is maximal but the converse need not hold. A greatest element, if it exists, is clearly unique, unlike a maximal element. *Least* and *minimal* elements are defined dually; e.g. \mathbb{N} with its usual ordering has a least element, but no greatest element, while \mathbb{Q} has neither a least nor a greatest element.

An *upper bound* of a subset X of S is an element $b \in S$ such that $x \leq b$ for all $x \in X$; here b may or may not belong to X . *Lower bounds* are defined dually, and a set is *bounded* if it has both an upper and a lower bound.

We now take up the question of the comparability of cardinals left open in the last section, i.e. whether the ordering of cardinals is in fact total. In terms of sets the question is whether, given two sets A, B , we can find an injective mapping from one of them to the other. In intuitive terms one might try to answer this question by choosing an element from each of A and B , say a_1, b_1 , and pairing them off, then choosing another pair of elements $a_2 \in A, b_2 \in B$ and pairing them off, and so on. For sets that are at most countable this solves the problem, but we have seen that there are uncountable sets, and here the procedure adopted is rather more problematic. One way to overcome the difficulty is to introduce the concept of a *well-ordering*:

An ordered set A is said to be *well-ordered* if every non-empty subset of A has a least element.

A well-ordered set is always totally ordered, as we see by applying the definition to 2-element subsets. It is also clear from the definition that any subset of a well-ordered set is again well-ordered. A countable set may be well-ordered simply by enumerating its elements, e.g. the natural order of the positive integers is a well-ordering, but there are many other well-orderings which do not put the countability into evidence, e.g. $\{2, 3, 4, \dots, 1\}$, where the order intended is that in which the numbers are written, or $\{1, 3, 5, \dots, 2, 4, 6, \dots\}$ or even $\{1, 2, 4, 6, \dots, 3, 9, 15, 21, \dots, 5, 25, 35, \dots, 7, 49, \dots\}$. By contrast, the negative numbers in their natural order form a set which is not well-ordered, although we can well-order it, e.g. by writing it in the opposite order.

For well-ordered sets it is possible to prove the comparability in a strong form. Let us call a subset A' of an ordered set A a *lower segment* if $u \in A', v \leq u$ implies $v \in A'$. This definition can be used for any ordered set, not necessarily well-ordered, or even totally ordered. In particular, for any element $a \in A$, the set $|a| = \{x \in A | x < a\}$ is a lower segment in A ; in a well-ordered set A every lower segment not the whole of A is of this form, for if A' is a proper lower segment of A and a is the first element of $A \setminus A'$, then $A' = |a|$.

Two ordered sets A, B are said to be *order-isomorphic* or of the same *order-type* if there is a bijection between them which preserves the ordering, $f:A \rightarrow B$ such that $x \leq y \Leftrightarrow xf \leq yf$.

LEMMA 2.1 *A well-ordered set cannot be order-isomorphic to one of its proper lower segments.*

Proof. Let A be well-ordered, $|a\rangle$ a proper lower segment and let $f:A \rightarrow |a\rangle$ be an order-isomorphism. Then clearly $af < a$; let a_0 be the least element of A such that $a_0f < a_0$. If we apply f to this inequality and recall that f preserves the order, we find that $a_0ff < a_0f$, so we have found an earlier element with the same property, namely a_0f . This contradiction shows that f cannot exist. ■

We now show that any two well-ordered sets can be compared.

THEOREM 2.2 *Let A, B be two well-ordered sets. Then one of them is order-isomorphic to a lower segment of the other.*

Proof. Let us call a pair of elements $a \in A$ and $b \in B$ *matched* if the corresponding lower segments $|a\rangle$ and $|b\rangle$ are order-isomorphic. Two distinct lower segments of A cannot be order-isomorphic, for one of them will be a lower segment of the other, and this would contradict Lemma 2.1. It follows that any element of B can be matched against at most one element of A and vice versa. Let A' be the set of elements of A that can be matched against elements of B , and B' the set of elements of B matched against elements of A . Then A' and B' are order-isomorphic, as we see by using the correspondence provided by the matching. Moreover, A' is a lower segment of A , for if $a \in A'$ and $a_1 < a$, let a be matched to $b \in B$; then a_1 is matched to the element of $|b\rangle$ which corresponds to it under the isomorphism between $|a\rangle$ and $|b\rangle$. Similarly B' is a lower segment of B . If $A' \neq A$, then $A' = |a'\rangle$ for some $a' \in A$; likewise, if $B' \neq B$, then $B' = |b'\rangle$ for some $b' \in B$, and by construction there is an order-isomorphism between $|a'\rangle$ and $|b'\rangle$, so that a' and b' are matched. But this is a contradiction, because $a' \notin |a'\rangle$; therefore we have either $A' = A$ or $B' = B$ (or both) and the conclusion follows. ■

The problem of comparing cardinals is thus reduced to the problem of well-ordering sets. If every set could be well-ordered, Th.2.2 would tell us that any two cardinals are comparable. Now it was proved by Zermelo (in 1908) that every set can be well-ordered, but he had to make an assumption which was somewhat less intuitive than the other axioms used; this is the following:

AXIOM OF CHOICE *Given a family of non-empty sets $\{A_i\}_{i \in I}$ there exists a function which associates with each set A_i a member of A_i .*

At first sight this is an innocent-sounding assumption, which acquires its force from the fact that it applies to collections of sets with arbitrary indexing set; only for finite families $\{A_1, \dots, A_n\}$ is no axiom needed. The axiom may be illustrated by the following example due to Bertrand Russell. A certain millionaire has infinitely many pairs of shoes and infinitely many pairs of socks. He wants to pick out one shoe from each pair: this causes no problems; he simply picks the left shoe each time. But when he wants to pick a sock from each pair, he needs the axiom of choice.

In some respects the axiom of choice occupies a position analogous to the parallel axiom in geometry (although set theory without the axiom of choice is not as interesting as non-Euclidean geometry). Like the continuum hypothesis it has been proved consistent with and independent of the other axioms of set theory (by K. Gödel in 1939 and P. J. Cohen in 1963 respectively).

A logical step at this point would be to prove that every set can be well-ordered, using the axiom of choice. In fact we shall introduce another axiom, known as *Zorn's lemma*, which is equivalent to the axiom of choice, and use it to prove the well-ordering theorem. This seems more appropriate in the present context, for it is Zorn's lemma rather than the axiom of choice that is used in algebra; we shall meet many examples later on.

ZORN'S LEMMA *Let A be a partially ordered set. If every chain in A has an upper bound, then A has a maximal element.*

A partially ordered set is said to be *inductive* if every chain in it has an upper bound. In particular, such a set must be non-empty, as we see by taking an upper bound of the empty chain. In this terminology Zorn's lemma states that every partially ordered set which is inductive has a maximal element.

This statement sounds plausible, but any attempt at a direct proof soon encounters the situation typical of the axiom of choice. The actual derivation of Zorn's lemma from the axiom of choice can be found in most books on set theory. For an excellent account we refer to Kaplansky (1972). Below, in Th. 2.3, we shall prove the well-ordering theorem (W) on the basis of Zorn's lemma (Z), and it is easy to prove the axiom of choice (C) from the well-ordering theorem: if $\{X_i\}$ is any family of non-empty sets, well-order $\bigcup X_i$ and assign to each X_i the element of it which comes first in the well-ordering. Thus $C \Rightarrow Z$, $Z \Rightarrow W$, $W \Rightarrow C$; so the three assertions, C, Z, W are all equivalent.

Later on we shall meet many situations where the hypotheses of Zorn's lemma are satisfied; for the moment we shall give an illustration where the hypotheses do not hold. Let A be an infinite set and let \mathcal{F} be the collection of all its finite subsets, partially ordered by inclusion. It is clear that \mathcal{F} has no maximal element, and by Zorn's lemma this means that \mathcal{F} must contain chains which have no upper bound in \mathcal{F} ; such chains are of course easily found. In verifying the hypotheses of Zorn's

lemma it is important to test arbitrary chains and not merely ascending sequences, as is shown by examples (cf. Ex. (3)).

THEOREM 2.3 *Every set can be well-ordered.*

Proof. The idea of the proof is to consider well-orderings of parts of the given set, make these well-orderings into a partially ordered set and show it to be inductive, so that Zorn's lemma can be applied.

Given a set A , let \mathcal{W} be the collection of all subsets of A that can be well-ordered; if a subset can be well-ordered in more than one way we list all the versions separately. For example, any finite subset of A can be well-ordered (usually in more than one way); this shows that \mathcal{W} is not empty (even if $A = \emptyset$, \mathcal{W} contains the set \emptyset as member). We order \mathcal{W} by writing $X \leq Y$ for $X, Y \in \mathcal{W}$ whenever X is a subset of Y and the inclusion mapping from X to Y is an order-isomorphism of X with a lower segment of Y ; in particular, the ordering of X is then the same as that induced by Y . It is clear that this defines a partial ordering on \mathcal{W} and we have to show that \mathcal{W} is inductive. Let $\{X_\lambda\}$ be a chain in \mathcal{W} , where λ runs over an indexing set (not necessarily countable); thus for any λ, μ either X_λ is a lower segment of X_μ or X_μ is a lower segment of X_λ . To get an upper bound for this chain we put $X = \bigcup X_\lambda$ and define an ordering on X as follows: let $x, y \in X$ and choose an index λ such that $x, y \in X_\lambda$. If $x \leq y$ in the ordering of X_λ , then the same is true in the ordering of X_μ for any μ such that $x, y \in X_\mu$; for of X_λ, X_μ , one is a lower segment of the other, with the same ordering. Thus we may without ambiguity put $x \leq y$ in X if this holds in some X_λ , and the relation on X so defined is easily seen to be a well-ordering of X , with each X_λ as a lower segment. Hence X is an upper bound of the given chain in \mathcal{W} , and this shows \mathcal{W} to be inductive.

We can now apply Zorn's lemma and obtain a maximal element X' in \mathcal{W} . We claim that $X' = A$; for if not, then there exists $z \in A \setminus X'$. We form $X'' = X' \cup \{z\}$ into a well-ordered set by taking the given order on X' and letting z follow all of X' . Then X'' is a member of \mathcal{W} which is strictly greater than X' , contradicting the maximality of X' . Hence $X' = A$ and this is the desired well-ordering of A . ■

This result allows us to conclude that any two cardinals can be compared; thus the relation ' \leq ' is a total ordering of cardinals. But we can say rather more than this. Th. 2.2 and 2.3 suggest a classification of well-ordered sets according to their order-type. Thus with every well-ordered set A we associate a symbol α , called its *ordinal number* or *order-type*, or simply *ordinal*, such that two well-ordered sets have the same ordinal precisely when they are order-isomorphic. Further, we can define a relation $\alpha \leq \beta$ between ordinals, whenever a set of type α is order-isomorphic to a lower segment of a set of type β . This is a well-defined relation on any set of ordinals, clearly transitive and by Lemma 2.1 antisymmetric, i.e. an ordering, which is total by Th. 2.2. In fact it is a well-ordering (cf. Ex. (7)).

With each ordinal number α a cardinal number $|\alpha|$ may be associated, namely

the cardinal of a well-ordered set of ordinal α . This is an order-preserving mapping from ordinals to cardinals, but not injective: to each finite cardinal there corresponds just one ordinal of the same type, but an infinite cardinal always corresponds to many different ordinals. However, we obtain a well-defined mapping by assigning to each cardinal the *least* ordinal which corresponds to it. For example, a countable set, \mathbb{N} say, may be well-ordered as $1, 2, 3, \dots$; this order-type is denoted by ω and is the least countable ordinal. Another ordering, not isomorphic to the first, is $2, 3, 4, \dots, 1$; it is denoted by $\omega + 1$. Similarly, $n + 1, n + 2, \dots, 1, 2, \dots, n$ has ordinal $\omega + n$. The type of $1, 3, 5, \dots, 2, 4, 6, \dots$ is written $\omega + \omega$, and generally, given ordinals α, β , we define $\alpha + \beta$ as the type of a well-ordered set of type α followed by one of type β . It is easily checked that such an arrangement gives rise to a well-ordered set whose type depends only on α and β . We observe that the addition of ordinal numbers is still associative, but no longer commutative: $1 + \omega = \omega \neq \omega + 1$.

We shall write 2ω for $\omega + \omega$ and generally $n\omega$ for $\omega + \omega + \dots + \omega$ to n terms. The limit of the sequence $\omega, 2\omega, 3\omega, \dots$, i.e. the first ordinal following all of them, is written ω^2 . We shall not pursue this topic further except to mention that every ordinal number α can be written in just one way as

$$\alpha = a_1\omega^{\alpha_1} + a_2\omega^{\alpha_2} + \dots + a_r\omega^{\alpha_r},$$

where r, a_1, \dots, a_r are natural numbers and $\alpha_1, \dots, \alpha_r$ is a decreasing sequence of ordinal numbers (for a proof see e.g. Sierpiński (1956)).

There is a particular situation allowing Zorn's lemma to be applied which frequently occurs in algebra. Let S be a set and P a property of certain subsets of S ; by a P -set we shall understand a subset with the property P . A property P of subsets of S is said to be of *finite character* if any subset T of S is a P -set precisely when all finite subsets of T are P -sets. For example, if S is a partially ordered set, then 'being totally ordered' is a property of finite character: a subset T of S is totally ordered iff every 2-element subset of T is totally ordered. On the other hand, being well-ordered is not a property of finite character in ordered sets, because every finite subset of a totally ordered set is well-ordered, but the set itself need not be well-ordered.

For a property of finite character there is always a maximal subset with this property:

PROPOSITION 2.4 *Let S be a set and P a property of subsets of S . If P is a property of finite character, then the collection of all P -sets in S has a maximal member.*

Proof. The result will follow by Zorn's lemma, if we can show that the set \mathcal{F} of all P -sets in S is inductive. Let $\{T_\alpha\}$ be a chain of P -sets and write $T = \bigcup T_\alpha$. If T fails to have the property P , then there is a finite subset $\{x_1, \dots, x_r\}$ of T which does not have P . Let $x_i \in T_{\alpha_i}$; since the T_α form a chain, there is a largest among the

sets $T_{\alpha_1}, \dots, T_{\alpha_r}$, say T' . But then $T' \supseteq \{x_1, \dots, x_r\}$, so T' does not have P , which is a contradiction. This shows \mathcal{F} to be inductive, and by Zorn's lemma it has a maximal member, and this is the desired maximal P -set. ■

For well-ordered sets there is a form of induction known as the *principle of transfinite induction*. This is embodied in the remark (which practically reproduces the definition) that any non-empty subset of a well-ordered set has a least element. Let us pause briefly to examine how a transfinite induction proof looks in practice. One can distinguish three kinds of ordinals: an ordinal number β may have an immediate predecessor α , so that $\beta = \alpha + 1$, or it may have no immediate predecessor. In that case it is either the first ordinal 1, or it is the first ordinal after an infinite set of ordinals, in which case it is called a *limit ordinal*. Thus our three classes are (i) 1, (ii) ordinals which have an immediate predecessor, and (iii) limit ordinals.

Now let A be a well-ordered set; if its ordinal is τ , the set may be indexed by the ordinals less than τ : $A = \{a_\alpha\}_{\alpha < \tau}$. Suppose that a subset X of A satisfies the following conditions: (i) $a_1 \in X$; (ii) if $a_\alpha \in X$, then $a_{\alpha+1} \in X$; and (iii) if λ is a limit ordinal less than τ , and $a_\alpha \in X$ for all $\alpha < \lambda$, then $a_\lambda \in X$.

Then $X = A$. For if X were a proper subset of A , let a_β be the least element of $A \setminus X$. Then $\beta > 1$, by (i); if β has an immediate predecessor α , then $a_\alpha \in X$ and $\beta = \alpha + 1$, hence by (ii), $x_\beta \in X$, a contradiction. So β must be a limit ordinal and by definition, $x_\alpha \in X$ for all $\alpha < \beta$. Hence by (iii), $a_\beta \in X$, again a contradiction. This proves that $X = A$, as asserted.

This analysis allows us to give an explicit description of well-ordered sets.

PROPOSITION 2.5 *Any well-ordered set A consists of a well-ordered set of countable sequences, possibly followed by a finite sequence.*

Proof. Consider the set L of limit ordinals of A , together with the first element. This is a well-ordered set, and each $\lambda \in L$ which does not come last in L is the first of a countable sequence. If L has no last element, then A consists of a family of countable sequences indexed by L . If L has a last element, then this is the first of a countable or finite sequence. Thus in either case A has the required form. ■

We conclude this section with a proof of a special case of a formula mentioned earlier, eqn. (3) of 1.1.

PROPOSITION 2.6 *For any infinite cardinal a ,*

$$\aleph_0 a = a. \tag{1}$$

Proof. We shall need the associative law of multiplication of cardinal numbers: $(ab)c = a(bc)$. This is easily proved by observing that each side may be regarded as

the cardinal number of the product set $A \times B \times C$, where A, B, C are sets of cardinals a, b, c respectively.

In the case where $a = \aleph_0$, (1) states that \mathbf{N}^2 is equipotent with \mathbf{N} , and this was proved in Th. 1.1. Secondly if a is of the form $\aleph_0 b$, for some cardinal b , then by the associative law,

$$\aleph_0 a = \aleph_0(\aleph_0 b) = \aleph_0^2 b = \aleph_0 b = a,$$

which proves (1) in this case. We complete the proof by showing that every infinite cardinal is of the form $\aleph_0 b$. This amounts to showing that every infinite set A is equipotent with a set of the form $\mathbf{N} \times B$, for a suitable set B .

Let A be an infinite set. By Th. 2.3 A can be well-ordered, and by Prop. 2.5, A consists of a well-ordered set of countable sequences, possibly followed by a finite sequence. Since A is infinite, at least one infinite sequence occurs, and we may rearrange A by taking the finite sequence from the end and putting it in front of the first sequence. The set A now consists entirely of countable sequences, i.e. well-ordered sequences of type ω . If they are indexed by a set B , it follows that A is equipotent with $\mathbf{N} \times B$, and this is what we had to show. ■

Exercises

- (1) Show that the axiom of choice is equivalent to the following axiom: Every surjective mapping has a left inverse.
- (2) Let Φ be a partial ordering relation on a set A . Show that there is a total order Φ' on A such that $\Phi \subseteq \Phi'$.
- (3) Let A be an uncountable set and \mathcal{F} the collection of all its countable subsets. Show that every ascending (countable) sequence of members of \mathcal{F} has an upper bound in \mathcal{F} , but that \mathcal{F} has no maximal element.
- (4) Show that any totally ordered set X has a well-ordered subset Y (where the ordering of Y is that induced by X) with the property: For each $x \in X$ there exists $y \in Y$ such that $x \leq y$.
- (5) Determine all order-automorphisms (i.e. order-preserving permutations) of \mathbf{Z} .
- (6) Find a well-ordering of the set \mathbf{Z} of all integers. Find well-orderings of \mathbf{N} of type 2ω , $2\omega + 1$, $\omega^2 + 1$.
- (7) Show that the set of all lower segments of a well-ordered set is well-ordered by inclusion. Deduce that any set of ordinals is well-ordered by \leq .
- (8) Check that $\alpha + \beta$ is well-defined and satisfies the associative law.
- (9) Ordinal multiplication may be defined by taking, for any ordinals α, β , sets A, B of type α, β respectively, and denoting by $\alpha\beta$ the ordinal of the product $A \times B$, ordered

lexicographically. Show that this multiplication is well-defined, and that it agrees with the following recursive definition: (i) $1\beta = \beta$; (ii) $(\alpha + 1)\beta = \alpha\beta + \beta$; (iii) if λ is a limit ordinal, then $\lambda\beta = \sup \{\gamma\beta \mid \gamma < \lambda\}$.

- (10) Show that $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$, $\alpha(\beta\gamma) = (\alpha\beta)\gamma$, but that in general, $\alpha(\beta + \gamma) \neq \alpha\beta + \alpha\gamma$. (*Hint.* For the inequality take β, γ finite and α infinite.)
- (11) From the recursive definition in Ex. (9) show that if $\alpha < \beta$, $\gamma > 0$, then $\alpha\gamma < \beta\gamma$. Deduce that $\gamma \neq 0$ and $\alpha\gamma = \beta\gamma$ implies $\alpha = \beta$; give examples to show that $\gamma\alpha = \gamma\beta$ does not imply $\alpha = \beta$.
- (12) Show that if β is a limit ordinal, then so is $\alpha + \beta$, for any ordinal α . Is $\beta + \alpha$ necessarily a limit ordinal? If $\alpha > 0$ and β is a limit ordinal, show that $\alpha\beta$ is a limit ordinal.

1.3 Categories

The notions of category and functor were introduced in **4.9** of Vol. 1, but in view of their importance we recall their definitions in some detail and describe some of their simpler properties. A *category* \mathcal{A} consists of a class of *objects* and a class of *morphisms* or *maps*. With each morphism α two objects are associated, its *source* and *target*; if they are X, Y respectively, we write $\alpha: X \rightarrow Y$ or $X \xrightarrow{\alpha} Y$ and say: α goes from X to Y . The collection of all morphisms from X to Y is written $\text{Hom}_{\mathcal{A}}(X, Y)$ or $\mathcal{A}(X, Y)$. Further, we can combine certain morphisms. Given $\alpha: X \rightarrow Y$, $\beta: Y \rightarrow Z$, so that the target of α is the source of β , we can compose α and β to a morphism from X to Z , denoted by $\alpha\beta$. These objects and morphisms are subject to the following rules:

- C.1 *$\mathcal{A}(X, Y)$ is a set and $\mathcal{A}(X, Y) \cap \mathcal{A}(X', Y') = \emptyset$ unless $X = X'$, $Y = Y'$.*
- C.2 *If $\alpha: X \rightarrow Y$, $\beta: Y \rightarrow Z$, $\gamma: Z \rightarrow T$, so that $(\alpha\beta)\gamma$ and $\alpha(\beta\gamma)$ are both defined, then $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.*
- C.3 *For each object X there exists a morphism $1_X: X \rightarrow X$ such that for each $\alpha: X \rightarrow Y$, we have $1_X\alpha = \alpha 1_Y = \alpha$.*

It is easily seen that 1_X is uniquely determined by these properties; it is called the *identity morphism* for X . Any morphism with a two-sided inverse is an *isomorphism*; two objects are *isomorphic* if there is an isomorphism between them.

Obvious examples of categories are *Ens*, the category of sets and mappings; more precisely, the morphisms in this case are triples (α, X, Y) , where the source and target are named, to distinguish (α, X, Y) from (α, X, Y') , where Y' is a subset of Y containing the image of X under α . Similarly we have *Gp*, the category of groups and homomorphisms; *Top*, the category of topological spaces and continuous mappings; *Rg*, the category of rings and homomorphisms. For each ring R we have a category \mathcal{M}_R whose objects are all right R -modules, while the

morphisms are all R -homomorphisms; the category ${}_R\mathcal{M}$ of left R -modules is defined correspondingly.

As we saw in Vol. 1 (p. 9), we cannot speak of the set of all sets without rapidly reaching contradictions; the simplest way out of this dilemma is to refer to the class of all sets, and to keep a distinction between classes and sets. A set may be thought of as a 'small' class; in this sense a category is said to be *small* if the class of its objects is a set. The categories listed above, Ens, Gp, Top, Rg, \mathcal{M}_R , ${}_R\mathcal{M}$, are not small. But any group (more generally, any monoid) can be regarded as a category with a single object, with multiplication consequently everywhere defined; this provides an example of a small category.

Given a category \mathcal{A} , a *subcategory* \mathcal{B} is a collection of objects and morphisms of \mathcal{A} which forms a category with respect to the composition in \mathcal{A} . Thus $\mathcal{B}(X, Y) \subseteq \mathcal{A}(X, Y)$ for any \mathcal{B} -objects X, Y ; if equality holds here, \mathcal{B} is called a *full* subcategory. Thus a full subcategory is determined once we have specified the objects.

From every category \mathcal{A} we obtain another category \mathcal{A}° , called its *opposite*, by reversing all the arrows. Thus \mathcal{A}° has the same objects as \mathcal{A} and to every \mathcal{A} -morphism $\alpha: X \rightarrow Y$ there corresponds a morphism $\alpha^\circ: Y \rightarrow X$ in \mathcal{A}° , with multiplication $(\alpha\beta)^\circ = \beta^\circ\alpha^\circ$, whenever both sides are defined.

A *functor* F from one category \mathcal{A} to another, \mathcal{B} , is a function F which assigns to each \mathcal{A} -object X a \mathcal{B} -object X^F and to each \mathcal{A} -morphism $\alpha: X \rightarrow X'$ a \mathcal{B} -morphism $\alpha^F: X^F \rightarrow X'^F$ such that

F.1 If $\alpha\beta$ is defined in \mathcal{A} , then $\alpha^F \cdot \beta^F$ is defined in \mathcal{B} and

$$\alpha^F \cdot \beta^F = (\alpha\beta)^F;$$

F.2 $1_X^F = 1_{X^F}$, for each \mathcal{A} -object X .

Thus a functor may be described succinctly as a homomorphism of categories. More precisely, the functor F defined above is called *covariant*; by a *contravariant* functor one understands a functor G from \mathcal{A} to \mathcal{B} which assigns to each \mathcal{A} -object X a \mathcal{B} -object X^G and to each \mathcal{A} -morphism $\alpha: X \rightarrow X'$ a \mathcal{B} -morphism $\alpha^G: X'^G \rightarrow X^G$ (note the reversed order) such that F.2 holds, while F.1 is replaced by

F.1° If $\alpha\beta$ is defined in \mathcal{A} then $\beta^G \cdot \alpha^G$ is defined in \mathcal{B} and equals $(\alpha\beta)^G$.

Thus a contravariant functor from \mathcal{A} to \mathcal{B} may be described as an *antihomomorphism* from \mathcal{A} to \mathcal{B} , or also as a homomorphism from \mathcal{A}° to \mathcal{B} (or from \mathcal{A} to \mathcal{B}°).

As an example of a functor we have the derived group G' of a group G , viz. the subgroup generated by all commutators $x^{-1}y^{-1}xy$ ($x, y \in G$). For every homomorphism $f: G \rightarrow H$ there is a homomorphism $f': G' \rightarrow H'$, obtained by restriction from f , and it is clear that $(fg)' = f'g'$, $1' = 1$.

In any ring R we define $\mathbf{U}(R)$ to be the group of units of R . Given a ring homomorphism $f:R \rightarrow S$, we again have an induced homomorphism $\bar{f}:\mathbf{U}(R) \rightarrow \mathbf{U}(S)$, and \mathbf{U} is a functor from \mathbf{Rg} to \mathbf{Gp} . On the other hand, the centre of a group cannot be regarded as a functor; if the centre of G is denoted by $Z(G)$, then a homomorphism $G \rightarrow H$ need not map $Z(G)$ into $Z(H)$, as we see by taking G to be an abelian subgroup of H not contained in $Z(H)$, for a suitable group H .

All the categories mentioned above are *concrete*, in the sense that there is a functor F to \mathbf{Ens} , such that the induced mapping of hom-sets $(X, Y) \rightarrow \mathbf{Ens}(X^F, Y^F)$ is injective. This functor F , associating with each group, ring, etc., its underlying set, is called the *forgetful* functor.

Given two categories \mathcal{A}, \mathcal{B} and two functors S, T from \mathcal{A} to \mathcal{B} , we define a *natural transformation* from S to T as a family of \mathcal{B} -morphisms $\varphi_X:X^S \rightarrow X^T$ for each \mathcal{A} -object X , such that for any \mathcal{A} -morphism $f:X \rightarrow Y$ we have $f^S \varphi_Y = \varphi_X f^T$:

$$\begin{array}{ccc} X^S & \xrightarrow{f^S} & Y^S \\ \downarrow \varphi_X & & \downarrow \varphi_Y \\ X^T & \xrightarrow{f^T} & Y^T \end{array}$$

A natural transformation with a two-sided inverse which is again a natural transformation is called a *natural isomorphism*. Here S and T were assumed covariant throughout, but the same definition applies when both S and T are contravariant.

Examples. If V is a finite-dimensional vector space over a field k , and $V^* = \text{Hom}(V, k)$ is the dual space, then $\dim V = \dim V^*$ and so the spaces V, V^* are isomorphic, but the isomorphism is not natural (it depends on the choice of bases in V, V^*). On the other hand, there is a natural isomorphism between V and its bidual V^{**} . Let us write $\langle x, \alpha \rangle$ for the value of $\alpha \in V^*$ at $x \in V$ and (f, α) for the value of $f \in V^{**}$ at $\alpha \in V^*$. Then a natural transformation from V to V^{**} is given by

$$x \mapsto \hat{x}, \quad \text{where } \hat{x} \in V^{**} \text{ is defined by } (\hat{x}, \alpha) = \langle x, \alpha \rangle. \quad (1)$$

We observe that we cannot expect to find a natural transformation from V to V^* because the correspondence $V \mapsto V^*$ is a contravariant functor. Now it can be shown that for finite-dimensional vector spaces the correspondence (1) is a natural isomorphism (cf. 4.6).

As another example, consider, for any group G , the quotient $G^{ab} = G/G'$. It is easily shown (cf. Vol. 1, p. 267) that G^{ab} is the universal abelian homomorphic image of G , in the sense that G^{ab} is abelian and there is a homomorphism $v_G:G \rightarrow G^{ab}$ such that any homomorphism $f:G \rightarrow A$ to an abelian group A can be

uniquely factored as $f = v_G f'$, where $f': G^{ab} \rightarrow A$. Here v_G is a natural homomorphism from G to G^{ab} (cf. Vol. 1, p. 267).

Two categories \mathcal{A}, \mathcal{B} are said to be *isomorphic* if there is a functor $T: \mathcal{A} \rightarrow \mathcal{B}$ with an inverse, i.e. a functor $S: \mathcal{B} \rightarrow \mathcal{A}$, such that $ST = 1$, $TS = 1$. For example, the category of abelian groups is isomorphic to the category of \mathbf{Z} -modules; it is well-known that every abelian group may also be considered as a \mathbf{Z} -module and vice versa. Nevertheless the notion of isomorphism between categories is rather restrictive; it leaves out of account the fact that isomorphic objects in a category are for many purposes interchangeable. For this reason the following notion of equivalence is more useful:

Two categories \mathcal{A}, \mathcal{B} are said to be *equivalent* if there are two covariant functors $T: \mathcal{A} \rightarrow \mathcal{B}, S: \mathcal{B} \rightarrow \mathcal{A}$ such that TS is naturally isomorphic to the identity functor on \mathcal{A} , and similarly ST is naturally isomorphic to the identity on \mathcal{B} . When this holds for contravariant functors, \mathcal{A} and \mathcal{B} are called *dual* or *anti-equivalent*. For example, the reversal op: $\mathcal{A} \rightarrow \mathcal{A}^o$ is a duality.

Any functor $T: \mathcal{A} \rightarrow \mathcal{B}$ defines for each pair of \mathcal{A} -objects X, Y a mapping

$$\mathcal{A}(X, Y) \rightarrow \mathcal{B}(X^T, Y^T). \quad (2)$$

T is called *faithful* if (2) is injective, *full* if (2) is surjective and *dense* if each \mathcal{B} -object is isomorphic to one of the form X^T , for some \mathcal{A} -object X . For an equivalence functor T , (2) is a bijection, so in this case T is full and faithful, and clearly it is also dense. Conversely, suppose that T is full, faithful and dense. Then (2) is an isomorphism; moreover, for each \mathcal{B} -object Z we can by density find an \mathcal{A} -object Z^S such that $Z^{ST} \cong Z$, and now we can use the isomorphism (2) to transfer any map between \mathcal{B} -objects to a map between the corresponding \mathcal{A} -objects. Thus we obtain

PROPOSITION 3.1 *A functor is an equivalence if and only if it is full, faithful and dense.* ■

To illustrate the result, let k be a field and consider vec_k , the category of all finite-dimensional vector spaces over k with linear mappings as morphisms. In vec_k we have the subcategory col_k consisting of all column vectors over k , i.e. all spaces k^n ($n \geq 0$). Let us choose, for each vector space V of dimension n , an isomorphism $\theta_V: V \rightarrow k^n$. Define $T: \text{col}_k \rightarrow \text{vec}_k$ as the inclusion functor and $S: \text{vec}_k \rightarrow \text{col}_k$ as follows: $V^S = k^n$, where $n = \dim V$, and given $f: U \rightarrow V$, we put $f^S = \theta_U^{-1} f \theta_V$. This definition ensures that θ is a natural transformation from vec_k to col_k . We may without loss of generality take θ_V to be the identity when $V = k^n$. In that case we have $TS = 1$, while ST is naturally isomorphic to 1, via θ . Thus vec_k is equivalent to col_k . We observe that col_k is a small category, so vec_k is equivalent to a small category, though not itself small. We also note that we cannot choose a smaller category than col_k , for it has only one object of any given isomorphism

type. A category with this property is said to be *skeletal*, and for any category, a skeletal subcategory equivalent to it is called its *skeleton*. It is clear that any category has a skeleton, for we can always choose a subcategory by taking one copy from each isomorphism class of objects.

We also recall the universal mapping property (Vol. 1, p. 108f.). An *initial* object in a category \mathcal{A} is an object I in \mathcal{A} such that there is just one morphism from I to any \mathcal{A} -object; thus $\mathcal{A}(I, X)$ always has just one element and in particular the only map $I \rightarrow I$ is the identity on I . A category may have more than one initial object, but they are all isomorphic. For if I, I' are both initial, then there exist unique morphisms $\alpha:I \rightarrow I', \beta:I' \rightarrow I$; hence $\alpha\beta:I \rightarrow I$ must be the identity on I , and likewise $\beta\alpha$ is the identity on I' ; therefore α is an isomorphism. An initial object in the opposite category \mathcal{A}° is called a *final object* of \mathcal{A} . What we have proved can be stated as follows.

PROPOSITION 3.2 *In any category any two initial (or any two final) objects are isomorphic, by a unique isomorphism.* ■

As an illustration consider the universal mapping property for free groups. Given any set X , we take F_X to be the free group on X . It has the property that there is a mapping $v:X \rightarrow F_X$ and any mapping $f:X \rightarrow G$ to a group G can be factored uniquely by v , thus $f = vf'$ for a homomorphism $f':F_X \rightarrow G$. This is a well known property of free groups, the *universal mapping property* (Vol. 1, p. 296). If $U:Gp \rightarrow Ens$ is the forgetful functor from groups to sets, we can, for a fixed set X , form the category (X, Gp^U) whose objects are maps from X to G^U , the set underlying a group G , and whose morphisms are commutative triangles arising from homomorphisms $f:G \rightarrow H$. This is the *comma category* based on X and U . The free group F_X with the canonical map $X \rightarrow F_X^U$ may be described as an initial object in the comma category. By Prop. 3.2, F_X is unique up to isomorphism.

As a further illustration we recall from Vol. 1 (p. 250) the factor theorem for groups; the proof is an easy verification (cf. p. 251, Vol. 1).

THEOREM 3.3 *Given any group G and a normal subgroup N of G , there exists a group G/N (the quotient of G by N) and a homomorphism $v:G \rightarrow G/N$ (the natural homomorphism) which is universal for homomorphisms of G with kernel containing N . In detail, every homomorphism $f:G \rightarrow H$ such that $N \subseteq \ker f$ can be factored uniquely by v , $f = vf'$. Here f' is injective if and only if $N = \ker f$.* ■

Exercises

- (1) Show that the set \mathbf{N}_0 consisting of all the natural numbers and 0 can be defined as a category, which is equivalent to the category of all finite sets and mappings. Verify that it is a skeleton.

- (2) Show that Ab is a full subcategory of Gp , and that Rg is a subcategory of ‘rings which may lack 1’, but not full.
- (3) Show that if I is a small category and \mathcal{A} is any category, then there is a category $\text{Fun}(I, \mathcal{A})$ whose objects are functors from I to \mathcal{A} , while the morphisms are natural transformations.
- (4) Let I be a small category. Show that the functor from I° to $\text{Fun}(I, \text{Ens})$ which maps X to $I(X, -)$ is full and faithful.
- (5) Let \mathcal{A} be a small category and for any \mathcal{A} -objects X, Y write $X \leqslant Y$ iff $\mathcal{A}(X, Y) \neq \emptyset$. Show that ‘ \leqslant ’ is a preordering on the object set of \mathcal{A} . Verify that, conversely, any preordered set can be made into a small category by introducing a morphism $\alpha: x \rightarrow y$ whenever $x \leqslant y$. Show that any skeleton of the resulting category is an ordered set.
- (6) Find a category with a skeleton which is not small.

1.4 Graphs

Many problems both in mathematics and elsewhere can best be solved diagrammatically; the diagrams involved are more or less subtle reformulations of the problem, and the efforts to solve these problems have given rise to the theory of graphs. We can do no more here than present the beginnings of the theory, but it seems appropriate to do so since the methods are often algebraic and graphs are increasingly being used in other parts of mathematics as well as in algebra itself.

A graph Γ consists of a pair of sets V, E . The members of V are the *points* or *vertices* of Γ while the members of E are its *edges*. With each edge of E two points are associated, its *endpoints*. Here the endpoints of an edge need not be distinct; if they coincide, the edge is a *loop*. Two edges may have the same pair of (distinct) endpoints, giving a *multiple edge*. Two vertices are *adjacent* if they are joined by an edge. Some simple examples are illustrated here, where the first two represent *simple* graphs, i.e. graphs without loops or multiple edges. Given a graph $\Gamma = \{V, E\}$, if V' is a subset of V and E' is a subset of E such that any edge in E' has its endpoints in V' , then $\{V', E'\}$ is again a graph, called a *subgraph* of Γ . A subgraph Γ' is said to be *full* if for any vertices p, q in Γ' all edges between p and q in Γ belong to Γ' .



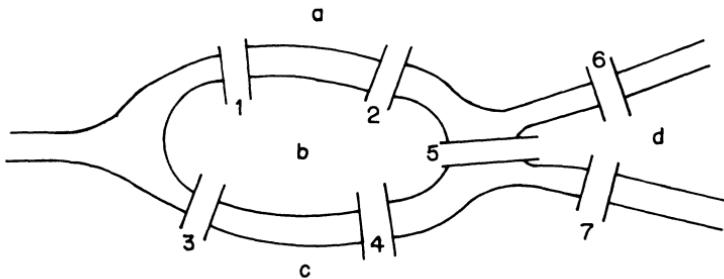
- b) Let S be any set; by the *complete graph* $C(S)$ on S we understand the graph with S as vertex set and an edge between each distinct pair of vertices. Every simple graph $\Gamma = \{V, E\}$ is clearly a subgraph of $C(V)$; the graph with vertex set V and the set of edges of $C(V)$ that are not in Γ is written Γ' and called the *complementary graph* of Γ .

Examples

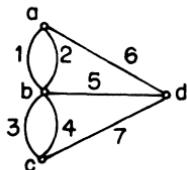
1. In a group of six people it is always possible to find either three people who know each other or three people who are all strangers to each other.

In order to prove this statement we represent the six individuals by points and join any two points representing acquaintances by an edge. We thus obtain a graph Γ , the ‘acquaintanceship graph’ of our group, and we have to show that either Γ contains three edges forming a triangle, or its complement Γ' does so. Take a vertex p_1 ; it is adjacent to each of the other five points in just one of Γ, Γ' . Hence it must be adjacent to three points in one of these graphs, say in Γ it is adjacent to p_2, p_3, p_4 . If two of p_2, p_3, p_4 are adjacent in Γ , then these two vertices together with p_1 form a triangle in Γ ; otherwise p_2, p_3, p_4 form a triangle in Γ' .

2. (The Königsberg bridge problem). A famous problem concerns the seven bridges of Königsberg (crossing the river Pregel), which are situated as shown.



The problem was to cross in the course of a single walk each bridge exactly once. It is not hard to convince oneself by trial and error that this is impossible; this was first proved by Euler in 1736. The first step is to represent the problem by a graph in which the areas are points and the bridges become edges.



For each vertex we define its *valency* as the number of edges ending in it (counting loops twice). A walk across all the bridges becomes a path which includes each edge just once. If this path begins and ends at the same point, each point has even valency; otherwise there are just two points of odd valency, the beginning and endpoint of our path. But in the above graph all four points have odd valency, so there cannot be such a path.

Formally we define a *path* of length n from p to q in a graph as a set of edges

e_1, \dots, e_n such that e_i has endpoints p_{i-1}, p_i and $p_0 = p, p_n = q$. A path from p to q is called a *cycle* and a graph without cycles (of positive length) is said to be *acyclic*.

We note that every finite partially ordered set may be considered as a graph by drawing an edge from p to q if q covers p , i.e. $p < q$ but $p < x < q$ for no x . In fact we thus obtain a *directed graph* or *digraph*, i.e. a graph in which the endpoints for each edge form an ordered pair, the *initial vertex* and the *final vertex*; the edges in a digraph are also called *arrows* and a finite digraph is sometimes called a *quiver*.

In a digraph only paths are allowed which go in the direction of each arrow. Such a graph clearly defines a small category in which the vertices are the objects and the paths are the morphisms.

It is clear that the digraph derived from an ordered set is acyclic; conversely, given any acyclic digraph Γ , we can define a partially ordered set on the vertex set of Γ by writing $p \leq q$ whenever there is a path from p to q . Thus finite partially ordered sets may be identified with directed acyclic graphs. Our first result, though nominally about ordered sets, is really graph-theoretical in nature.

THEOREM 4.1 (Dilworth's theorem) *Let S be a finite partially ordered set. Then the minimum number of disjoint chains into which S can be decomposed is the width of S , i.e. the maximum number of elements in an anti-chain of S .*

Proof. We use induction on $|S|$. Suppose that m is the width of S ; clearly we cannot express S as a disjoint union of fewer than m chains, and we have to show that S can be expressed as a disjoint union of m chains. Let S' be the set obtained by omitting a minimal element c from S . Then S' has width m' , where $m' = m - 1$ or $m' = m$.

(i) $m' = m - 1$. By the induction hypothesis S' can be written as a disjoint union of $m - 1$ chains, hence $S = S' \cup \{c\}$ can be written as a disjoint union of m chains.

(ii) $m' = m$. Using again induction on $|S|$, we can write S' as a disjoint union of m chains, say $S' = C_1 \cup \dots \cup C_m$. Let b_i be the least element in C_i ; then $\{b_1, \dots, b_m\}$ is an anti-chain, but $\{b_1, \dots, b_m, c\}$ contains $m + 1$ elements and so cannot be an anti-chain. Thus c is comparable with b_i for some i , say $i = 1$, and since c is minimal in S , we have $c < b_1$. Put $C'_1 = C_1 \cup \{c\}$; then $S = C'_1 \cup C_2 \cup \dots \cup C_m$ and this expresses S as a union of m disjoint chains, as claimed. ■

Dilworth's theorem can be used to prove P. Hall's theorem on distinct representatives:

THEOREM 4.2 (P. Hall's theorem) *Let S_1, \dots, S_m be subsets of a finite set T . Then there is a family of distinct elements a_1, \dots, a_m in T such that $a_i \in S_i$ provided that the intersection of any k of the subsets S_i contains at least k elements ($k = 1, 2, \dots, m$).*

Proof. Let $T = \{a_1, \dots, a_n\}$ and define a partial ordering on the set W consisting of all the S_i and a_j by writing $x \leq y$ precisely when $x = y$ or $y = S_i$ and $x = a_j \in S_i$. Suppose that we have an anti-chain in W :

$$\{S_1, \dots, S_r, a_1, \dots, a_s\}. \quad (1)$$

By hypothesis $S_1 \cap \dots \cap S_r$ contains at least r elements, which must be distinct from a_1, \dots, a_s because (1) is an anti-chain. It follows that $r + s \leq n$, and so every anti-chain in W has at most n elements; in fact this bound is reached for $\{a_1, \dots, a_n\}$. By Th. 4.1 we can decompose our set W into n disjoint chains. Each S_i is in just one of the chains and each chain contains an a_j . Thus $a_j \in S_i$ and we have a system of distinct representatives. ■

Remarks

1. Let S_1, \dots, S_m and T_1, \dots, T_n be two families of disjoint subsets of a set. If any k of the S_i between them meet at least k of the T_j then we can find $a_i \in S_i$ such that distinct a_i belong to distinct T_j 's. For example, in a finite group G with a subgroup H we can take the S_i to be the left cosets and the T_j to be the right cosets. We thus obtain a set of common representatives for the left and right cosets of H (this was Hall's original purpose in proving his theorem).

2. The assertion of Th.4.2 has been formulated more picturesquely by P. Halmos as the 'marriage theorem': In a group of m men and n women, if any k men between them are acquainted with at least k of the women, then each man can be married off to an acquaintance.

A graph is said to be *connected* if any two of its points can be joined by a path; if, moreover, there is a unique path joining any two points, the graph is called a *tree*. Clearly a tree can also be characterized as a connected graph which is acyclic. It is an important fact that any connected graph contains a tree which includes all its vertices.

THEOREM 4.3 *Let Γ be a connected graph. Then there is a subgraph Γ_0 including all the vertices of Γ which is a tree. Moreover, for any finite tree (V, E) we have*

$$|V| = |E| + 1. \quad (2)$$

Proof. We observe that a graph Γ_1 is a tree iff any full subgraph on finitely many vertices is a tree, for if Γ_1 contains a cycle, this already appears in a finite subgraph, and if Γ_1 fails to be connected, then the same is true of a full 2-element subgraph. It follows by Prop. 2.4 that Γ contains a maximal subgraph Γ_0 which is a tree (i.e. a maximal subtree). We claim that Γ_0 contains all vertices of Γ . For otherwise there is a vertex p adjacent to a vertex in Γ_0 but not itself in Γ_0 . Since Γ is connected, there is an edge e from p to $q \in \Gamma_0$. Then the graph obtained by

adjoining p and the edge e to Γ_0 is still a tree, but it contains Γ_0 as a proper subtree, contradicting the maximality of the latter. Thus Γ_0 contains all vertices of Γ and it is the required tree.

Now let $\Gamma = \{V, E\}$ be a finite tree; in Γ we can find a vertex p_0 which is adjacent to only one other vertex. To find p_0 we start from any vertex of Γ along a path and continue as far as possible without traversing an edge more than once. As long as all the vertices we reach have valency greater than 1 we can continue, and we never pass a vertex twice because Γ is a tree. Since Γ is finite, this process must come to a halt, and it can only do so when we reach a vertex of valency 1; this is the required vertex p_0 . If we omit p_0 and the single edge ending at p_0 from Γ we obtain a tree $\Gamma' = \{V', E'\}$ with fewer vertices than Γ . By induction we have $|V'| = |E'| + 1$, and since $|V| = |V'| + 1$, $|E| = |E'| + 1$, we obtain (2). ■

Our final result in this section, Ramsey's theorem, is a far-reaching generalization of an earlier example involving the acquaintanceship graph (p. 22). For brevity let us call a set or a subset consisting of r elements an r -set, respectively an r -subset, and write $\mathcal{P}_r(S)$ for the set of all r -subsets of S ; for example, every r -set has exactly one r -subset and one 0-subset. The earlier example showed that if the collection of all 2-subsets of a 6-set S is partitioned in any way into two disjoint sets A_1, A_2 , then either S contains a 3-subset all of whose 2-subsets lie in A_1 or it contains a 3-subset all of whose 2-subsets lie in A_2 . In terms of graphs, if in a complete graph on six vertices each edge is painted either blue or red, then there is a complete subgraph on three vertices in which all edges have the same colour. More generally, given integers $p, q \geq 2$, there exists an integer N such that if in a complete graph on N vertices the edges are painted blue or red, then there is either a complete blue subgraph on p vertices or a complete red subgraph on q vertices. This is just the special case $r = 2$ of the following theorem:

THEOREM 4.4 (Ramsey's theorem) *Let p_1, p_2, r be integers such that $0 \leq r \leq p_i$ ($i = 1, 2$). Then there exists an integer $N(p_1, p_2, r)$ with the following property: If in any set S of at least $N(p_1, p_2, r)$ elements the family of r -subsets is partitioned into two disjoint sets A_1 and A_2 , then for $i = 1$ or 2 , S contains a p_i -subset all of whose r -subsets lie in A_i .*

In symbols we have $\mathcal{P}_r(S) = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$, and the assertion is that there is a p_i -subset T of S such that $\mathcal{P}_r(T) \subseteq A_i$ for $i = 1$ or 2 .

Proof. We shall use induction on r , and for given r , on p_1 and p_2 . For $r = 0$ the result holds trivially: there is only one 0-subset of any set and this lies in A_1 or A_2 . We may therefore assume that $r > 0$.

Firstly we note that $N(p_1, r, r) = p_1$. For if S has at least p_1 ($\geq r$) elements, and the family of r -subsets of S has been partitioned into disjoint sets A_1 and A_2 , then either $A_2 \neq \emptyset$ and so S contains an r -subset whose unique r -subset lies in A_2 , or

$A_2 = \emptyset$ and any p_1 -subset of S has all its r -subsets in A_1 . This shows that $N(p_1, r, r) \leq p_1$ and it is easily seen that the inequality here cannot be strict. Now a similar argument shows that $N(r, p_2, r) = p_2$.

We put $q_1 = N(p_1 - 1, p_2, r)$, $q_2 = N(p_1, p_2 - 1, r)$ and claim that $N(p_1, p_2, r) \leq N(q_1, q_2, r - 1) + 1$. For let S be a set with at least $N(q_1, q_2, r - 1) + 1$ elements and let A_1, A_2 be a given partition of the family of r -subsets of S . Fix $c_0 \in S$, write $S' = S \setminus \{c_0\}$ and define a partition $\{B_1, B_2\}$ of the family of $(r - 1)$ -subsets of S' by taking an $(r - 1)$ -subset T to B_i whenever $T \cup \{c_0\} \in A_i$. Since S' has at least $N(q_1, q_2, r - 1)$ elements, we can apply the induction hypothesis. Two cases can arise: (i) S' has a q_1 -subset U all of whose $(r - 1)$ -subsets are in B_1 . If U contains a p_2 -subset all of whose r -subsets lie in A_2 , then the conclusion follows. Otherwise, by the definition of q_1 , U has a $(p_1 - 1)$ -subset whose r -subsets are all in A_1 , and adjoining c_0 we get a p_1 -subset all of whose r -subsets are in A_1 . The second case (ii) is that S' has a q_2 -subset all whose $(r - 1)$ -subsets are in B_2 ; this can be dealt with by symmetry, so the conclusion holds for S . ■

The least value of $N(p_1, p_2, r)$ is sometimes called the *Ramsey number*. For example, the example given earlier shows that $N(3, 3, 2) \leq 6$, and it is easily checked that equality holds here. For $r = 1$ we have $N(p_1, p_2, 1) = p_1 + p_2 - 1$ and the theorem states that when a set of at least $p_1 + p_2 - 1$ elements is partitioned into two disjoint subsets A_1, A_2 , then A_i has at least p_i elements for either $i = 1$ or 2.

Exercises

- (1) Construct all simple graphs with four edges and no isolated points (there are 11, up to isomorphism). Construct all trees with five edges (there are six).
- (2) Let S be a partially ordered set in which each chain and each anti-chain is finite. Define $l(a)$ for $a \in S$ as the minimum of the lengths of maximal chains below a , and show that $A_n = \{a \in S \mid l(a) = n\}$ is an anti-chain. By considering the maximal elements of S , write S as a union of a finite number of the A_n and hence show that S must be finite.
- (3) Prove Dilworth's theorem in the infinite case: Every partially ordered set of width m can be written as the disjoint union of m chains.
- (4) Prove König's lemma: Every infinite connected graph in which all points have finite valency has an infinite path. (*Hint.* Choose p_1, p_2, \dots, p_n so that there is a path along these points and p_n can be connected to infinitely many points by paths not passing through p_{n-1} .) Show that the result need not hold if there is at least one point of infinite valency.
- (5) Let Γ be a finite graph, not necessarily connected but for which each connected component is a tree (such Γ is sometimes called a 'forest'). Show that Γ has $|V| - |E|$ connected components. (*Hint.* Use induction on the number of edges.)

- (6) Show that $N(p, q, 0) = \max \{p, q\}$, $N(p, q, 1) = p + q - 1$, $N(p, q, 2) \leq \binom{p+q-2}{p-1}$.
- (7) Show that of five points in a plane, no three of which are collinear, there are four that form vertices of a convex quadrilateral.
- (8) Given n points in a plane of which no three are collinear, show that if all the quadrilaterals formed from these points are convex, then there is a convex n -gon with these points as vertices.
- (9) Show that, for any n , if at least $N(n, 5, 4)$ points in the plane are given, no three of which are collinear, then there are n points in the given set forming a convex n -gon. (Erdős–Szekeres. For $n = 3$ take $N = 3$; otherwise use Th. 4.4 and Exs. (7)–(8).)
- (10) Show that there is a number $N = N(p_1, \dots, p_t, r)$ such that if for any set of at least N elements the family of subsets is partitioned into disjoint sets A_1, \dots, A_t then, for some $i = 1, \dots, t$, S contains a p_i -subset all of whose r -subsets lie in A_i .
- (11) Show that if the r -subsets of an infinite set S are partitioned into disjoint sets A_1, \dots, A_t , then there is an infinite subset of S whose r -subsets all lie in A_i , for some i (*Hint.* Adapt the proof of Th. 4.4.)
- (12) Show that an infinite partially ordered set contains an infinite chain or an infinite anti-chain. (*Hint.* Use Ex.(11).)

Further exercises on Chapter 1

- (1) Show that the union of a finite family of finite sets is finite.
- (2) Show that if α, β are cardinals of which at least one is infinite, then $\alpha + \beta = \max \{\alpha, \beta\}$.
- (3) Show that a set is finite iff every collection of subsets has either a maximal or a minimal member.
- (4) (Sierpiński) A collection T of sets will (in this exercise) be called a *tower* if $\emptyset \in T$ and $X, Y \in T \Rightarrow X \cup Y \in T$. Show that a set A is finite iff A belongs to every tower T such that $\{x\} \in T$ for all $x \in A$.
- (5) Prove that Zorn's lemma is equivalent to Hausdorff's maximal principle: Given a collection A of sets, any chain in A is contained in a maximal chain.
- (6) (Tarski) Show that a set is finite iff it can be ordered so that both the ordering and its opposite are well-orderings.
- (7) Define an *upper segment* in an ordered set as a lower segment in the opposite ordering. Verify that the upper segments are precisely the complements of the lower segments. If a set

A is well-ordered, and order-isomorphic to every non-empty upper segment of itself, what can be said about A ? Show that A cannot be of type 2ω .

(8) Let A be a countable set. Show that the set of all equivalences on A with finitely many equivalence classes is uncountable, but the subset of those equivalences in which only one class is infinite is countable.

(9) Let S be a set of n elements and $\mathcal{C} = \{C_i\}$ a collection of subsets of S such that $C_i \neq C_j$ for $i \neq j$ and $C_i \cap C_j \neq \emptyset$. Show that \mathcal{C} has at most 2^{n-1} members.

(10) Let A be a well-ordered set. Show that any interval of the form $\{x \in A \mid a \leq x < b\}$ is order-isomorphic to a lower segment of A .

(11) Let S be a set of n elements. Show that $\mathcal{P}(S)$ has $n!$ maximal chains and that each r -subset of S is contained in exactly $r!(n-r)!$ maximal chains. Show further that if $\mathcal{P}(S)$ has an anti-chain containing v_r r -subsets, then

$$\sum v_r \binom{n}{r}^{-1} \leq 1.$$

Deduce Sperner's lemma: The maximal length of any anti-chain in $\mathcal{P}(S)$ is the number of $[n/2]$ -subsets in S (D. Lubell).

(12) (Yoneda's lemma) Let $F: \mathcal{A} \rightarrow \text{Ens}$ be a functor and for any $p \in X^F$ define a natural transformation $\dot{p}: h^X = \mathcal{A}(X, -) \rightarrow F$ by the rule: if $\alpha \in \mathcal{A}(X, Y)$, then $\alpha \mapsto p\alpha^F$ maps $Yh^X (= \mathcal{A}(X, Y))$ to Y^F . Verify that this is indeed a natural transformation and prove that the resulting mapping $X^F \rightarrow \text{Nat}(h^X, F)$ to the set of all natural transformations is an isomorphism. (*Hint.* Define the inverse $\tau \mapsto (1_X)\tau \in X^F$.)

(13) Let Γ be a graph and Δ a subgraph. The graph Δ^c whose edges are the edges of Γ not in Δ and whose vertices are the vertices of Γ that are either not in Δ or incident with an edge of Γ not in Δ is called the *complement* of Δ in Γ . Show that Δ^{cc} is obtained from Δ by deleting the vertices of Δ that are incident with an edge of Γ not in Δ but not incident with any edge in Δ .

(14) Show that in any finite graph the number of points of odd valency is even.

(15) A finite graph is said to be *Eulerian* if it is connected and has a closed path (i.e. cycle) including each edge once. Show that a connected graph is Eulerian iff each vertex has even valency. Find a condition for a graph to have a path (not necessarily closed) that includes each edge just once.

(16) For any finite graph $\Gamma = (V, E)$ define p_0 as the number of its connected components and p_1 as the least number of edges which need to be removed to make each component into a tree. Show that $|V| - |E| = p_0 - p_1$. Show also that if $v(x)$ is the valency of $x \in V$, then

$$\sum_{x \in V} [v(x) - 2] = 2(p_1 - p_0).$$

- (17) Let S be a finite set and \mathcal{F} a family of subsets of S whose union is S itself. The *intersection graph* of \mathcal{F} is defined as the simple graph with \mathcal{F} as vertex set, and where $X, Y \in \mathcal{F}$ are joined by an edge whenever $X \neq Y$ and $X \cap Y \neq \emptyset$. Show that every finite simple graph is the intersection graph of an appropriate family of sets.
- (18) Let Γ be a finite graph and for each vertex p define $d(p)$ as the length of the longest simple path starting at p . Show that for a tree, $d(p)$ assumes its least value either at a single vertex or at several adjacent vertices.
- (19) With every finite graph Γ one associates its *adjacency matrix* A , a square matrix whose rows and columns are indexed by the vertices of Γ and whose (i,j) -entry is the number of edges from i to j . Show that A is symmetric and that Γ is connected iff there is no permutation matrix P such that $P^{-1}AP$ is the diagonal sum of two matrices. Interpret the entries of A^n .
- (20) Deduce Zorn's Lemma from the well-ordering theorem.

2

Lattices

The subsets of a set permit operations quite similar to those performed on numbers. If for the moment we denote the union of two subsets A, B by $A + B$ and their intersection by AB , a notation which will not be used later (despite some historical precedents), then we have laws like $AB = BA$, $A(B + C) = AB + AC$, similar to the familiar laws of arithmetic, as well as new laws such as $A + A = A$, $A + BC = (A + B)(A + C)$. The algebra formed in this way is called a *Boolean algebra*, after G. Boole who introduced it around the middle of the 19th century, and who made the interesting observation that Boolean algebras could also be used to describe the propositions of logic.

The more general notion of a lattice was first used by Dedekind to study the relations between ideals in rings of numbers. The lattice concept helps to unify a number of disparate ideas and this chapter deals with the basic properties of lattices and Boolean algebras. The latter have recently found applications in switching theory, but we shall not enter on this aspect. We shall concentrate on applications to algebra, and describe such important ideas as chain conditions and Möbius functions in the general setting of partially ordered sets.

2.1 Definitions; modular and distributive lattices

The definitions relating to ordered sets, recalled in 1.2, form the foundation for much that follows. Whereas sets of numbers, like \mathbb{N} or \mathbb{Q} , are totally ordered, this property is not shared by most partially ordered sets. Examples are the set $\mathcal{P}(A)$ of all subsets of a set A (with more than one element) relative to inclusion, or the set \mathbb{N} of natural numbers relative to divisibility. In both these cases any two elements have a least upper bound or *supremum*, briefly sup: given x, y , there exists z such that (i) $x \leq z$, $y \leq z$ and (ii) $x \leq z'$, $y \leq z'$ implies $z \leq z'$. Dually any two elements have a greatest lower bound or *infimum*, briefly inf, defined similarly. Thus if X and Y are subsets of A , their sup is $X \cup Y$ and their inf is $X \cap Y$; if m, n are natural numbers, their sup (relative to divisibility) is their least common multiple while their inf is their highest common factor.

A partially ordered set in which any two elements have a supremum and an infimum is called a *lattice*. The sup of x, y is written $x \vee y$ and is called the *join* of x and y ; their inf is written $x \wedge y$ and is called the *meet* of x and y . It is clear that the

lattice concept is self-dual, so that the dual of a lattice, obtained by reversing the ordering, is again a lattice. As examples of lattices we have the set \mathbb{N} of natural numbers under divisibility, as well as $\mathcal{P}(A)$, the set of all subsets of A under inclusion; further, every totally ordered set is trivially a lattice, with $x \wedge y = x$, $x \vee y = y$ if $x \leqslant y$.

Partially ordered sets are often represented as directed graphs: the elements of the set form the vertices, and edges are drawn so that $a \leqslant b$ holds precisely when b is higher than a and there is a descending path from b to a . Such diagrams are mainly used for finite sets; for \mathbb{N} or \mathbb{Z} the structure can be hinted at in a partial diagram, but the case of \mathbb{Q} or \mathbb{R} would be more difficult. Some examples are given in Figs. 1–3, where Figs. 1 and 2 are lattices, but not Fig. 3.

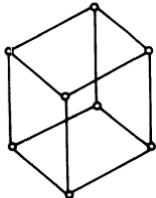


Fig. 1

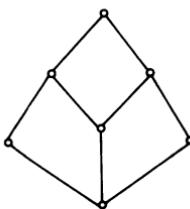


Fig. 2

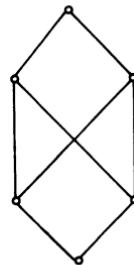


Fig. 3

A lattice may be regarded as a set with two binary operators, \vee and \wedge . These operators satisfy a number of laws, reminiscent of the laws governing addition and multiplication of numbers, and these laws can be used to give an alternative definition of lattices.

PROPOSITION 1.1 *Let L be a lattice. Then for any $a, b, c \in L$,*

$$a \vee (b \vee c) = (a \vee b) \vee c, \quad a \wedge (b \wedge c) = (a \wedge b) \wedge c \quad (\text{associative law}) \quad (1)$$

$$\bullet \quad a \vee b = b \vee a, \quad a \wedge b = b \wedge a \quad (\text{commutative law}) \quad (2)$$

$$a \wedge (a \vee b) = a, \quad a \vee (a \wedge b) = a, \quad (\text{absorptive law}), \quad (3)$$

$$a \vee a = a, \quad a \wedge a = a \quad (\text{idempotent law}). \quad (4)$$

Conversely, if L is a set with two binary operators \vee , \wedge satisfying (1)–(3), then (4) also holds and a partial ordering may be defined on L by the rule

$$a \leqslant b \quad \text{if and only if} \quad a \vee b = b. \quad (5)$$

Relative to this ordering L is a lattice such that the join of a, b is $a \vee b$ and the meet is $a \wedge b$.

Proof. By the definition of $a \vee b$ as sup we see that the unique sup of a, b can be written either as $a \vee b$ or as $b \vee a$, and a similar remark applies to $a \wedge b$, hence (2). Likewise the sup of a, b, c may be written $a \vee (b \vee c)$ or $(a \vee b) \vee c$, hence (1). Now (3) holds because $a \wedge b \leq a \leq a \vee b$, and (4) is a trivial consequence of the definition, but we observe that it also follows from (3): if in the first eqn. (3) we replace b by $a \wedge a$ we get $a = a \wedge [a \vee (a \wedge a)] = a \wedge a$, by the second eqn. (3). This proves the first idempotent law; the second follows by duality.

Now let L be a set with two operators \vee, \wedge satisfying (1)–(3); then (4) also holds, as we have just seen, and moreover,

$$a \vee b = b \Leftrightarrow a \wedge b = a. \quad (6)$$

For if $a \vee b = b$, then by (3), $a = a \wedge (a \vee b) = a \wedge b$, and the converse follows by duality (and the commutative law (2)). If we define the relation ' \leq ' by (5) we have a partial ordering: if $a \leq b, b \leq c$, then $a \vee b = b, b \vee c = c$; hence by (1), $c = b \vee c = (a \vee b) \vee c = a \vee (b \vee c) = a \vee c$, so $a \leq c$. Further, $a \leq a$ by (4), and if $a \leq b, b \leq a$, then $b = a \vee b = b \vee a = a$, by (2).

By the definition of $a \leq b$ and (2), $a \vee b$ is an upper bound of a, b . If c is another upper bound, then $a \leq c, b \leq c$, hence $c = a \vee c = b \vee c$, and so $c = a \vee (b \vee c) = (a \vee b) \vee c$, i.e. $a \vee b \leq c$, which shows $a \vee b$ to be the least upper bound, i.e. the sup. Now by duality and (6) it follows that the inf of a, b is $a \wedge b$. ■

Any subset of an ordered set S is again ordered, but it need not be a lattice, even if S is one. Guided by Prop. 1.1, we define a *sublattice* of a lattice L as a subset M which admits the operators \vee, \wedge of L , i.e. given $a, b \in M$, we have $a \vee b, a \wedge b \in M$. It is clear that M is again a lattice. In terms of the partial ordering the definition may be expressed as follows: M is a sublattice of L , if for any $a, b \in M$, their sup and inf (taken in L) again lie in M . We note that it may very well be possible for a subset of L to be a lattice without being a sublattice of L . For example, let L be a group, $\mathcal{P}(G)$ the set of all subsets of G and $\text{Lat}(G)$ the set of all subgroups of G . As we shall soon see, $\text{Lat}(G)$ is again a lattice with respect to the ordering by inclusion, as is $\mathcal{P}(G)$, but $\text{Lat}(G)$ is not usually a sublattice of $\mathcal{P}(G)$, because the union $H \cup K$ of two subgroups need not be a subgroup.

It is clear from the definition that in a lattice L any intersection of sublattices is again a sublattice. Thus we can define the sublattice *generated* by a subset X of L as the intersection of all sublattices containing X . As in the case of subgroups, this sublattice can be obtained by repeated application of the lattice operations to the elements of X .

As for groups or rings we define a *homomorphism* of lattices as a mapping $f: L \rightarrow L'$ between lattices L, L' such that

$$(a \vee b)f = af \vee bf, \quad (a \wedge b)f = af \wedge bf. \quad (7)$$

It is clear that a lattice-homomorphism preserves the ordering: $a \leq b$ implies $af \leq bf$, but not every order-preserving mapping between lattices is a lattice-

homomorphism. For example, the mapping $\alpha_c: x \mapsto x \vee c$ (for a fixed element c) in any lattice is order-preserving:

$$a \leq b \Rightarrow a \vee c \leq b \vee c, \quad (8)$$

and although α_c satisfies the first eqn. (7), it does not generally satisfy the second (this is in fact the distributive law, to be discussed later). However, an order-preserving bijection with an order-preserving inverse between lattices is always a lattice-isomorphism, because the sets are then order-isomorphic and the lattice operations can be defined in terms of the ordering (as in (5)).

In any lattice each finite (non-empty) subset has a sup and an inf, as an easy induction shows. Explicitly, the sup and inf of a_1, \dots, a_n are given by

$$a_1 \vee \cdots \vee a_n \quad \text{and} \quad a_1 \wedge \cdots \wedge a_n$$

respectively. Here we may omit brackets, by associativity, and the order of the factors is immaterial, by commutativity.

The notions of sup and inf can also be defined for infinite subsets, but they will not necessarily exist in a general lattice. A lattice L in which every subset has a sup and an inf is said to be *complete*. In particular such a lattice L has a greatest element ($\sup L$ or $\inf \emptyset$), denoted by 1, and a least element ($\inf L$ or $\sup \emptyset$), denoted by 0. For example, every finite lattice is complete and so is $\mathcal{P}(A)$, for any set A . The following criterion for completeness is often useful:

PROPOSITION 1.2 *If L is a partially ordered set such that every subset has an inf, then L is a complete lattice.*

For, given $X \subseteq L$, let Y be the set of all upper bounds of X in L and set $y = \inf Y$. Any element of X is a lower bound of Y , hence $x \leq y$ for all $x \in X$. If also $x \leq z$ for all $x \in X$, then $z \in Y$, by the definition of Y and so $y \leq z$; therefore $y = \sup X$. ■

For example, the set of all subgroups of a group G is partially ordered by inclusion and if $\{H_\lambda\}$ is a family of subgroups, then their intersection $\bigcap H_\lambda$ is again a subgroup; thus every subset of our set has an inf. Applying Prop.1.2, we see that it is a complete lattice, which we denote by $\text{Lat}(G)$. The inf of a family $\{H_\lambda\}$ of subgroups is their intersection, while their sup is the least subgroup containing all the H_λ , i.e. the subgroup generated by all the H_λ . We see that although the inf in $\text{Lat}(G)$ is the same as in $\mathcal{P}(G)$, the sup in general is not. In a similar way the submodules of an R -module M form a lattice $\text{Lat}_R(M)$, with $A \cap B$ as meet and $A + B$ as join.

Many of the lattices we shall meet in the sequel have the following property, called the *modular law*:

$$a \vee (b \wedge c) = (a \vee b) \wedge c \quad \text{for all } a, b, c \in L \text{ such that } a \leq c. \quad (9)$$

A lattice L satisfying (9) is said to be *modular* (Dedekind's name was 'Dualgruppe'

vom Modultypus'). For example, let G be a group and consider the set $\mathcal{N}(G)$ of all normal subgroups of G . This is a lattice for the ordering by inclusion, with $H \wedge K = H \cap K$, $H \vee K = HK$. In fact it is a modular lattice, i.e.

$$A(B \cap C) = AB \cap C, \quad \text{whenever } A \subseteq C. \quad (10)$$

For we have $A, B \cap C \subseteq AB, C$, hence $A(B \cap C) \subseteq AB \cap C$. Now take $c \in AB \cap C$, say $c = ab$, where $a \in A$, $b \in B$; then $b = a^{-1}c \in C$, hence $b \in B \cap C$ and so $c = ab \in A(B \cap C)$.

The lattice of *all* subgroups of a group will not in general be modular, and $H \vee K$ need not equal HK (cf. Ex. (10)). The relation (10) holds more generally for normal subgroups with operators, in particular it holds for submodules of a module (which accounts for the name of (9)).

To see why modular lattices are more tractable, let us return to a general lattice L for a moment. With $a, b \in L$ such that $a \leq b$ we can associate the *interval*

$$[a, b] = \{x \in L \mid a \leq x \leq b\}.$$

Such an interval need not be a chain, but it is always a sublattice of L , with least element a and greatest element b . More generally, for any $a, b \in L$ we can form the intervals $I = [a \wedge b, a]$ and $J = [b, a \vee b]$. Let us define a mapping $\alpha: I \rightarrow J$ by

$$\alpha: x \mapsto x \vee b,$$

and a mapping $\beta: J \rightarrow I$ by

$$\beta: y \mapsto y \wedge a.$$

For any $x \in I$ we have $x\alpha\beta = (x \vee b) \wedge a$. Here $x \leq a$; hence if L is modular, then $x\alpha\beta = (x \vee b) \wedge a = x \vee (b \wedge a) = x$. Thus $\alpha\beta = 1$ and dually $\beta\alpha = 1$, i.e. in a modular lattice the mappings α, β are mutually inverse. Since α and β are both order-preserving, it follows that I and J are isomorphic as lattices. But in concrete cases the explicit form of the mappings α, β often tells us more than this.

Let us call two intervals I and J that are related as in Fig. 4 *perspective*, and two intervals related by a series of perspectivities *projective*. The second isomorphism theorem of group theory (Vol. 1, p. 252 and 4.1 below) shows that in the lattice $\mathcal{N}(G)$ of all normal subgroups of G , perspective intervals define isomorphic quotients. Hence the same is true of projective intervals and the usual proof of the Schreier refinement theorem (Vol. 1, p. 257) shows that any two normal chains have refinements in which corresponding intervals are projective, and hence define isomorphic quotient groups (however, this theorem does not operate within $\mathcal{N}(G)$).

To obtain a criterion for modularity, let us call two elements x, y in an interval $[a, b]$ *complementary* and call each a relative *complement* (in $[a, b]$) of the other if $x \wedge y = a$, $x \vee y = b$.

PROPOSITION 1.3 *A lattice L is modular if and only if for each interval I of L ,*

any two elements of I which are comparable and have a common complement in I are equal.

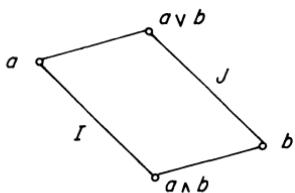


Fig. 4

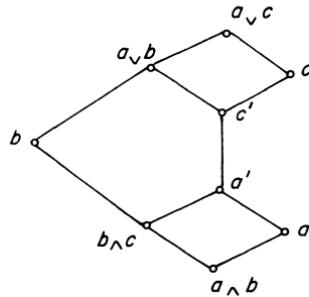


Fig. 5

Proof. Given a, b, c in any lattice L , if $a \leq c$, then $a \vee (b \wedge c) \leq a \vee b$ and $a \vee (b \wedge c) \leq c$, hence

$$a \vee (b \wedge c) \leq (a \vee b) \wedge c. \quad (11)$$

Therefore L is non-modular iff the inequality (11) is strict for at least one triple (a, b, c) such that $a \leq c$. When $a = c$, the two sides of (11) are equal by the absorptive law (3), so we may assume that $a < c$. Suppose first that strict inequality holds in (11). Put $a' = a \vee (b \wedge c), c' = (a \vee b) \wedge c$; then by (11)

$$a \leq a' < c' \leq c, \quad (12)$$

and $b \wedge c' = b \wedge (a \vee b) \wedge c = b \wedge c, a' \vee b = a \vee (b \wedge c) \vee b = a \vee b$. Moreover, $c' \leq a \vee b$, hence $b \vee c' \leq a \vee b \leq b \vee c'$ by (11); therefore $b \vee c' = a \vee b$, and dually, $a' \wedge b = b \wedge c$. This shows that a' and c' have the common complement b in $[b \wedge c, a \vee b]$, and by (12) they are comparable but distinct (cf. Fig. 5). Conversely, if a', c' are distinct elements which are comparable and have a common complement in $[u, v]$, say $a' \wedge b = c' \wedge b = u, a' \vee b = c' \vee b = v$ and $u \leq a' < c' \leq v$, then

$$a' \vee (b \wedge c') = a' < c' = (a' \vee b) \wedge c',$$

hence L is not modular. ■

The property characterizing modularity involves only five elements, namely the endpoints of the interval, an element and its two complements. Thus we have

COROLLARY 1.4 *A lattice is modular if and only if it does not contain a sublattice isomorphic to the pentagon lattice in Fig. 6.* ■

For example, the lattice of Fig. 1 is modular, but not that of Fig. 2.

If A is any set, then a lattice of subsets of A , i.e. a sublattice of $\mathcal{P}(A)$, is modular,

but not every modular lattice can be represented as a lattice of subsets. For example, the Klein 4-group gp $\{a, b \mid a^2 = b^2 = (ab)^2 = 1\}$ has the subgroup lattice shown in Fig. 7 (the ‘diamond’ lattice), but it cannot be represented as a lattice of subsets, as the reader can discover by a little experimentation.

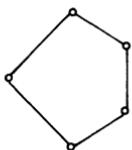


Fig. 6

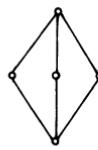


Fig. 7

It turns out that the lattices $\mathcal{P}(A)$ satisfy a further law not holding in the diamond lattice, namely the distributive law. There are two dual forms of this law, which however are equivalent; moreover they imply the modular law. Before proving this fact we note that in any lattice we have $a \vee b \geq a \wedge c$, $b \wedge c$ and $c \geq a \wedge c$, $b \wedge c$ for any a, b, c ; hence

$$(a \vee b) \wedge c \geq (a \wedge c) \vee (b \wedge c). \quad (13)$$

Now the distributive law is expressed by equality in (13). More precisely we have the following:

PROPOSITION 1.5 *In any lattice L the following conditions are equivalent:*

- (α) $(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c)$ for all $a, b, c \in L$;
- (α^*) $(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)$ for all $a, b, c \in L$;
- (β) $(a \vee b) \wedge c \leq a \vee (b \wedge c)$ for all $a, b, c \in L$.

Proof. If (α) holds, then

$$(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c) \leq a \vee (b \wedge c),$$

i.e. (β). Conversely, assume (β): $(a \vee b) \wedge c \leq a \vee (b \wedge c)$. Applying $\wedge c$ to both sides and using (β) again, we obtain

$$(a \vee b) \wedge c \leq [(b \wedge c) \vee a] \wedge c \leq (b \wedge c) \vee (a \wedge c).$$

The reverse inequality holds by (13), hence we obtain (α). Thus $(\alpha) \Leftrightarrow (\beta)$, and since (β) is self-dual, we also have $(\alpha^*) \Leftrightarrow (\beta)$. ■

A lattice satisfying the three equivalent conditions of this proposition is said to be *distributive*; specifically, either (α) or (α^*) is called the *distributive law*. From (β) it is clear that every distributive lattice is modular.

There is a criterion for distributivity analogous to Prop. 1.3: a lattice is distributive iff relative complements in any interval are unique. We shall only need the necessity of this condition.

PROPOSITION 1.6 *In a distributive lattice, relative complements in any interval are unique.*

Proof. Let $a \wedge b = a' \wedge b = u$, $a \vee b = a' \vee b = v$; then

$$a = a \wedge v = a \wedge (a' \vee b) = (a \wedge a') \vee (a \wedge b) = a \wedge a';$$

hence $a \leqslant a'$. By symmetry $a' \leqslant a$ and hence $a' = a$. ■

If we single out the five elements involved we obtain the following alternative formulation:

A lattice is distributive if and only if it does not contain a sublattice isomorphic to the pentagon lattice in Fig. 6 or the diamond lattice in Fig. 7.

For a proof of this criterion see Birkhoff (1967) or Cohn (1981) (see also Ex. (12) below).

Exercises

- (1) In any lattice, if $a \leqslant a'$, $b \leqslant b'$, then $a \vee b \leqslant a' \vee b'$, $a \wedge b \leqslant a' \wedge b'$.
- (2) Find the smallest partially ordered set in which any two elements have an upper bound and a lower bound but which is not a lattice.
- (3) Find all lattices on at most five elements. Which of them are anti-isomorphic with themselves? Which admit a non-identity automorphism? Which are modular, or distributive?
- (4) Show that the least element 0 in a lattice (if it exists) is characterized by $0 \wedge x = 0$, $0 \vee x = x$, and give a corresponding characterization of the greatest element.
- (5) Let L be a modular lattice. Show that if $a, b, c \in L$ satisfy $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, then the sublattice generated by a, b, c is distributive.
- (6) Show that in any modular lattice the sublattice generated by two chains is distributive.
- (7) Let L be a system with two binary operators \vee , \wedge and a particular element 1 in L satisfying (i) $a \wedge a = a$, (ii) $a \vee 1 = 1 \vee a = 1$, (iii) $a \wedge 1 = 1 \wedge a = a$, (iv) $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, $(b \vee c) \wedge a = (b \wedge a) \vee (c \wedge a)$ for all $a, b, c \in L$. Show that L is a distributive lattice with greatest element 1.
- (8) Show that in a topological space the closed sets form a complete distributive lattice.
- (9) (O.Ore) Show that, for any group G , $\text{Lat}(G)$ is distributive iff G is locally cyclic (i.e. any finitely generated subgroup is cyclic).

- (10) Let $G = \text{Sym}_4$, $H = \{1, (1\ 2)\}$, $K = \{1, (1\ 2\ 3\ 4)\}$; show that G is generated by H and K , and deduce that $H \vee K \neq HK$, where $H \vee K$ is the join of H and K in $\text{Lat}(G)$. Show that $\text{Lat}(G)$ is not modular.
- (11) Show, by examining the lengths of maximal chains in $\text{Lat}(\text{Alt}_4)$, that this lattice is not modular.
- (12) Show that in any lattice L ,

$$(x \wedge y) \vee (y \wedge z) \vee (z \wedge x) \leq (x \vee y) \wedge (y \vee z) \wedge (z \vee x) \quad \text{for all } x, y, z \in L, \quad (\text{i})$$

with equality iff L is distributive. Denote the two sides of (i) by u, v respectively and put $x' = (x \wedge v) \vee u$, $y' = (y \wedge v) \vee u$, $z' = (z \wedge v) \vee u$. If L is modular but not distributive, choose $x, y, z \in L$ such that the inequality (i) is strict and verify that u, v, x', y', z' form a diamond lattice. Deduce that a modular lattice is distributive iff it does not contain a diamond lattice as sublattice.

2.2 Chain conditions

Although most of the lattices we shall meet are infinite, many of them satisfy finiteness conditions; these take several equivalent forms. We state them for any partially ordered set:

PROPOSITION 2.1 *In any partially ordered set S the following conditions are equivalent:*

(a) (*Ascending chain condition*) *Every ascending chain becomes stationary: if*

$$a_1 \leq a_2 \leq \cdots, \quad (1)$$

then there exists n_0 such that $a_m = a_n$ for all $m, n \geq n_0$.

(b) *Every strictly ascending chain terminates: if*

$$a_1 < a_2 < \cdots, \quad (2)$$

then the chain has only finitely many terms.

(c) (*Maximum condition*) *Every non-empty subset of S has a maximal element.*

Proof. (a) \Rightarrow (b) follows because any chain (2) can become stationary only by terminating. To prove (b) \Rightarrow (c), let M be a non-empty subset of S . Pick $a_1 \in M$; if a_1 is not maximal in M , we can find $a_2 \in M$ such that $a_2 > a_1$, and generally, for each $a_n \in M$, either a_n is maximal or there exists $a_{n+1} \in M$ such that $a_{n+1} > a_n$. Thus we obtain a chain (2) which must terminate, by (b), and the last element is maximal in M .

(c) \Rightarrow (a). Given (1), let a_n be maximal in the set $\{a_1, a_2, \dots\}$, then $a_n \geq a_m$ for all m , hence $a_n = a_{n+1} = \cdots$, so (1) becomes stationary. ■

We note that the axiom of choice was used in the deduction (b) \Rightarrow (c); it can be shown that this is indispensable. Thus without the axiom of choice the maximum

condition is stronger than the ascending chain condition, but in the presence of the axiom of choice both are equivalent (cf. Hodges 1974).

There is a useful induction principle holding in sets with maximum condition:

PROPOSITION 2.2 (Noetherian induction) *Let S be a partially ordered set with maximum condition. If X is a subset of S which contains any element a of S whenever it contains all elements $x \in S$ such that $x > a$, then $X = S$.*

For, consider the complement X' of X in S . If $X' \neq \emptyset$, let c be a maximal element of X' . Then any $x > c$ must be in X , hence by hypothesis, $c \in X$, which contradicts the fact that $c \in X'$. Therefore X' is empty, and $X = S$, as claimed. ■

By duality we obtain from Prop. 2.1 the equivalence of the minimum condition and (two forms of) the descending chain condition, and as in Prop. 2.2 we obtain an induction principle for sets with minimum condition. For well-ordered sets, i.e. totally ordered sets with minimum condition, this is just the principle of transfinite induction.

Every finite ordered set clearly satisfies both maximum and minimum conditions; the converse is false, as any infinite set shows whose elements are all incomparable. Even for modular lattices the converse need not hold (cf. Ex.(2)), but it does hold for distributive lattices, as we shall see in 2.3. For the moment we shall describe the modular lattices satisfying both chain conditions. Given a chain C between certain points p, q , any chain from p to q which includes C is called a *refinement* of C ; clearly C is a maximal chain from p to q iff it has no proper refinement.

PROPOSITION 2.3 *In a partially ordered set S with both chain conditions, every chain is finite and can be refined by inserting further terms to a maximal chain between the given endpoints.*

Proof. By the minimum condition, every chain in S has a minimal element, necessarily unique. Given a chain C in S , let a_1 be its least element, and generally define a_v as the least element of $C \setminus \{a_1, \dots, a_{v-1}\}$. Then

$$a_1 < a_2 < \dots, \tag{3}$$

and by the maximum condition this chain terminates. If the last term is a_n , it follows that $C = \{a_1, \dots, a_n\}$, hence C is finite. Next, given a chain (3), let b_1 be minimal in S such that $a_1 < b_1 \leq a_2$. If $b_1 < a_2$, we choose $b_2 \in S$ such that b_2 is minimal subject to $b_1 < b_2 \leq a_2$. Continuing in this way, we obtain a chain

$$a_1 < b_1 < b_2 < \dots \leq a_2,$$

which cannot be refined further. By the maximum condition it must terminate, which can happen only when $b_k = a_2$ for some k . We now have a maximal chain

from a_1 to a_2 ; by induction on the number of terms in (3) we can find a maximal chain from a_2 to a_n , and together with the part found this provides a maximal chain from a_1 to a_n . ■

Let us define the *length* of a chain as the number of its links, thus

$$a = a_0 < a_1 < \cdots < a_n = b \quad (4)$$

has length n . In general there is no reason why two maximal chains between given endpoints should have the same length. For example, in the pentagon lattice of Fig. 6 there are two maximal chains in the lattice of lengths 2 and 3. But if we have two chains between points a and b in a modular lattice, then the proof of the Schreier refinement theorem (Vol. 1, p. 257) shows that they have refinements whose links can be paired off in such a way that corresponding links are projective. In particular, if both chains are maximal, it follows that both have the same length. Thus we have

PROPOSITION 2.4 *Let L be a modular lattice and $a \leq b$ in L such that there is a maximal chain from a to b , of length n . Then every chain between a and b has length at most n , and every maximal chain between a and b has length exactly n .* ■

In view of this result we can define the *length* of a modular lattice as the supremum of the lengths of its chains. Modular lattices of finite length can be described as follows:

COROLLARY 2.5 *A modular lattice has finite length if and only if it satisfies both chain conditions.*

Proof. Clearly any modular lattice of finite length satisfies both chain conditions. Conversely, if a modular lattice L satisfies both chain conditions, take a maximal element c in L . Then $x \leq c \vee x = c$ for any $x \in L$, hence c is in fact the greatest element 1 in L , and dually L has a least element 0. By Prop. 2.4, there is a maximal chain from 0 to 1; this is finite, of length n say, and no chain can be longer. ■

A modular lattice has finite length if there is an element c such that all chains below c are finite, and likewise all chains above c . But sometimes we shall want the corresponding assertion when only one of the chain conditions holds.

PROPOSITION 2.6 *Let L be a modular lattice and $c \in L$. Denote by L_c , ${}_c L$ the sublattices of elements $\leq c$ and $\geq c$ respectively. If both L_c and ${}_c L$ satisfy the maximum condition, then so does L .*

Proof. Let

$$a_1 \leq a_2 \leq \cdots \quad (5)$$

be an ascending chain in L . Then $a_1 \wedge c \leq a_2 \wedge c \leq \dots$, and by hypothesis this becomes stationary. Likewise for $a_1 \vee c \leq a_2 \vee c \leq \dots$; thus we may choose n_0 such that for $m, n \geq n_0$

$$a_m \wedge c = a_n \wedge c = u \quad \text{say}, \quad a_m \vee c = a_n \vee c = v \quad \text{say}.$$

Hence a_m, a_n have a common complement c in $[u, v]$, and since $a_m \leq a_n$ for $m \leq n$, it follows by Prop. 1.3 that $a_m = a_n$ for $m, n \geq n_0$, so (5) becomes stationary, as claimed. ■

The results of this section can be applied to modules over a ring by observing that, for any R -module M , the set of all submodules forms a lattice, $\text{Lat}(M)$ say, under the partial ordering by inclusion, with intersection as meet and sum (not union) as join. As we have seen in 2.1 (cf. also Vol. 1, p. 256), this lattice is modular, and the Jordan–Hölder theorem and Schreier refinement theorem proved in Vol. 1 were based on this fact.

An R -module M is said to be *Noetherian* (after E. Noether) if $\text{Lat}(M)$ satisfies the maximum condition, and *Artinian* (after E. Artin) if $\text{Lat}(M)$ satisfies the minimum condition. If we apply this terminology to R itself, regarded as a right R -module, we obtain the notion of a *right Noetherian* or *right Artinian* ring, as a ring with maximum or minimum condition respectively on right ideals. Left Noetherian and Artinian rings are defined similarly, using left ideals.

The maximum condition for modules can be stated in another way:

PROPOSITION 2.7 *An R -module M is Noetherian if and only if all its submodules are finitely generated.*

Proof. \Rightarrow . Let N be any submodule of M and choose $x_1, x_2, \dots \in N$ such that on writing $N_i = x_1R + \dots + x_{i-1}R$, we have $x_i \notin N_i$. Then

$$0 = N_1 \subset N_2 \subset \dots; \tag{6}$$

by the ascending chain condition this must break off, at N_r say, and this means that $N = N_r$, so N is finitely generated.

\Leftarrow . Given an ascending chain (6) of submodules in M , write $N = \sum N_i$ and choose a finite generating set x_1, \dots, x_r of N . If $x_i \in N_{v_i}$ and $v = \max\{v_1, \dots, v_r\}$, then $x_i \in N_v$ for $i = 1, \dots, r$, hence $N_v = N$ and (6) breaks off; this shows M to be Noetherian. ■

If a ring R is right Noetherian (or Artinian), then any cyclic right R -module, being of the form R/a for some right ideal a of R , is again Noetherian (or Artinian), because the submodules of R/a are in natural (order-preserving) bijection with the right ideals of R containing a , by the third isomorphism theorem (4.1 below and Vol. 1, pp. 253, 303). But we can say more than this:

THEOREM 2.8 *Let R be a right Noetherian (resp. Artinian) ring. Then any finitely generated right R -module is again Noetherian (resp. Artinian).*

Proof. We have just seen that the result holds for cyclic modules. Now use induction on the number of generators of M : let M be generated by n elements and denote by M' the submodule generated by $n - 1$ of these. Then $M'' = M/M'$ is cyclic and we have the exact sequence

$$0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0.$$

Assume that R is right Noetherian; then $\text{Lat}(M')$ satisfies the maximum condition, by the induction hypothesis, and $\text{Lat}(M'')$ satisfies the maximum condition because M'' is cyclic. Now the lattice of submodules of M containing M' corresponds to $\text{Lat}(M'')$ (by the third isomorphism theorem) and so satisfies the maximum condition; hence by Prop. 2.6, $\text{Lat}(M)$ satisfies the maximum condition and this proves M to be Noetherian. The Artinian case is proved similarly. ■

In a principal right ideal domain every right ideal is certainly finitely generated, hence we have

COROLLARY 2.9 *Over a principal right ideal domain every finitely generated right module is Noetherian.* ■

A finitely generated module over any ring has an important maximality property, which follows from Zorn's lemma (in the form Prop. 1.2.4).

PROPOSITION 2.10 *Let R be any ring and M a finitely generated R -module. Then every proper submodule of M is contained in a maximal proper submodule.*

Proof. Let M be generated by u_1, \dots, u_r . Given a proper submodule N , consider the set \mathcal{C} of all subsets X of M which together with N generate a proper submodule. A set X_1 fails to lie in \mathcal{C} precisely if u_1, \dots, u_r each is a linear combination of elements of X_1 and N , and this condition involves only finitely many elements of X_1 . Therefore X_1 lies in \mathcal{C} precisely when all its finite subsets lie in \mathcal{C} so \mathcal{C} is of finite character. Hence there is a maximal subset X in \mathcal{C} ; the submodule generated by $X \cup N$ again lies in \mathcal{C} and so by maximality is equal to X , and now X is the desired maximal proper submodule containing N . ■

This result may be applied to R itself. As right R -module, R is generated by the single element 1 and the submodules are right ideals. By a *maximal* right ideal is meant a right ideal maximal in the set of all proper right ideals. Similarly for left ideals and for two-sided ideals, regarding R as R -bimodule. Then we obtain

THEOREM 2.11 (Krull's theorem) *Let R be any ring. Then any proper right ideal is contained in a maximal right ideal. In particular, every non-trivial ring has maximal right ideals. A corresponding result holds for left ideals, and for two-sided ideals.* ■

This result makes essential use of the existence of a unit element in rings; in the absence of 1 it need not hold (cf. Ex. (9)).

In lattices with a chain condition there is a decomposition lemma which is frequently used. It is convenient to formulate it more generally for ordered monoids (cf. 3.1 in Vol. 1, or Ch. 11 below). By a *partially ordered monoid* we understand a monoid which is partially ordered as a set and such that $x \leqslant x'$, $y \leqslant y'$ implies $xy \leqslant x'y'$. We shall further require the condition

$$xy \leqslant x, \quad xy \leqslant y. \quad (7)$$

Examples are lattices with 1 and with $x \wedge y$ as operation, or \mathbb{N} with the ordering opposite to the usual one. An element c in such a monoid is called *irreducible* if $c \neq 1$ and c cannot be written as a product of two elements that are $> c$.

LEMMA 2.12 (decomposition lemma) *Let M be a partially ordered monoid satisfying the maximum condition and (7). Then every element of M can be written as a product of irreducible elements.*

Proof. Denote by I the subset of elements of S which cannot be expressed as a product of a finite number of irreducible elements; we have to show that I is empty. If $I \neq \emptyset$, take a maximal element c in I ; then c cannot be irreducible and $c \neq 1$ because 1 is the product of the empty family. Hence $c = ab$ for some $a > c$, $b > c$. By the maximality of c , a and b are products of irreducible elements, say $a = a_1 \cdots a_r$, $b = b_1 \cdots b_s$ and it follows that $c = a_1 \cdots a_r b_1 \cdots b_s$. This contradicts the choice of c and it shows that $I = \emptyset$. Thus every element of S is a product of irreducible elements. ■

Examples

1. Let L be a lattice with maximum condition. Then L has a greatest element 1 and it satisfies the hypothesis of the lemma with respect to the operation $x \wedge y$. In this case the irreducible elements are called *meet-irreducible* and the lemma tells us that in L every element can be written in the form

$$c = a_1 \wedge \cdots \wedge a_r, \quad \text{where } a_i \text{ is meet-irreducible.}$$

2. Dually, if the operation is $x \vee y$, the irreducible elements are called *join-irreducible*. Hence in a lattice with minimum condition every element can be written as

$$c = b_1 \vee \cdots \vee b_r, \quad \text{where } b_i \text{ is join-irreducible.}$$

3. In \mathbb{N} the minimum condition holds for the usual ordering and $xy \geq x, y$; the irreducible elements in this case are the prime numbers, so we can apply the lemma to deduce that every natural number can be written as a product of prime numbers. Later in Ch. 9 we shall apply Lemma 2.12 in a similar situation, which generalizes this case.

In a partially ordered set with minimum condition there is a relation between anti-chains and lower segments which is sometimes useful.

PROPOSITION 2.13 *Let S be a partially ordered set with minimum condition. Then there is a natural bijection between lower segments and anti-chains: To each lower segment L there corresponds the anti-chain L° consisting of all minimal elements of the complement L' of L ; to each anti-chain A there corresponds the complement of the upper segment $[A]$ generated by A , $A^\circ = [A]'$.*

Proof. From the definitions it is clear that for any lower segment L we have $L \subseteq L^{\circ\circ}$. If $x \notin L$, then $x \in L'$, hence by the minimum condition there is a minimal element a of L' such that $a \leq x$, so $x \in [L^\circ]$ and therefore $x \notin L^{\circ\circ}$. Now take an anti-chain A ; then it is again clear that $A \subseteq A^{\circ\circ}$. To prove equality here we note that $A^\circ = [A]'$; hence for any $x \in A^{\circ\circ}$, x is a minimal element of $A^\circ = [A]$. By definition this means that $x \geq a$ for some $a \in A$, but by the minimality of x , $x = a$, so $x \in A$, as claimed. ■

Our final result does not strictly deal with a chain condition but it is a construction arising from an ascending chain of groups.

PROPOSITION 2.14 *Let $\{G_n\}$ be a sequence of groups such that G_n is a subgroup of G_{n+1} . Then the union $\bigcup G_n$ is a group with each G_n as subgroup.*

Proof. On the union $G = \bigcup G_n$ we can define a group structure in just one way. Namely if $x, y \in G$, then $x \in G_r$, $y \in G_s$ where $r \leq s$ say. Hence $x, y \in G_s$ and the product xy is defined in G_s . Moreover, for any $n \geq s$, G_s is a subgroup of G_n , so the product xy is the same in G_n as in G_s . In this way we define a multiplication on G and it is easily checked that G is a group with respect to this multiplication, with each G_n as subgroup. ■

It is clear that a corresponding result holds for sequences of rings, or of fields.

Exercises

- (1) Show that a partially ordered set in which each chain has at most m and each anti-chain at most n elements, has at most mn elements.
- (2) Given an example of an infinite modular lattice of length 2.

- (3) Let M be a finitely generated module (over any ring). Show that the union of any countable strictly ascending sequence of submodules without last term is a proper submodule.
- (4) Show that a module (over any ring) is finitely generated iff it cannot be expressed as a union of a well-ordered chain without last term of proper submodules. Deduce another proof of Props. 2.7 and 2.10. Does the result still hold if instead of chains we take countable ascending sequences of proper submodules?
- (5) Show that a lattice in which every lower segment is principal satisfies the maximum condition.
- (6) Show that every quotient of a finitely generated module is finitely generated, but not necessarily every submodule.
- (7) Give an example of a non-zero \mathbf{Z} -module without proper maximal submodules.
- (8) Let R be a non-trivial ring such that every right R -module is free. Show that R is a skew field. (*Hint.* See Th. 4.4.8.)
- (9) Show that the ‘ring without 1’ defined on the additive group of rational numbers by the multiplication $xy = 0$ has no maximal ideals.
- (10) Prove without the axiom of choice that in a module with maximum condition all submodules are finitely generated.

2.3 Boolean algebras

We have seen that in any distributive lattice complements in an interval, when they exist, are unique. Thus if L is a distributive lattice with 0 and 1 in which every element has a complement, we can regard the process of associating with each element x its complement x' as a *unary* operator (i.e. an operator with one argument). A complemented distributive lattice with greatest and least element is called a *Boolean algebra* (after G. Boole). Thus a Boolean algebra is a set with two binary operators \vee , \wedge , a unary operator ' and constants 0, 1, satisfying the laws (1)–(4) of 2.1 and the distributive law, as well as the equations

$$x \wedge x' = 0, \quad x \vee x' = 1. \tag{1}$$

Clearly $0' = 1$, $1' = 0$, and since x' is the unique complement of x ,

$$x'' = x. \tag{2}$$

We note the following consequence, known as *De Morgan's laws*:

$$(x \wedge y)' = x' \vee y', \quad (x \vee y)' = x' \wedge y'. \tag{3}$$

For we have $(x' \vee y') \wedge (x \wedge y) = [x' \wedge (x \wedge y)] \vee [y' \wedge (x \wedge y)] = 0$, and

$(x' \vee y') \vee (x \wedge y) = [(x' \vee y') \vee x] \wedge [(x' \vee y') \vee y] = 1$. Hence $x' \vee y'$ is the complement of $x \wedge y$ and we obtain the first of eqns. (3); the second follows by duality.

We also note

$$x \leqslant y \Leftrightarrow x \wedge y' = 0 \Leftrightarrow x' \vee y = 1. \quad (4)$$

For if $x \leqslant y$, then $x \wedge y' \leqslant y \wedge y' = 0$, hence $x \wedge y' = 0$. Conversely, if $x \wedge y' = 0$, then $x \vee y = (x \vee y) \wedge 1 = (x \vee y) \wedge (y' \vee y) = (x \wedge y') \vee y = y$, and so $x \leqslant y$. This proves the first equivalence in (4); the second follows by applying (3).

Examples

1. The set of all subsets $\mathcal{P}(X)$ of any set X is a Boolean algebra if we put $0 = \emptyset$, $1 = X$ and for $Y \in \mathcal{P}(X)$ take Y' to be the complement of Y in X . More generally, any system of subsets of X closed under complements and finite unions is a Boolean algebra; the closure under finite intersections follows by (3) and $0, 1$ are present as the empty union and its complement. This is merely a subalgebra of $\mathcal{P}(X)$, also called a *field of sets* in X .

Let X be any infinite set. Then the subalgebra of $\mathcal{P}(X)$ generated by all the finite subsets of X consists of all the finite subsets of X and their complements, called the *cofinite* subsets of X .

We remark that a sublattice of $\mathcal{P}(X)$ need not be a Boolean algebra, because it may not be closed under complements.

2. In any interval $[a, b] = I$ of a distributive lattice L , the set of elements of I which have a complement in I forms a Boolean algebra.

3. The 2-element lattice is a Boolean algebra. If the elements are a, b and $a \vee b = b$ say, then $a < b$, and we get a Boolean algebra by putting $a = 0, b = 1, 0' = 1, 1' = 0$. This lattice will usually be denoted by **2**. The 1-element lattice is also a Boolean algebra, called the *trivial* algebra; it is usually excluded from consideration. Thus **2** is the smallest non-trivial Boolean algebra.

4. Let R be any ring. Then the set $\mathcal{B}(R)$ of all central idempotents of R (i.e. elements x in the centre such that $x^2 = x$) is a Boolean algebra under the operations

$$x \wedge y = xy, \quad x \vee y = x + y - xy, \quad x' = 1 - x.$$

This example will be taken up again in Prop. 3.1.

5. The set of all propositions in logic (cf. 1.1 in Vol. 1) forms a Boolean algebra, taking \wedge, \vee to be conjunction and disjunction respectively, and x' the negation $\neg x$, and for 1, 0 propositions T, F known to be true and false respectively (e.g. T = ‘someone is looking at this page now’ and $F = T'$). As we saw in Vol. 1, each proposition A has a truth-value $f(A)$, and it is easily verified that f is a homomorphism of the Boolean algebra of all propositions into **2**.

The next result shows that every Boolean algebra may be obtained as the algebra of idempotents of some commutative ring R ; in fact R may be chosen to consist

entirely of idempotents. Let us define a *Boolean ring* as a ring satisfying the identity

$$x^2 = x. \quad (5)$$

Such a ring is necessarily commutative. For we have $2x = (2x)^2 = 4x^2 = 4x$, hence $2x = 0$ and so

$$x = -x. \quad (6)$$

Next we have $x + y = (x + y)^2 = x^2 + xy + yx + y^2$; hence $xy + yx = 0$, i.e. by (6)

$$xy = yx.$$

Thus every Boolean ring is commutative and of characteristic 2.

PROPOSITION 3.1 (i) *Given any ring R , the set $\mathcal{B}(R)$ of its central idempotents forms a Boolean algebra relative to the operations*

$$x \wedge y = xy, \quad x \vee y = x + y - xy, \quad x' = 1 - x. \quad (7)$$

(ii) *On any Boolean algebra B define the two operations*

$$x + y = (x \wedge y') \vee (x' \wedge y), \quad xy = x \wedge y. \quad (8)$$

Then B forms a Boolean ring $\mathcal{R}(B)$ relative to these operations.

Moreover, $\mathcal{B}(\mathcal{R}(B)) \cong B$ for any Boolean algebra B and $\mathcal{R}(\mathcal{B}(R)) \cong R$ for any Boolean ring R .

The first operation (8) is called the *symmetric difference* of x and y ; in the case where $B = \mathcal{P}(X)$, $Y + Z$ represents the subset of X consisting of all elements that are either in Y or in Z but not in both (this is also called ‘addition mod 2’).

Proof. The verification that $\mathcal{B}(R)$ is a Boolean algebra relative to the operations (7) and that $\mathcal{R}(B)$ is a Boolean ring relative to the operations (8) is routine and may be left to the reader.

Now let B be a Boolean algebra and put $B_1 = \mathcal{B}(\mathcal{R}(B))$. Since $\mathcal{R}(B)$ is a Boolean ring, B_1 and B are the same set, and both are Boolean algebras with the same definition of \wedge . In B_1 we have $x' = 1 - x$; from (8) and the fact that $\mathcal{R}(B)$ is a Boolean ring we see that $1 - x = (1 \wedge x') \vee (0 \wedge x) = x'$. So $x \wedge y$ and x' agree on B and B_1 , hence so does $x \vee y = (x' \wedge y')$.

Next take a Boolean ring R and put $R_1 = \mathcal{R}(\mathcal{B}(R))$. Again R and R_1 are the same set; both are Boolean rings with the same definition of product, and if $+_1$ denotes the sum in R_1 , then we have $x +_1 y = x(1 - y) + (1 - x)y = x + y$; therefore $R_1 \cong R$, as we had to show. ■

It is clear that the Boolean algebra homomorphisms correspond precisely to ring homomorphisms; we have here an example of a category equivalence (cf. 1.3):

\mathcal{R} and \mathcal{B} are mutually inverse functors defining an equivalence between the categories of Boolean rings and Boolean algebras.

It is clear that the duality holding in general lattices extends to Boolean algebras. Thus from any law holding in Boolean algebras we obtain another law by interchanging \vee , \wedge and 0, 1.

Given Boolean algebras A and B , by a *dual homomorphism* one understands a mapping $f: A \rightarrow B$ such that

$$x'f = (xf)', \quad (x \vee y)f = xf \wedge yf, \quad (x \wedge y)f = xf \vee yf. \quad (9)$$

Dual isomorphisms, etc., are defined similarly. From De Morgan's laws (3) we see that every Boolean algebra admits a dual automorphism, the *natural duality*, defined by the complementation mapping: $x \mapsto x'$. In a non-trivial Boolean algebra this duality has no fixed point, and its square is the identity; this shows that every finite (non-trivial) Boolean algebra has an even number of elements.

The Boolean algebra **2** possesses a remarkable property, its functional completeness: we shall find that every function of n variables on **2** with values in **2** can be expressed as a Boolean polynomial. Here a *Boolean polynomial* in x_1, \dots, x_n is defined by the following rules:

1. Each x_i is a Boolean polynomial.
2. If u is a Boolean polynomial, then so is u' .
3. If u, v are Boolean polynomials, so are $u \vee v$ and $u \wedge v$.

For example, $(x_1 \wedge x'_2) \vee (x''_3 \wedge x_1)$ is a Boolean polynomial. Two Boolean polynomials are said to be *equal* if we can pass from one to the other by applying the laws of Boolean algebras, viz. (1)–(4) of **2.1** and eqns. (1)–(3) above. It is clear that equal Boolean polynomials are equal as functions on any Boolean algebra; below we shall prove a converse: two Boolean polynomials which define the same function on a given non-trivial Boolean algebra must be equal as polynomials. Clearly it will be enough to prove this for the smallest non-trivial algebra **2**. The result is proved by finding a normal form for polynomials, which is shown to be unique; more precisely, we shall find two normal forms dual to each other.

To describe these normal forms, let us define a *minterm* in x_1, \dots, x_n as

$$X_1 \wedge \cdots \wedge X_n, \quad \text{where } X_i \text{ is } x_i \text{ or } x'_i \quad (i = 1, \dots, n). \quad (10)$$

Now a Boolean polynomial f is said to be in *disjunctive normal form* if it has the form

$$f = f_1 \vee \cdots \vee f_r, \quad (11)$$

where each f_λ is a minterm in x_1, \dots, x_n . For example the disjunctive normal form for $x'_1 \wedge x_2$ is $x'_1 \wedge x_2$ while for $x_1 \vee x'_2$ it is $(x_1 \wedge x'_2) \vee (x'_1 \wedge x_2) \vee (x_1 \wedge x_2)$.

Dually we define a *maxterm* in x_1, \dots, x_n as

$$X_1 \vee \cdots \vee X_n, \quad \text{where } X_i \text{ is } x_i \text{ or } x'_i \quad (i = 1, \dots, n), \quad (12)$$

and the *conjunctive normal form* for f is

$$f = f_1 \wedge \cdots \wedge f_r, \quad (13)$$

where each f_λ is a maxterm in x_1, \dots, x_n . For example the conjunctive normal form for $x'_1 \wedge x_2$ is $(x'_1 \vee x_2) \wedge (x'_1 \vee x'_2) \wedge (x_1 \vee x_2)$, while for $x_1 \vee x'_2$ it is $x_1 \vee x'_2$. With these definitions we have

LEMMA 3.2 *Any Boolean polynomial in x_1, \dots, x_n is equal to a polynomial in disjunctive normal form, and likewise to a polynomial in conjunctive normal form.*

Proof. Let p be any Boolean polynomial in x_1, \dots, x_n . By the distributive law we can write p as a disjunction (11), where each f_λ is a conjunction of some of $x_1, \dots, x_n, x'_1, \dots, x'_n$. By the idempotent law we can omit repetitions and if both x_i and x'_i occur in f_λ , then f_λ is 0 and so may be omitted, and x''_i can be replaced by x_i . In order to obtain a disjunction of minterms we order the variables in f_λ (by the commutative law), say in ascending order; if neither x_i nor x'_i occurs, then since $f_\lambda = (f_\lambda \wedge x_i) \vee (f_\lambda \wedge x'_i)$, we can replace f_λ by $f_\lambda \wedge x_i$ and $f_\lambda \wedge x'_i$. In this way we obtain an expression (11) for p , where each f_λ is a minterm in x_1, \dots, x_n , so p has been expressed in disjunctive normal form. The result for conjunctive normal forms follows by duality. ■

Let us consider a minterm (10). It takes on the value 1 for just one set of values in $\mathbf{2}^n$, viz. $a = (a_1, \dots, a_n)$, where

$$a_i = \begin{cases} 1 & \text{if } X_i = x_i, \\ 0 & \text{if } X_i = x'_i. \end{cases} \quad (14)$$

From (14) we see that for each n -tuple $a \in \mathbf{2}^n$ there is just one minterm taking the value 1 at a and 0 elsewhere; this minterm will be denoted by ε_a . Thus ε_a is the characteristic function of the 1-point subset $\{a\}$ of $\mathbf{2}^n$. More generally, if S is any subset of $\mathbf{2}^n$, then its characteristic function can be written as a Boolean polynomial:

$$\chi_S = \bigvee_{a \in S} \varepsilon_a. \quad (15)$$

Now any function $f: \mathbf{2}^n \rightarrow \mathbf{2}$ is the characteristic function of a subset of $\mathbf{2}^n$, namely $1f^{-1} = \{a \in \mathbf{2}^n \mid f(a) = 1\}$. Hence (15) can be used to express f as a Boolean polynomial; moreover this expression is unique, since f is completely determined by $1f^{-1}$. Thus every function on $\mathbf{2}$ can be represented by exactly one Boolean polynomial in disjunctive normal form. A dual argument applies to conjunctive normal forms. Let us sum up our conclusion:

THEOREM 3.3 *Each Boolean polynomial is equal to a unique expression in disjunctive normal form (11) and to a unique expression in conjunctive normal form*

(13) and it defines a function on each Boolean algebra. Moreover, on the Boolean algebra **2**, every function of n variables is given by a Boolean polynomial but on any Boolean algebra of more than two elements there are functions not represented by Boolean polynomials.

We have seen that every Boolean polynomial can be put in disjunctive normal form (Lemma 3.2) and that every function on **2** has a unique expression in disjunctive normal form. This shows that the disjunctive normal form is unique (since it is unique on **2**). Now the number of mappings from 2^n to **2** is 2^{2^n} , so this must be the number of expressions in disjunctive normal form (this is also easily verified directly). But on a Boolean algebra with b elements there are b^{bn} functions of n arguments, so for $b > 2$ not all of them can be written as Boolean polynomials. ■

The fact that on **2** every function is a Boolean polynomial is expressed by saying that **2** is *functionally complete*.

As an example consider the function given by the table

000	001	010	100	011	101	110	111
<hr/>							
0	1	1	0	0	0	0	1

The corresponding polynomial in disjunctive normal form is

$$(x \wedge y' \wedge z) \vee (x' \wedge y' \wedge z) \vee (x \wedge y \wedge z).$$

We note that this may be simplified to give $(y' \wedge z) \vee (x \wedge y \wedge z)$ or also $(x' \wedge y' \wedge z) \vee (x \wedge z)$. Thus uniqueness holds only for the full disjunctive normal form, where each variable occurs in each minterm.

We also note the interpretation of Th. 3.3 in terms of the propositional calculus (Example 5, p. 46). It states that every assignment of truth values to n propositions can be realized by a propositional function of n variables.

An important help in understanding finite Boolean algebras is a representation theorem which shows them to have the form $\mathcal{P}(X)$ for finite sets X . We shall establish this result in the slightly more general context of distributive lattices. Let P be a partially ordered set and denote by P^* the set of its lower segments. Each element c of P defines a principal lower segment $|c| = \{x \in P \mid x \leq c\}$; by identifying c with $|c|$ we may regard P as a subset of P^* . It is clear that P^* is a lattice with union and intersection as its operations; thus it is a sublattice of $\mathcal{P}(P)$ and hence is distributive. In this way every partially ordered set can be embedded in a distributive lattice. It is a remarkable fact that every finite distributive lattice is of this form.

THEOREM 3.4 *Let L be a distributive lattice of finite length. Then there is a finite partially ordered set P , unique up to order-isomorphism, such that $L \cong P^*$.*

Moreover, if P and L correspond in this way, then the length of L equals the number of elements of P .

Proof. Denote by L^* the set of all join-irreducible elements of L , partially ordered by inclusion. By Lemma 2.12, each $c \in L$ can be represented by the set of all join-irreducible elements below it, and the sets of join-irreducible elements occurring in this way are just the lower segments of L^* ; thus L is order-isomorphic to the set L^{**} of these lower segments, as claimed.

To show that P is uniquely determined by P^* we shall verify that $P^{**} \cong P$. Consider $\alpha \in P^*$; by definition, α is a lower segment in P . If a_1, \dots, a_r are the different maximal elements of α , then $x \in \alpha$ iff $x \leq a_1$ or \dots or $x \leq a_r$. Hence

$$\alpha = [a_1] \vee \dots \vee [a_r],$$

and it follows that α is join-irreducible in P^* iff it is principal. Thus P^{**} , the set of join-irreducible elements of P , is just the set of principal lower segments of P . But the latter is order-isomorphic to P , as we saw; therefore $P^{**} \cong P$, as claimed.

Finally, if P and L correspond and P has n elements, then we can form a maximal chain in L by picking a minimal element $a_1 \in P$, next a minimal element a_2 in $P \setminus \{a_1\}$ and so on; therefore each maximal chain in L has length n . ■

It is not hard to see that this correspondence between ordered sets and lattices is a contravariant functor in each direction, providing a duality, i.e. an anti-equivalence, between the category of finite partially ordered sets and order-homomorphisms and finite distributive lattices and lattice-homomorphisms. Here the natural isomorphism from P to P^{**} takes the following form. With each $x \in P$ we associate an element $\hat{x} \in P^{**}$ defined as a mapping $P^* \rightarrow 2$ by

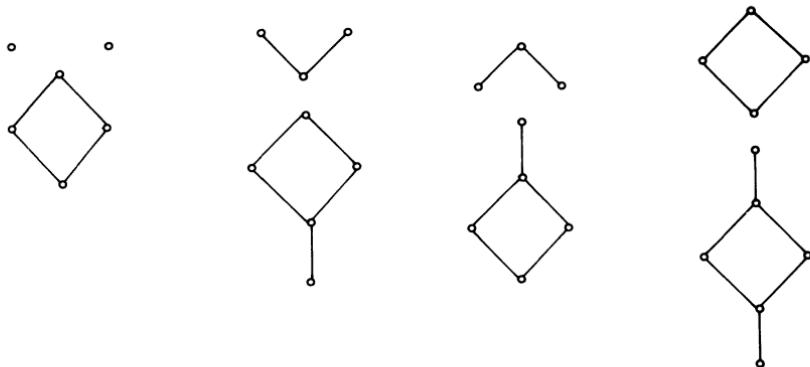
$$\hat{x}(\alpha) = x\alpha \quad \text{for all } \alpha \in P^*.$$

By definition of P^* , if $x \leq y$ in P , then $x\alpha \leq y\alpha$ for all $\alpha \in P^*$, hence $\hat{x}(\alpha) \leq \hat{y}(\alpha)$, and so $\hat{x} \leq \hat{y}$. This shows that the natural mapping $P \rightarrow P^{**}$ is an order-isomorphism. In a similar way it can be shown that the natural mapping $L \rightarrow L^{**}$ is a lattice-isomorphism.

In any pair P, L that correspond, P is generally simpler than L . For example, the free distributive lattice on three generators has length 6 and consists of 18 elements. The corresponding partially ordered set is the three-cornered crown:



Other examples of partially ordered sets and their corresponding lattices are given below:



Let L be a distributive lattice of finite length n . By Th. 3.4, the set L^* of its join-irreducible elements has n elements and L is isomorphic to a sublattice of 2^{L^*} . In particular it follows that L is finite:

COROLLARY 3.5 *Any distributive lattice L of finite length n is a sublattice of $\mathcal{P}(P)$, where P , the set of join-irreducible elements of L , has n elements. In particular, $|L| \leq 2^n$. ■*

The bound can be attained, since for a totally unordered set P of n elements 2^P , the set of all order-homomorphisms from P to 2 , has exactly 2^n elements.

Every Boolean algebra is a distributive lattice and it is natural to ask which partially ordered sets correspond to finite Boolean algebras in the duality of Th. 3.4. This is answered by the next result. To state it let us define an *atom* in a Boolean algebra as an element a such that $a > 0$ but no element x satisfies $a > x > 0$. It is clear that any atom is join-irreducible; conversely, if p is join-irreducible, then $p \neq 0$ and since $p = p \wedge 1 = p \wedge (a \vee a') = (p \wedge a) \vee (p \wedge a')$ for any $a \in L$, we have either $p = p \wedge a$, i.e. $p \leq a$, or $p = p \wedge a'$, i.e. $p \leq a'$, and so $p \wedge a = 0$. Therefore a cannot satisfy $0 < a < p$, and it follows that p is an atom.

THEOREM 3.6 *Let L be a finite distributive lattice and $P = L^*$ the associated partially ordered set. Then L is complemented (and hence a Boolean algebra) if and only if P is totally unordered. Hence any finite Boolean algebra, of length n , has 2^n elements and is isomorphic to $\mathcal{P}(A)$, where A , the set of its atoms, has n elements.*

Proof. If P is totally unordered, every subset of P is a lower segment, hence $P^* \cong \mathcal{P}(P)$, and this is the Boolean algebra of all subsets of P , with 2^n elements. Conversely, if L is a Boolean algebra, then the join-irreducible elements are just its atoms, as we have just seen, and it is clear that the atoms of L form a totally unordered set. ■

Theorem 3.6 (or rather, the Boolean algebra case of Th. 3.4) has been

generalized to arbitrary Boolean algebras by M. H. Stone, whose representation theorem establishes a duality between the category of all Boolean algebras and the category of all Boolean spaces, i.e. totally disconnected Hausdorff spaces. The space corresponding to a Boolean algebra B is the set of homomorphisms $B \rightarrow \mathbf{2}$, as subset of $\mathbf{2}^B$ with the product topology, while the algebra corresponding to a space T is the set of continuous mappings $T \rightarrow \mathbf{2}$, as subset of the product $\mathbf{2}^T$.

Exercises

- (1) In any Boolean algebra show that $(x \vee y) \wedge (x' \vee z) = (x' \wedge y) \vee (x \wedge z)$.
- (2) Show that in (9) the first two formulae imply the third.
- (3) Prove the general distributive law:
$$\left(\bigvee_I a_i \right) \wedge \left(\bigvee_J b_j \right) = \bigvee_{I \times J} (a_i \wedge b_j),$$
for any finite sets I, J .
- (4) Show that any Boolean algebra with ascending chain condition is finite.
- (5) Show that in any Boolean algebra $x + y = (x \vee y) + (x \wedge y)$, where addition has been defined by (8).
- (6) Express the following functions in conjunctive normal form for x, y, z :
 - (i) $(x \wedge y' \wedge z) \vee (x \wedge y \wedge z') \vee (x \wedge y' \wedge z') \vee (x' \wedge y)$;
 - (ii) $(x \wedge y \wedge z') \vee (x' \wedge y' \wedge z) \vee (x \wedge y' \wedge z') \vee (x' \wedge y' \wedge z') \vee (x' \wedge y \wedge z)$; and for x, y, z, u :
 - (iii) $(x' \wedge y \wedge z) \vee (y' \wedge z) \vee (x' \wedge z \wedge u) \vee (z \wedge u')$.
- (7) Verify that for any finite distributive lattice L , the natural mapping $L \rightarrow L^{**}$ is a lattice isomorphism.
- (8) Show that a finite partially ordered set P is order-isomorphic to P^* precisely when it is totally ordered.
- (9) Show that two Boolean algebras with the same (finite) number of elements are isomorphic.
- (10) In a Boolean algebra let $a \leqslant x \leqslant b$; show that $(x' \vee a) \wedge b$ is a relative complement of x in $[a, b]$. Verify that for an element x with a complement x' in a modular lattice this remains true, but not in general lattices.
- (11) Show that a Boolean ring R is simple iff $R \cong \mathbf{2}$. Deduce that an ideal I in a Boolean ring R is a maximal ideal iff $R/I \cong \mathbf{2}$.
- (12) Let B be a Boolean algebra. Show that for $a, b \in B$ there is a homomorphism of B into $\mathbf{2}$

which maps a to 1 and b to 0, unless $a \leq b$. Deduce that B can be embedded in 2^I , for some set I .

2.4 Möbius functions

With each partially ordered set we can associate an algebra, its incidence algebra, which turns out to be useful in answering enumeration questions. The following brief account is based on the elegant treatment by G.-C. Rota (1964).

Let P be a partially ordered set and consider the collection \mathcal{I}_P of square matrices with entries in \mathbf{R} (or any given commutative integral domain), whose rows and columns are indexed by P :

$$A = (a_{ij}), \quad \text{where } a_{ij} = 0 \text{ unless } i \leq j \quad (i, j \in P). \quad (1)$$

For example, if $P = \{a, b, c\}$ with $a < c$, $b < c$, and the matrices of \mathcal{I}_P are indexed by P in the order a, b, c , then \mathcal{I}_P consists of all upper triangular 3×3 matrices (1) with $a_{12} = 0$. For any finite P an easy induction shows that the elements of P can be arranged as a sequence so that i precedes j whenever $i \leq j$, but this is not essential for our purpose (it means in effect that A in (1) can be written as an upper triangular matrix).

If P is finite, so that the matrices in \mathcal{I}_P have finitely many rows and columns, \mathcal{I}_P is closed under the usual addition and multiplication of matrices and so forms a linear algebra over \mathbf{R} , called the *incidence algebra* of P . Thus let $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij}) \in \mathcal{I}_P$; if $C = AB$, then

$$c_{ik} = \sum_j a_{ij}b_{jk}, \quad (2)$$

and here it is enough to confine the summation to the indices j such that $i \leq j \leq k$, by (1). For example, for a totally ordered set P , \mathcal{I}_P is just the set of all upper triangular matrices. We can also allow P to be infinite, provided that each interval $[i, k]$ is finite, for then the summation (2) contains only finitely many non-zero terms, for any pair i, k . A partially ordered set in which all intervals are finite will be called *locally finite*. We begin by noting conditions for a matrix in \mathcal{I}_P to be invertible.

PROPOSITION 4.1 *Let P be a locally finite partially ordered set and \mathcal{I}_P its incidence algebra. Then $A \in \mathcal{I}_P$ is invertible if and only if all its diagonal elements are invertible.*

Proof. Let $A = (a_{ij})$ have the inverse $A^{-1} = (a'_{ij})$. Then by (2),

$$a'_{ii}a_{ii} = 1; \quad (3)$$

hence the condition is necessary. When it holds, we can define a'_{ii} by (3). Given $i, k \in P$, $i < k$, assume that a'_{ij} has already been defined for all j such that $i \leq j < k$;

then we can determine a'_{ik} uniquely from

$$\sum a'_{ij} a_{jk} = 0, \quad (4)$$

for the only unknown term in (4) is a'_{ik} and it occurs with the coefficient a_{kk} , which is invertible by hypothesis. In this way we can determine a'_{ik} for all $i \leq k$, and together with the equation $a'_{ik} = 0$ when $i < k$ this defines $A' = (a'_{ik})$ uniquely to satisfy $A'A = I$. Hence A' is the required inverse. ■

We note that when $A = (a_{ij})$ is given and $A^{-1} = (a'_{ij})$, then a'_{hk} depends only on the values of a_{ij} for i, j in the interval $[h, k]$. As an example for Prop. 4.1, consider the zeta-matrix $Z = (z_{ij})$ defined by

$$z_{ij} = \begin{cases} 1 & \text{if } i \leq j, \\ 0 & \text{otherwise.} \end{cases}$$

By Prop. 4.1, Z has an inverse Z^{-1} , also denoted by M and called the *Möbius matrix*. Its importance stems from its use in the inversion formula.

THEOREM 4.2 (Möbius inversion formula) *Let P be a locally finite partially ordered set with a greatest element. Given any function $f(i)$ on P , define g by the equation*

$$g(i) = \sum_{j \geq i} f(j). \quad (5)$$

Then on writing $Z^{-1} = (m_{ij})$, we have

$$f(i) = \sum m_{ij} g(j) = \sum g(h) m_{hi}. \quad (6)$$

To prove (6) we note that (5) can be written as $g = Zf$, where $f = (f(i))$ is regarded as a column vector, and similarly for g , and the summation is over the interval $[i, \omega]$, where ω is the greatest element of P . It follows that $f = Z^{-1}g = Mg$, and this is the first eqn. (6). The second equation follows similarly, by regarding f, g as row vectors. ■

As an illustration consider the set \mathbb{N} of natural numbers, partially ordered by divisibility; this is infinite, but locally finite, with greatest element 1, if we set $i \leq j$ whenever $j| i$. The Z -matrix is $z_{ij} = 1$ if $j|i$ and 0 otherwise. Since z_{ij} only depends on i/j , we shall write $z_{ij} = \zeta(i/j)$; then $\zeta(n) = 1$ for all $n \in \mathbb{N}$ and $\zeta(\alpha) = 0$ for $\alpha \in \mathbb{Q} \setminus \mathbb{N}$. The inverse $M = (m_{ij})$ can again be written $\mu(i/j)$, where $\mu(n)$, called the *Möbius function*, is given by

$$\mu(n) = \begin{cases} (-1)^r & \text{if } n \text{ is the product of } r \text{ distinct primes,} \\ 0 & \text{if } n \text{ is divisible by the square of a prime.} \end{cases} \quad (7)$$

To establish (7), we note the formula $\sum z_{ij} m_{jk} = \delta_{ik}$; translated to ζ and μ , this

becomes

$$\sum_{rs=n} \zeta(r)\mu(s) = \delta_{n1}.$$

Hence we obtain

$$\mu(1) = 1, \quad \sum_{d|n} \mu(d) = 0 \quad \text{for } n > 1.$$

Given $n > 1$, let us single out a prime p dividing n and write $n = mp^t$, where $p \nmid m$. Then

$$0 = \sum_{d|n} \mu(d) = \sum_{c|m} [\mu(c) + \mu(cp) + \cdots + \mu(cp^t)],$$

and this will be satisfied if for each factor c of m , the sum shown vanishes. Taking $t = 1$ we find $\mu(cp) = -\mu(c)$, and for $t > 1$, $\mu(cp^t) = 0$, and this leads to (7), by an induction on n .

Th. 4.2 in this case reduces to the classical Möbius inversion formula:

Let f be any function on \mathbf{N} . If g is defined by the equation

$$g(n) = \sum_{d|n} f(d),$$

then f is given in terms of g by the formula

$$f(n) = \sum_{d|n} g(d)\mu(n/d), \quad \text{or also} \quad f(n) = \sum_{d|n} \mu(d)g(n/d).$$

The simplest partially ordered set with more than one element is the 2-element set **2**. Its incidence algebra consists of all triangular 2×2 matrices; in particular,

$$Z = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad Z^{-1} = M = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

We can regard **2** as $\mathcal{P}(U)$, where U is a 1-element set. More generally, let S be any finite set and consider $\mathcal{P}(S)$. We assert that for $X \subseteq Y \subseteq S$,

$$m_{XY} = (-1)^{|Y| - |X|}. \quad (8)$$

For if we determine m_{XY} by (4), we have $m_{XY} = -\sum m_{XZ}$, where the summation is over all Z such that $X \subseteq Z \subset Y$. Put $|Y| - |X| = r$ and assume the result for values less than r . There are $\binom{r}{i}$ subsets Z such that $X \subseteq Z \subset Y$ and $|Z| = |X| = i$; therefore by the induction hypothesis,

$$m_{XY} = -1 + \binom{r}{1} - \binom{r}{2} + \cdots + (-1)^r \binom{r}{r-1} = (-1)^r - (1-1)^r = (-1)^r,$$

and so (8) follows.

The Möbius inversion formula of Th. 4.2 can be applied to $\mathcal{P}(S)$ as follows. Let f_X be any function on $\mathcal{P}(S)$ and put

$$g_X = \sum_{Y \supseteq X} f_Y.$$

Then by (6) and (8), we have

$$f_X = \sum_{Y \supseteq X} (-1)^{|Y| - |X|} g_Y. \quad (9)$$

Let us put

$$G_r = \sum_{|Z|=r} g_Z. \quad (10)$$

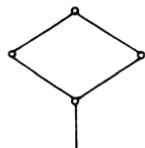
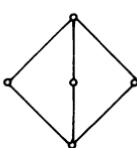
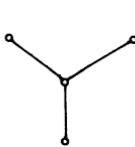
Then on taking $X = \emptyset$ in (9) we find, if $|S| = n$,

$$f_\emptyset = G_0 - G_1 + G_2 - \cdots + (-1)^n G_n. \quad (11)$$

This is known as the *sieve principle* or the *principle of inclusion-exclusion*. If we think of S as a number of properties which a collection of objects may or may not possess, and f_X is the number of objects having precisely the properties in X , then g_X is the number of objects having at least the properties in X (and possibly others). Now (11) states that to find the number of objects having none of the given properties we take the number of objects, subtract, for each property, the number of objects having that property; then for each pair of properties, add the objects having these properties (because they will have been subtracted twice); for each triple of properties, subtract.... To give an example, of two dozen suitors arriving at the Royal Court, 12 were short, 11 were fair, 14 were plain, three were short and fair, seven were short and plain, four were fair and plain, and none were all three. By (10) we see that of the 24 suitors, $24 - (12 + 11 + 14) + (3 + 7 + 4) = 1$ was tall, dark and handsome.

Exercises

- (1) An interviewer questioned 47 people and reported: 22 were male, 18 were married, 19 were retired, five were male and married, four retired and married and two male and retired. How would you test these data for their reliability?
- (2) Show that the Möbius matrix for $\{1, 2, \dots, n\}$ with the natural order is $I - N$, where $N = (n_{ij})$, $n_{ij} = \delta_{ji+1}$.
- (3) Find the Möbius matrices of



- (4) Let S_1, S_2 be two partially ordered sets and $S = S_1 \times S_2$ with the ordering $(x_1, x_2) \leqslant (y_1, y_2)$ iff $x_i \leqslant y_i$ ($i = 1, 2$). Show that if μ_i is the Möbius function for S_i , then the Möbius function μ for S is given by $\mu(i_1, i_2; j_1, j_2) = \mu_1(i_1, j_1)\mu_2(i_2, j_2)$. Deduce another proof of (8).
- (5) A function f on \mathbb{N} is called *multiplicative* if $f(ab) = f(a)f(b)$ whenever a, b are coprime. Prove from the definition that the Möbius function on M is multiplicative, and hence derive (7).
- (6) For any multiplicative function f on \mathbb{N} show that

$$\sum_{d|n} \mu(d)f(d) = \prod_{p|n} [1 - f(p)],$$

where p runs over all the primes.

Show that the Euler function $\varphi(n)$ indicating the number of integers less than and prime to n satisfies $\sum_{d|n} \varphi(d) = n$, and is multiplicative. Deduce that $\varphi(n) = n \prod (1 - p^{-1})$, where the product is taken over all primes p dividing n .

- (7) Prove that the elements of a finite partially ordered set P can be arranged as a sequence a_1, a_2, \dots, a_n so that $a_i < a_j$ implies $i < j$, and show that the number of ways of doing this is equal to the number of maximal chains in P^* (the set of lower segments in P).
- (8) Show that the inverse of the ζ -function $\zeta(s) = \sum n^{-s}$ is $\zeta(s)^{-1} = \sum \mu(n)n^{-s}$, where $\mu(n)$ is the Möbius function.
- (9) Show that for any finite partially ordered set, $Z = I + N$, where N is a nilpotent matrix. Hence obtain the formula $M = I - N + N^2 - \dots$ for the Möbius matrix. What happens in the case of an infinite (but locally finite) set?

Deduce that $m_{ij} = \sum (-1)^v \lambda_v$, where λ_v is the number of chains of length v from i to j and $\lambda_0 = 1$.

- (10) Show that for any locally finite partially ordered set the number of chains from i to j is the (i, j) -entry of $(2 - Z)^{-1}$.

Further exercises on Chapter 2

- (1) Let \mathcal{C} be a family of subsets of a set S and assume that \mathcal{C} is closed under arbitrary intersections. Prove that \mathcal{C} is a lattice, but not in general a sublattice of $\mathcal{P}(S)$.
- (2) Let L be a complete lattice and f an order-homomorphism of L into itself. Define $A = \{x \in L \mid x \leqslant xf\}$; show that if $a = \sup A$, then af is an upper bound for A , and deduce that $af = a$. Show that the fixed points of f again form a complete lattice with respect to the order induced by L . Is this lattice necessarily a sublattice of L ?
- (3) Let A, B be partially ordered sets such that A is order-isomorphic to a lower segment of B and B is order-isomorphic to an upper segment of A . Show that there is a bijection $f: A \rightarrow B$ such that $x < y \Rightarrow xf \not\geqslant yf$ for all $x, y \in A$. Deduce the Schröder–Bernstein theorem (cf.

1.2). (*Hint.* If $g:A \rightarrow B$, $h:B \rightarrow A$ are the given order-homomorphisms and $\mathcal{S}(A)$ is the set of all lower segments of A , define $x\theta = ((xg)'h)'$ for $x \in \mathcal{S}(A)$, where $'$ denotes the complement. Verify that θ is an order-homomorphism of a complete lattice and use Ex. (2).)

- (4) Let A, B be totally ordered sets. Show that if A is order-isomorphic to a lower segment of B and B is order-isomorphic to an upper segment of A , then A, B are order-isomorphic.
- (5) Show that the finite unions of half-closed intervals $[a, b]$, including $(-\infty, b)$ and $[a, \infty)$, form a field of sets in \mathbf{R} .
- (6) A partially ordered set is said to be *graded* if there is a function $f(x)$ with integer values such that $x < y \Rightarrow f(x) < f(y)$ and if y covers x (i.e. $x < y$ and there is no element between x and y) then $f(y) = f(x) + 1$. In any partially ordered set S with least element 0, define the *height function* $h(x)$ as the sup of lengths of chains from 0 to x . Show that all maximal chains between the same endpoints have the same length iff S is graded by h . Moreover, in this case all bounded chains are finite.
- (7) Show that in a modular lattice, if h is defined as in Ex. (6), then $h(x \vee y) + h(x \wedge y) = h(x) + h(y)$. Conversely, show that in a lattice of finite length this condition implies modularity.
- (8) A lattice L is said to be *lower semimodular* if whenever a covers b and c ($b \neq c$), then b and c both cover $b \wedge c$. Show that in a lower semimodular lattice the height function h satisfies $h(x \vee y) + h(x \wedge y) \geq h(x) + h(y)$.
Show that the lattice of subgroups of a finite p -group is lower semimodular. (*Hint.* Use Th. 3 of 9.8 on p. 289 of Vol. 1: Every maximal subgroup of a group of order p^m has order p^{m-1} , and observe that $(B:B \cap C) = (BC:C)$ for any subgroups B, C .)
- (9) Show that a complemented modular lattice is Noetherian iff it is Artinian. Does this remain true if modularity is not assumed?
- (10) A subset of a lattice is called an *ideal* if it is a lower segment and is closed under \vee ; it is *principal* if it has the form $[a] = \{x \in L \mid x \leq a\}$. Show that the ideals form a lattice and the principal ideals form a sublattice isomorphic to L . Show that if L satisfies the ascending chain condition, then every ideal is principal.
- (11) Show that any injective endomorphism of an Artinian module is an automorphism, and dually any surjective endomorphism of a Noetherian module is an automorphism. Give examples to show that Artinian and Noetherian cannot be interchanged.
- (12) Show that a module M is finitely generated iff the set of all finitely generated submodules of M has a maximal element.
- (13) Show that in a finitely generated module M any generating set contains a finite subset which generates M .
- (14) Show that every non-trivial commutative ring has a homomorphism onto a field.

- (15) Let R be a ring which is not right Noetherian. Show that among the right ideals of R that are not finitely generated there is a maximal one.
- (16) Let V be a three-dimensional space over \mathbf{F}_2 and X_i ($i = 1, 2, 3$) the set of all vectors with i -component $\neq 0$. Show that the lattice of subsets generated by X_1, X_2, X_3 has 18 elements (this is the free distributive lattice on three free generators; cf. Ch. 1 of Vol. 3).
- (17) Verify that a subset of a Boolean algebra is an ideal iff it is an ideal of the corresponding Boolean ring. Show that an ideal in $\mathcal{P}(S)$ is maximal iff for each subset X of S the ideal contains either X or its complement. (*Hint.* Use Ex. (11) of 2.3.)
- (18) (E. V. Huntington) Let A be a set with a binary operator \vee and a unary operator ' and define $a \wedge b = (a' \vee b')$. Show that if (i) $a \vee b = b \vee a$, (ii) $a \vee (b \vee c) = (a \vee b) \vee c$ and (iii) $(a \wedge b) \vee (a \wedge b') = a$, then A is a Boolean algebra.
- (19) (M. Sheffer) Let A be an algebra with a binary operator $a|b$ and define $a' = a|a$, $a \vee b = (a|b)', a \wedge b = a'|b'$. Show that if this operation satisfies (i) $(b|a)|(b'|a) = a$, (ii) $[(c'|a)|(b'|a)]' = a|(b|c)$ and (iii) for some $e \in A$ and all $a \in A$, $e|a = a'$, then A is a Boolean algebra with respect to these operations.
- (20) Show that any two countably infinite Boolean algebras without atoms are isomorphic.
- (21) Show that for any partially ordered set P , the set P^* of lower segments is a distributive lattice and there is a natural order-homomorphism $P \rightarrow P^{**}$.
- (22) Let B be a Boolean algebra. Show that $B \cong [0, a] \times [0, a']$, for any $a \in B$. Deduce that $\mathbf{2}$ is the only indecomposable Boolean algebra, and that any finite Boolean algebra has the form $\mathbf{2}^n$.
- (23) Show that in any modular lattice of finite length, $[a \wedge b, a]$ and $[b, a \vee b]$ have the same length. Deduce that if a is the join of n atoms, then $[0, a]$ has length at most n .
- (24) Let P be a finite partially ordered set and $L = P^*$ the corresponding distributive lattice. Show that P and L have isomorphic automorphism groups.
- (25) Let P_n be the set of all rational numbers whose denominators are divisible at most by powers of the first n primes. Show that each P_n is a ring and that P_{n-1} is a subring of P_n . What is $\bigcup P_n$?
- (26) (W. H. Gottschalk) For any Boolean polynomial $f(x) = f(x_1, \dots, x_n)$ show that the operations $\alpha: f(x) \mapsto f(x)'$ and $\beta: f(x) \mapsto f(x')$ are automorphisms. Show that the group generated by α, β is the Klein 4-group.
- (27) Let X be an infinite set and \mathcal{F} the collection of all finite subsets. Show that \mathcal{F} is an ideal in $\mathcal{P}(X)$ and verify that $\mathcal{P}(X)/\mathcal{F}$ has no atoms and is not complete, as a lattice.

(28) A Boolean algebra is called *atomic* if every element > 0 contains an atom. Show that a complete atomic Boolean algebra is isomorphic to $\mathcal{P}(A)$, where A is the set of its atoms; in particular, every finite Boolean algebra is of this form (cf. Th. 3.6).

(29) Show that any generating set of 2^r as Boolean ring has cardinal at least $\log_2 r$, and that this estimate is best possible.

(30) Fix $r \geq 1$ and for $i = 1, \dots, r$ define $e_i: 2^r \rightarrow 2$ as the i th coordinate function; thus if $a = (a_i)$, then $a e_i = a_i$. For any $a \in 2^r$ denote by ε_a the projection of 2^r on the factor corresponding to a . Show that $e_j \varepsilon_a = a_j$ and deduce that every mapping from $\{e_1, \dots, e_r\}$ to a Boolean algebra B extends to a unique homomorphism from 2^r to B . Show also that 2^r is generated by e_1, \dots, e_r . (This is the universal mapping property for 2^r , characterizing it as the free Boolean algebra on e_1, \dots, e_r ; cf. Ch. 1 of Vol. 3.)

(31) Let P be a finite partially ordered set with Möbius matrix M . Show that the Möbius matrix of the set with the opposite ordering (and the same indexing of rows and columns) is M^T , the transpose of M .

(32) (L. Weisner) Let L be a finite lattice and $b > 0$ in L . Show that the Möbius matrix satisfies, for any $a \in L$,

$$\sum_{x \vee b = a} m_{0x} = 0.$$

(33) Let S be a set of n elements and for $X \subseteq S$ let $f(X)$ be the number of permutations of S whose set of fixed points is precisely X . Show that the number of permutations fixing every point of X is $\sum f(Y)$, where the sum is over all $Y \supseteq X$. Deduce that the number of *derangements* of S , i.e. permutations without fixed point, is

$$n! \sum_{v=0}^n (-1)^v / v!.$$

(Thus the probability that a permutation has no fixed point is, for large n , close to e^{-1} .)

(34) Let G be a finite abelian group and consider the Möbius matrix $M = (m_{AB})$ for $\text{Lat}(G)$. Show that m_{AB} vanishes unless $A \subseteq B$, and in that case it depends only on the quotient group B/A . Writing $m_{AB} = \mu(B/A)$, show that if $G = G_1 \times \dots \times G_r$ is the representation of G as the direct product of its primary components, then $\mu(G) = \mu(G_1) \cdots \mu(G_r)$, and for a p -group C , $\mu(C) = (-1)^k p^{k(k-1)/2}$ or 0 according as C is elementary abelian, of order p^k say, or not.

3

Field theory

The study of algebraic equations goes back to antiquity; the Babylonians knew two millenia BC how to solve quadratic equations as well as some particular cases of higher degrees, but the general solution of cubic and quartic equations was not accomplished until the 16th century. This gave rise to much activity to try to solve equations of higher than the fourth degree; here ‘solving’ means giving a formula, involving repeated radicals, for the roots in terms of the coefficients (the custom of using letters for the coefficients had been introduced by F. Viète in 1591). After several abortive attempts it was gradually realized that such a formula might not exist, and this was finally proved by N. H. Abel in 1826. Within a few years E. Galois developed the correspondence between equations and the groups of permutations of their roots, which shed a very clear light on the subject. The theory of Galois forms the core of modern field theory, and it leads to an explicit description of field extensions.

Until Steinitz’s fundamental paper of 1910 the complications arising in finite characteristic had not been contemplated, for although finite fields also went back to Galois, their algebraic extensions offered no difficulty. Today the problems of ‘inseparability’ arise mainly in algebraic geometry over a finite field; our account here will deal only with the basic facts, reserving a fuller treatment for Vol. 3.

3.1 Fields and their extensions

We recall (2.4 and 6.1 in Vol. 1) that a *field* is a ring in which $1 \neq 0$ and every non-zero element has an inverse. It follows that a field F has exactly two ideals, 0 and F , and hence every homomorphism from a field to a non-trivial ring is injective (because the kernel is a proper ideal). Obvious examples of fields are the rational numbers \mathbf{Q} and, for every prime number p , the field $\mathbf{F}_p = \mathbf{Z}/p$ of p elements (Vol. 1, 2.4).

Let us also recall the notion of characteristic, basic in all that follows. In any field F consider the multiples of 1: $1, 1+1, 1+1+1, \dots$. We abbreviate these expressions as $1, 2, 1, 3, 1, \dots$ and remark that two cases can arise:

(i) $n \cdot 1 \neq 0$ for all $n > 0$. Then F is said to have *characteristic 0*. In this case the homomorphism $n \mapsto n \cdot 1$ provides an embedding of \mathbf{Z} in F ; since F is a field, we can

invert the elements $n.1$ and so form the subfield P generated by 1 . This is the *prime subfield* of F ; clearly it is isomorphic to \mathbf{Q} . It is contained in every subfield of F and so is the least subfield.

(ii) For some $n > 0$ we have $n.1 = 0$. Let n_0 be the least positive integer such that $n_0.1 = 0$; clearly $n_0 > 1$. If $n_0 = rs$, then $(r.1)(s.1) = 0$ and $(rs).1 = 0$; hence $r.1 = 0$ or $s.1 = 0$, say the former. Then $r \geq n_0$ by the definition of n_0 and this shows that n_0 is a prime number, p say. This prime number is called the *characteristic* of F in this case; thus for any field the characteristic is either 0 or a prime number. When F has prime characteristic p , the elements $0, 1, 2.1, \dots, (p-1).1$ already form a subfield P , isomorphic to \mathbf{F}_p , which is the prime subfield now. We can sum up our results as follows:

THEOREM 1.1 *Every field F has a least subfield P , the prime subfield of F , which is contained in every subfield of F . Either (i) F has characteristic 0 and $P \cong \mathbf{Q}$, or (ii) F has characteristic p , a prime number, and then $P \cong \mathbf{F}_p$.* ■

We shall sometimes denote the characteristic of F by $\text{char } F$.

Let F be a field and k a subfield; we shall also describe k as the *ground field*, F as an *extension* of k and write* F/k for the field F considered as an extension of k . This way of considering a field is often useful. In the language of 11.7 of Vol. 1 (cf. also 5.1 below) F is a k -algebra and, in particular, it is a vector space over k . Its dimension is called the *degree* of F over k and is written $[F:k]$ or $\dim_k F$; we shall use this notation sometimes even if F is only a vector space. The dimension is a positive integer or ∞ ; for example, $[\mathbf{C}:\mathbf{R}] = 2$, $[\mathbf{R}:\mathbf{Q}] = \infty$, and of course for every field k , $[k:k] = 1$. By a *finite* extension one understands an extension of finite degree. We remark that for every extension F/k , F and k have the same prime subfield, and hence the same characteristic.

Let F/k be a finite extension, of degree n say. Then there is a basis u_1, \dots, u_n for F over k and relative to this basis each element $a \in F$ can be uniquely expressed in the form

$$a = \sum \alpha_i u_i, \quad \text{where } \alpha_i \in k. \quad (1)$$

In order to know F completely we need only know k and the expressions (1) for the n^2 products $u_i u_j$, but it is not particularly easy to work with these n^2 expressions and we shall soon find a simpler way of describing the extension F/k . A *k -homomorphism* (of extensions of k) is a homomorphism which is k -linear, or equivalently, one which reduces to the identity on k .

The following product formula is a basic tool in the study of field extensions.

PROPOSITION 1.2 *Let F/k be a field extension. Given any vector space V over F , we can regard V also as a vector space over k , its dimension $[V:k]$ is finite if and*

*The risk of confusion with quotient rings R/α is small since fields have no non-trivial quotient rings.

only if both $[V:F]$ and $[F:k]$ are finite, and when this is so, we have

$$[V:k] = [V:F][F:k]. \quad (2)$$

Proof. Suppose first that F/k is finite, with a basis u_1, \dots, u_m say, and that V is finite-dimensional over F , with basis v_1, \dots, v_n . For clarity we shall use latin suffixes for the range $1, \dots, m$ of the first basis, and greek suffixes for the range $1, \dots, n$ of the F -basis of V . Thus the u_i form a basis for F/k and the v_λ form an F -basis of V ; we shall prove that the set of products $u_i v_\lambda$ is a k -basis for V . This will prove that $[V:k] = mn$, and (2) follows in this case.

Let $x \in V$; since V is spanned by the v_λ over F , we have

$$x = \sum a_\lambda v_\lambda \quad \text{for some } a_\lambda \in F. \quad (3)$$

We express each a_λ in terms of the u_i : $a_\lambda = \sum \alpha_{\lambda i} u_i$, where $\alpha_{\lambda i} \in k$, and insert these values in (3):

$$x = \sum \alpha_{\lambda i} u_i v_\lambda.$$

This shows that the mn products $u_i v_\lambda$ span V over k . To prove their linear independence, let us assume a relation

$$\sum \alpha_{\lambda i} u_i v_\lambda = 0. \quad (4)$$

On writing $a_\lambda = \sum \alpha_{\lambda i} u_i$, we see that this is of the form $\sum a_\lambda v_\lambda = 0$, where $a_\lambda \in F$, and hence by the linear independence of the v_λ we have $a_\lambda = 0$ ($\lambda = 1, \dots, n$). This means that $\sum \alpha_{\lambda i} u_i = 0$ for all λ , and since the u_i are linearly independent over k , we conclude that $\alpha_{\lambda i} = 0$ for all λ, i . Hence the relation (4) is trivial and it follows that the $u_i v_\lambda$ are linearly independent over k ; this then shows that they form a k -basis for V .

We observe that in order to prove that the $u_i v_\lambda$ are linearly independent we needed only the linear independence of the sets $\{u_i\}$, $\{v_\lambda\}$, not the fact that they were bases. This means that for any m linearly independent elements of F over k and any n linearly independent elements of V over F we can construct mn linearly independent elements of V over k . Suppose now that V is of finite dimension N over k ; then N is a bound for the number of elements in a linearly independent set of vectors in V over k ; hence $mn \leq N$ and it follows that both $[V:F]$ and $[F:k]$ are finite. Thus $[V:k]$ is finite iff both $[V:F]$ and $[F:k]$ are, and when this is so, (2) holds. ■

We note particularly the case where V is an extension field E of F . Thus whenever we have a tower of fields

$$k \subseteq F \subseteq E, \quad (5)$$

then we have the *product formula*:

$$[E:k] = [E:F][F:k], \quad (6)$$

whenever either side is finite.

As in the case of Lagrange's theorem for groups (Vol. 1, p. 53), this has the useful consequence that for any tower of fields as in (5), the degree $[F:k]$ is a divisor of $[E:k]$. For example, an extension E/k of prime degree has no proper subextensions. However, an extension without proper subextensions need not be of prime degree; examples are easily constructed with the help of Galois theory, as we shall see in 3.6.

Consider any field extension F/k . Given elements $\alpha_1, \dots, \alpha_n$ of F , we write $k[\alpha_1, \dots, \alpha_n]$ for the *subring* generated by $\alpha_1, \dots, \alpha_n$ over k and $k(\alpha_1, \dots, \alpha_n)$ for the *subfield* generated by $\alpha_1, \dots, \alpha_n$ over k . For a finite extension we actually have equality:

$$k(\alpha_1, \dots, \alpha_n) = k[\alpha_1, \dots, \alpha_n], \quad (7)$$

for the right-hand side is a finite-dimensional k -algebra without zero divisors, so the multiplication by any non-zero element c is a linear transformation which is injective and hence invertible (cf. also Vol. 1, p. 359). Thus any non-zero element has an inverse and the right-hand side of (7) is a field, necessarily equal to the left-hand side.

Let us take a closer look at a *simple* extension, $n = 1$. Such an extension has the form $k(\alpha)$, where $\alpha \in F$. Consider first the subring $k[\alpha]$; we have a unique k -homomorphism λ from the polynomial ring $k[x]$ to $k[\alpha]$ which maps x to α . This is clearly surjective and it gives rise to the exact sequence

$$0 \rightarrow \ker \lambda \rightarrow k[x] \xrightarrow{\lambda} k[\alpha] \rightarrow 0. \quad (8)$$

There are two cases to consider:

(i) λ is injective; it is then an isomorphism: $k[\alpha] \cong k[x]$. In this case we say that α is *transcendental* over k . The field of fractions $k(\alpha)$ of $k[\alpha]$ is called a *purely transcendental* extension of k .

(ii) λ is not injective; in this case α is called *algebraic* over k . For complex numbers these terms are used with reference to the rational numbers \mathbf{Q} as ground field; e.g. $\sqrt{2}$ and $\sqrt{-1}$ are algebraic numbers, while e and π are transcendental.

We observe that in (8) the kernel of λ represents the set of all polynomials in x which vanish for $x = \alpha$. Thus α is algebraic precisely when it satisfies a non-trivial polynomial equation over k . If \mathfrak{a} is the kernel of λ , then by (8), $k[\alpha] \cong k[x]/\mathfrak{a}$. We claim that $k[\alpha]$ is finite-dimensional over k . For \mathfrak{a} , being a non-zero ideal, contains a polynomial $f = c_0 x^n + c_1 x^{n-1} + \dots + c_n$ which is non-zero, say $c_0 \neq 0$. By definition we have

$$c_0 \alpha^n + \dots + c_n = 0, \quad (9)$$

and since we can divide by c_0 , this equation shows α^n to be linearly dependent on $1, \alpha, \dots, \alpha^{n-1}$ over k . On multiplying (9) by α , we see that α^{n+1} is linearly dependent on $\alpha, \alpha^2, \dots, \alpha^n$ and hence on $1, \alpha, \dots, \alpha^{n-1}$. An easy induction shows that all powers of α are linearly dependent on $1, \alpha, \dots, \alpha^{n-1}$ and since $k[\alpha]$ is spanned by the powers of α over k , it follows that $k[\alpha]$ is finite-dimensional over k . By the

previous remark we see that $k[\alpha]$ is already a field. When (9) holds, α is called a *root* of the equation $f = 0$; it is a *zero* (or also a *root*) of f .

To study the kernel of λ more closely, we recall from **10.6** of Vol. 1 (p. 326) that it is a principal ideal, generated by a monic polynomial. Explicitly, let p be a non-zero polynomial of least degree with α as zero; then any $f \in \ker \lambda$ can be written as $f = pq + r$, where $\deg r < \deg p$. Putting $x = \alpha$, we find $0 = f(\alpha) = p(\alpha)q(\alpha) + r(\alpha) = r(\alpha)$. Thus α is a zero of r , and by the minimality of $\deg p$, r must vanish, so that $f = pq$ and $\ker \lambda = (p)$. Hence $\ker \lambda$ is generated by the polynomial of least degree with α as zero. If this polynomial p has degree n , then $k[\alpha]$ is spanned by $1, \alpha, \dots, \alpha^{n-1}$ and these elements are also linearly independent over k ; for any dependence would give an equation of degree less than n satisfied by α . Thus $1, \alpha, \dots, \alpha^{n-1}$ form a basis of $k[\alpha]$ and we see that $[k[\alpha]:k] = n = \deg p$. We note that p may always be taken monic, and it is then uniquely determined. For if α satisfied two different monic equations of degree n , it would also satisfy their difference, an equation of lower degree, contradicting the minimality of n . The monic polynomial p of least degree satisfied by α is called the *minimal polynomial* for α over k , and $\deg p$ is called the *degree* of α over k . We summarize these results as follows.

PROPOSITION 1.3 *Let E/k be any field extension and $\alpha \in E$. Either*

- (i) α is transcendental over k ; then $k(\alpha) \cong k(x)$, where x is an indeterminate, or
- (ii) α is algebraic over k . In that case $k(\alpha) = k[\alpha]$ and if p is the minimal polynomial for α over k , of degree n , then p is irreducible over k and $k(\alpha) \cong k[x]/(p)$, $[k(\alpha):k] = n$. Moreover, p is uniquely determined as the monic polynomial of least degree satisfied by α .

It only remains to prove that p is irreducible (i.e. an atom in $k[x]$, cf. Vol. 1, p. 153). This follows because $\deg p \geq 1$ and if $p = fg$, then

$$0 = p(\alpha) = f(\alpha)g(\alpha),$$

hence $f(\alpha) = 0$ or $g(\alpha) = 0$; therefore either f or g has degree at least n and the other factor then has degree 0, i.e. it is a unit. Hence p is indeed irreducible. ■

Of course over a larger field p may become reducible; e.g. over $k[\alpha]$, p has the factor $x - \alpha$, and this does not lie in $k[x]$ unless $\alpha \in k$ and p has degree 1.

An extension E/k is called *algebraic* if all its elements are algebraic over k ; otherwise it is called *transcendental* (thus a transcendental extension may well contain algebraic elements). It is clear that every finite extension is algebraic, for if $[E:k] = n$, then for any $\alpha \in E$, the extension $k(\alpha)/k$ has a degree dividing n , by Prop. 1.2. However, not every algebraic extension is finite, as we shall see in **3.3**. But there is a converse for finitely generated extensions:

PROPOSITION 1.4 *Let E/k be an extension generated by a finite number of*

algebraic elements over k . Then $[E:k]$ is finite; in particular, the sum and product of algebraic elements are algebraic.

Proof. Suppose that E is generated by $\alpha_1, \dots, \alpha_r$ over k . By induction on r , $[k(\alpha_2, \dots, \alpha_r):k] < \infty$ and since α_1 is algebraic over k , it is algebraic over $k(\alpha_2, \dots, \alpha_r)$, so $[E:k(\alpha_2, \dots, \alpha_r)] < \infty$. Hence by the product formula (Prop. 1.2) we find that $[E:k] < \infty$, as claimed. It follows that any element of E is algebraic over k ; in particular, this holds for $\alpha_1 + \alpha_2, \alpha_1\alpha_2$. ■

COROLLARY 1.5 *An extension is algebraic if it is generated by algebraic elements.*

For assume that E/k is generated by a set A of algebraic elements, and let $c \in E$. Then c lies in the subextension generated by a finite subset of A , say $c \in k(\alpha_1, \dots, \alpha_r)$, where $\alpha_i \in A$. Since the α_i are algebraic, $k(\alpha_1, \dots, \alpha_r)$ is finite over k and so c is algebraic over k , as claimed. ■

As an application of the product formula (6) we briefly indicate how to prove the impossibility of certain geometrical constructions. In Euclid's *Elements* the only allowable constructions are by ruler and compasses. This enables us to construct from a given length all multiples and their submultiples, i.e. all lengths commensurate with a given length. This means that all rational numbers can be constructed. Next, using the fact that a triangle inscribed in a circle on a diameter as base is right-angled, we can extract square roots of differences of squares (by Pythagoras' theorem); hence we can find arbitrary square roots by the formula

$$c = [(c+1)/2]^2 - [(c-1)/2]^2.$$

Each time we extract a square root, we enlarge our field by an extension of degree 2, if at all. Therefore all extensions reached in this way have as degree a power of 2. It follows that a real number α is constructible by ruler and compasses—it 'admits quadrature'—only when its degree over \mathbb{Q} is a power of 2. This condition, that $[\mathbb{Q}(\alpha):\mathbb{Q}] = 2^m$, though necessary, is not sufficient for constructability; we shall meet the precise condition later, in Ex. (1) of 3.11.

With the help of the above remark several ruler-and-compass constructions proposed in ancient Greece can be shown to be insoluble.

(i) Quadrature of the circle. This is the problem of constructing a square equal in area to the area of a circle of radius 1. What has been said shows that π would have to be algebraic over \mathbb{Q} of degree 2^m , for some $m \geq 0$. In fact it was proved by Lindemann in 1882 that π is transcendental over \mathbb{Q} , so the quadrature of the circle is impossible. The proof that π is transcendental will not be given here, cf. Jacobson (1985) or Lang (1984).

(ii) Duplication of the cube (Delian problem). The oracle at Delos revealed that to avert a plague, it would be necessary to double the size of a certain altar,

built in the shape of a cube. Since a cube of side a has volume a^3 , this entailed the solution of the equation $(ax)^3 = 2a^3$, or $x^3 = 2$. This is an equation of degree 3, irreducible over \mathbf{Q} , so $2^{1/3}$ does not admit quadrature*.

(iii) Trisection of an angle. Let α be a given angle and write $\alpha = 3\beta$. Then we must solve $\cos(3\beta) = \cos \alpha = \lambda$ say, or $4\cos^3 \beta - 3\cos \beta = \lambda$, i.e.

$$4x^3 - 3x - \lambda = 0.$$

For example, for $\lambda = \frac{1}{2}(\alpha = 60^\circ)$ the equation becomes, on putting $y = 2x$,

$$y^3 - 3y - 1 = 0.$$

This equation is irreducible over \mathbf{Q} , as we see by putting $y = z + 1$ and applying Eisenstein's criterion to the resulting equation (cf. Vol. 1, p. 167). It follows that the angle 60° cannot be trisected by ruler and compasses.

Of course, none of these impossibility proofs affect the practical constructibility, which is possible to any degree of accuracy.

Exercises

- (1) Find a basis for $\mathbf{Q}(\sqrt{2}, \sqrt{3})/\mathbf{Q}$. Find the minimal polynomial for $\sqrt{2} + \sqrt{3}$ and use it to show that $\mathbf{Q}(\sqrt{2}, \sqrt{3}) = \mathbf{Q}(\sqrt{2} + \sqrt{3})$.
- (2) Let $k(x)$ be the field of rational functions in an indeterminate x . Show that every element of $k(x)$ which is not in k is transcendental over k .
- (3) Show that every automorphism of a field leaves the prime subfield elementwise fixed.
- (4) Let F/k be a field extension and σ a homomorphism of F into a field F' . Show that $[F^\sigma:k^\sigma] = [F:k]$.
- (5) Show that an endomorphism of F/k , as k -algebra, is a field endomorphism of F leaving k elementwise fixed.
- (6) Show that if F/k is a field extension of finite degree, then every endomorphism of F/k is an automorphism. Give examples to show that this may fail to hold for arbitrary extensions; does it hold for arbitrary algebraic extensions?
- (7) Show that a field extension is algebraic iff every subalgebra is a field.
- (8) Show that for $k \subseteq E \subseteq F$, if F/E and E/k are algebraic, then so is F/k .
- (9) Show that $k(x)/k((x^3 + 2)/(x^2 + x - 1))$, where x is an indeterminate, is algebraic and find a basis.

*It is reported that at first a new altar in the shape of a cube with twice the side of the old altar was built, whereupon the plague got worse!

- (10) Verify in detail that the magnitudes constructible by ruler and compasses are just those obtainable using only square roots and rational operations.
- (11) Give an algebraic proof that every angle can be bisected by using only ruler and compasses.
- (12) Show that in any field extension F/k , the set F_0 of elements of F that are algebraic over k is a subfield.

3.2 Splitting fields

We have seen in 3.1 that every algebraic element over a field k is the zero of a polynomial. Conversely, suppose that we are given a polynomial f over k ; will f have a zero in k , or in a suitable extension of k ? There may be no zero in k itself; for example, the real polynomial $x^2 + 1$ has no real zeros, but it does have a zero in the field of complex numbers. We shall see that essentially the same method used to construct $\sqrt{-1}$ works in the general case.

THEOREM 2.1 (Kronecker) *Let f be a polynomial of positive degree over a field k . Then there exists an extension E/k in which f has a zero.*

Proof. The quotient ring $R = k[x]/(f)$ is a non-zero finite-dimensional k -algebra, its dimension being $\deg f$. Let α be a proper ideal of maximal dimension in R ; then α is a maximal ideal, hence $E = R/\alpha$ is a field, and still a k -algebra. Thus we have a homomorphism $k \rightarrow E$, necessarily injective. By identifying k with its image in E we may regard the latter as an extension of k . If $x \mapsto \alpha$ in the homomorphism $k[x] \rightarrow R \rightarrow E$, then $f(\alpha) = 0$, so f has a zero in E . ■

This proof, depending on the construction of a maximal ideal in $k[x]/(f)$, is not very explicit; in particular, we should like to know when the ideal (f) will itself be maximal, and to what extent the construction is unique. Both questions are answered by the next result.

PROPOSITION 2.2 *Let f be a non-zero polynomial over a field k . Then $k[x]/(f)$ is a field if and only if f is irreducible.*

Given two isomorphic fields k_1, k_2 , let f_1 be an irreducible polynomial over k_1 and f_2 the corresponding polynomial over k_2 . Suppose that f_i has a zero α_i in an extension E_i of k_i ($i = 1, 2$). Then the isomorphism from k_1 to k_2 can be extended to a unique isomorphism between $k_1(\alpha_1)$ and $k_2(\alpha_2)$ such that α_1 maps to α_2 .

Proof. We know that $R = k[x]/(f)$ is finite-dimensional over k , so by Prop. 3 of 11.7, Vol. 1 (see also 5.1 below) it will be a field iff it is an integral domain, which is the case iff (f) is a prime ideal. But this is so precisely when f is prime, i.e.

irreducible (because $k[x]$ is a unique factorization domain (UFD), cf. Vol. 1, p. 158 or **9.2** below).

For the second part let k be a field isomorphic to k_1 and k_2 and let $\varphi_i: k \rightarrow k_i$ be two isomorphisms such that $\varphi_1^{-1}\varphi_2$ is the given isomorphism between k_1 and k_2 . Further denote by f the polynomial over k which corresponds to f_i over k_i under the isomorphism φ_i . The isomorphism φ_1 extends to a unique homomorphism $\varphi'_1: k[x] \rightarrow k_1[\alpha_1]$ that maps x to α_1 . Since $f_1(\alpha_1) = 0$, this homomorphism maps $f(x)$ to 0, and so can be factored by the natural homomorphism $k[x] \rightarrow k[x]/(f)$ to give a homomorphism $\lambda_1: k[x]/(f) \rightarrow k_1[\alpha_1]$, by the factor theorem (Th. 1.3.3).

$$\begin{array}{ccc} & k[x]/(f) & \\ \lambda_1 \swarrow & & \searrow \lambda_2 \\ k_1[\alpha_1] & \xrightarrow{\quad} & k_2[\alpha_2] \end{array}$$

Clearly this is surjective, and since f is irreducible, the left-hand side is a field. Hence $\ker \lambda_1 = 0$, so λ_1 is injective and so is an isomorphism. The residue class of x (mod f), say \bar{x} , maps to α_1 under this isomorphism. Similarly there is an isomorphism $\lambda_2: k[x]/(f) \rightarrow k_2[\alpha_2]$ mapping \bar{x} to α_2 , and therefore $\lambda_1^{-1}\lambda_2$ is an isomorphism from $k_1[\alpha_1]$ to $k_2[\alpha_2]$ extending $\varphi_1^{-1}\varphi_2$ and mapping α_1 to α_2 . It is plainly unique since it is prescribed on a generating set of $k_1[\alpha_1]$. ■

With the help of this result we see more clearly what happens in the construction of a root for a given equation $f = 0$ (Th. 2.1). Given a polynomial f over k , we split off an irreducible factor p , say $f = pg$. Then $F = k[x]/(p)$ is a field, and denoting the residue class of x in F by α , we have $p(\alpha) = 0$, hence $f(\alpha) = p(\alpha)g(\alpha) = 0$. For example, over the real field \mathbf{R} , the polynomial $x^2 + 1$ is irreducible, and it leads to the construction of the complex numbers in the form $\mathbf{C} = \mathbf{R}[x]/(x^2 + 1)$.

We note that the construction of an extension E in which a given polynomial f has a zero is not usually unique, but depends on which irreducible factor of f is chosen. Thus if we are given $f = x^4 - 4$ over \mathbf{Q} , we have the factorization

$$x^4 - 4 = (x^2 + 2)(x^2 - 2),$$

where the factors on the right are irreducible over \mathbf{Q} , and correspondingly there are two non-isomorphic extensions of \mathbf{Q} containing a root of $x^4 - 4 = 0$. On the other hand, the equation $x^4 + 1 = 0$ is irreducible over \mathbf{Q} and so all extensions generated by a root of this equation over \mathbf{Q} are isomorphic.

Over the complex numbers every non-constant polynomial has a zero; this is the content of the ‘fundamental theorem of algebra’, which will be proved in **3.3**. It follows that every irreducible polynomial over \mathbf{C} is linear (cf. **6.7** in Vol 1, where

fields with this property were called ‘algebraically closed’). Now whatever the field, by repeating Kronecker’s construction (Th. 2.1) we can split any polynomial into linear factors. We shall now carry out this process in detail, but first we need a definition.

Given a polynomial f over a field k , suppose that in some extension E/k , f can be expressed as a product of linear factors:

$$f = a_0(x - \alpha_1) \dots (x - \alpha_n) \quad (\alpha_1, \dots, \alpha_n \in E, a_0 \in k^\times). \quad (1)$$

Then we shall say: f splits over E , and E is called a splitting field of f over k . If f splits over E but over no smaller field, then E is called a minimal splitting field of f over k . Given any splitting field E of f , we need only take $k(\alpha_1, \dots, \alpha_n)$, where the α ’s are as in (1), to obtain a minimal splitting field. We now show that splitting fields always exist and the minimal ones are unique up to isomorphism.

THEOREM 2.3 *Let k be a field and f a non-zero polynomial in $k[x]$. Then there exists a splitting field of f over k . If $\deg f = n$, then any minimal splitting field E satisfies $[E:k] \leq n!$.*

Further, let k' be a field isomorphic to k , f' the polynomial over k' corresponding to f over k and E' a minimal splitting field for f' over k' . Then the given isomorphism between k and k' can be extended to an isomorphism between E and E' .

Proof. Existence: We begin by factorizing f over k :

$$f = p_1 p_2 \dots p_r, \quad (2)$$

where each p_i is irreducible over k . If each p_i is linear, then k is itself a splitting field of f ; otherwise we adjoin a zero of some non-linear factor p_i , using Th. 2.1. In the resulting extension we can write p_i as the product of a linear factor and another factor, so we have increased the number of linear factors in a complete factorization of f . If there is another non-linear factor left, we repeat the procedure; after a finite number of steps we obtain an extension in which f splits completely and this is the desired splitting field. To estimate the degree of the extension, at the i th stage we have at least i linear factors, so the irreducible factor to be split has degree at most $n - i$, and the degree increases by a factor of at most $n - i$. We start at stage 0, so the degree is at most $n(n - 1) \dots 2 \cdot 1 = n!$ as claimed.

Uniqueness: We use induction on the degree $[E:k]$. If this is 1, then $E = k$, $E' = k'$ and there is nothing to prove. If $[E:k] > 1$, then some p_i in (2), say p_1 , has degree greater than 1, and f' has a corresponding factorization $f' = p'_1 \dots p'_r$ over k' . Take any zero α of p_1 in E and any zero α' of p'_1 in E' ; by Prop. 2.2, $k[\alpha] \cong k'[\alpha']$ and $[E:k] = [E:k[\alpha]] \cdot [k[\alpha]:k] > [E:k[\alpha]]$. Clearly E is a minimal splitting field for f over $k[\alpha]$ and E' is a minimal splitting field for f' over $k'[\alpha']$; hence by induction, $E \cong E'$ under the isomorphism mapping $k[\alpha]$ to $k'[\alpha']$. ■

Examples

1. $x^4 + 1 = 0$. Let θ be a root over \mathbf{Q} ; then so are $-\theta$, θ^{-1} and $-\theta^{-1}$, and these roots are distinct, for $\theta \neq -\theta$, because $\theta \neq 0$ and if $\theta = \pm \theta^{-1}$, then $\theta^2 = \pm 1$, hence $\theta^4 + 1 = 2 \neq 0$. Thus we have a complete factorization

$$x^4 + 1 = (x - \theta)(x + \theta)(x - \theta^{-1})(x + \theta^{-1})$$

over $\mathbf{Q}(\theta)$; the latter is therefore a minimal splitting field. We remark that every automorphism of $\mathbf{Q}(\theta)$ can be described by a permutation of the roots of $x^4 + 1 = 0$; the group of all these permutations is the Galois group of the extension $\mathbf{Q}(\theta)/\mathbf{Q}$, to be studied later on, in 3.6. The same holds over \mathbf{F}_p , the field of p elements, when $p \neq 2$. For $p = 2$, $x^4 + 1 = (x + 1)^4$ splits already over \mathbf{F}_2 .

2. $x^3 - 2 = 0$. Over the complex numbers we have

$$x^3 - 2 = (x - \alpha)(x - \omega\alpha)(x - \omega^2\alpha),$$

where $\alpha = \sqrt[3]{2}$, $\omega = (-1 + \sqrt{-3})/2$. Clearly $\mathbf{Q}(\alpha, \omega\alpha) = \mathbf{Q}(\omega, \alpha)$ is a splitting field, but neither $\mathbf{Q}(\alpha)$ nor $\mathbf{Q}(\omega)$ will do, for $\mathbf{Q}(\alpha)$ is real and $\mathbf{Q}(\omega)$ does not contain α . We note that $[\mathbf{Q}(\omega, \alpha):\mathbf{Q}] = 6$, for ω, α have degrees 2, 3 respectively over \mathbf{Q} ; hence $[\mathbf{Q}(\omega, \alpha):\mathbf{Q}]$ is a multiple of 6, and it equals 6 because it is spanned by $1, \alpha, \alpha^2, \omega, \omega\alpha, \omega\alpha^2$ over \mathbf{Q} .

Splitting fields naturally lead to an important class of extensions.

DEFINITION A field extension E/k is said to be *normal* if it is algebraic and every irreducible polynomial over k which has a zero in E splits completely in E .

It should be noted that normality is a property of extensions, not of fields. If $k \subseteq F \subseteq E$ and E/k is normal, then E/F is also normal, but F/k need not be. For example, in the second of the above examples, $\mathbf{Q}(\omega, \alpha)/\mathbf{Q}$ is normal and so is $\mathbf{Q}(\omega, \alpha)/\mathbf{Q}(\alpha)$, but $\mathbf{Q}(\alpha)/\mathbf{Q}$ is not, because $x^3 - 2$ has a zero in $\mathbf{Q}(\alpha)$ without splitting completely. Our next result provides a description of the finite normal extensions.

PROPOSITION 2.4 *A finite extension E/k is normal if and only if it is a minimal splitting field of a polynomial over k .*

Proof. Suppose that E is a minimal splitting field for a polynomial f over k and denote the zeros of f in E by $\alpha_1, \dots, \alpha_n$. Given any irreducible polynomial p over k which has a zero β in E , we have to show that p splits in E . Let β' be another zero of p in some extension of E ; we shall show that $\beta' \in E$. Since p is irreducible over k , we have $k(\beta) \cong k(\beta')$, by Prop. 2.2. Now E is clearly also a splitting field of f over $k(\beta)$, while $E(\beta')$ is a splitting field of f over $k(\beta')$, and as

minimal splitting fields they are isomorphic and $[k(\beta):k] = [k(\beta'):k]$, $[E:k(\beta)] = [E(\beta'):k(\beta')]$. Hence

$$\begin{aligned}[E:k] &= [E:k(\beta)][k(\beta):k] \\ &= [E(\beta'):k(\beta')][k(\beta'):k] \\ &= [E(\beta'):k]\end{aligned}$$

But E is a subspace of $E(\beta')$; since both have the same degree over k , we have $E = E(\beta')$, and so $\beta' \in E$, as claimed.

Conversely, let E/k be a finite normal extension. We can write $E = k(\alpha_1, \dots, \alpha_n)$, taking e.g. a basis of E over k for the α 's. Let p_i be the minimal polynomial of α_i over k ; by hypothesis each p_i splits into linear factors over E . Hence the same is true of $f = p_1 \dots p_n$ and E is generated by the zeros of f ; hence it is a minimal splitting field of f over k . ■

In exactly the same way one can show that a general extension is normal iff it is a minimal splitting field for a set of polynomials.

Given any finite extension F/k , say $F = k(\alpha_1, \dots, \alpha_n)$, let p_i be the minimal polynomial for α_i over k and put $f = p_1 \dots p_n$. Any normal extension of k containing F must be a splitting field of f over k ; if E is a minimal splitting field, it is normal over k and is contained in any normal extension containing F . Thus E/k may be described as the least normal extension containing F/k ; it is called the *normal closure* of F/k . The construction shows that it is unique up to F -isomorphism. It may also be described as the field obtained by adjoining to F all the zeros of all irreducible polynomials over k which have a zero in F . Let us note another useful property of normal extensions.

COROLLARY 2.5 *Given a tower of finite extensions $k \subseteq F \subseteq E$, if E/k is normal, then any k -homomorphism $\varphi: F \rightarrow E$ extends to a k -automorphism of E .*

Proof. By Prop. 2.4 E is a minimal splitting field of a polynomial f over k , hence also over F . Write $F' = F\varphi$; then E is also a minimal splitting field of f over F' . By Th. 2.3, the isomorphism $\varphi: F \rightarrow F'$ can be extended to an automorphism of E , which must leave k pointwise fixed, because φ does. ■

Two elements or two subextensions of a normal extension E/k are said to be *conjugate* if there is a k -automorphism of E which transforms one into the other. We observe that a subextension is normal precisely when it is equal to all its conjugates (cf. Th. 6.2 below). This definition shows the truth of

COROLLARY 2.6 *In a finite normal extension F/k the conjugates of a given element are permuted transitively by the k -automorphisms of F .* ■

Exercises

- (1) Find the degree of a minimal splitting field of f over \mathbf{Q} in the following cases: (i) $f = x^4 - 1$, (ii) $f = x^4 + 1$, (iii) $f = x^4 + 2$, (iv) $f = x^4 + 4$.
- (2) Find the degree of a minimal splitting field of $x^6 + 1$ over \mathbf{Q} and over \mathbf{F}_2 .
- (3) Show that a minimal splitting field over k for a polynomial of degree n is generated over k by any $n - 1$ of its zeros.
- (4) Show that $x^4 - 2x^2 - 2$ is irreducible over \mathbf{Q} , and find two pairs of zeros which generate non-isomorphic extensions.
- (5) Find the normal closure of $\mathbf{Q}(8^{1/n})$ over \mathbf{Q} .
- (6) Show that if E and F are normal extensions of k within a field U , then EF , the subfield generated by E and F , and $E \cap F$ are normal over k .
- (7) Show that the field generated by a root of $x^4 - 2 = 0$ over \mathbf{Q} is not normal. Deduce that a normal extension of a normal extension need not be normal.
- (8) Show that E is a normal closure of F/k iff $E \otimes_k F$ is a direct product of $[F:k]$ fields isomorphic to E (cf. 4.7 for tensor products).

3.3 The algebraic closure of a field

In 3.2 we defined a splitting field for any polynomial. More generally let \mathcal{F} be any set of polynomials over a field K . An extension E/k will be called a *splitting field* for the set \mathcal{F} if every $f \in \mathcal{F}$ splits completely over E . As before we can define minimal splitting fields and it is easily seen that E is a minimal splitting field for the set \mathcal{F} iff each $f \in \mathcal{F}$ splits over E and E is generated over k by the set of all zeros of all the members of \mathcal{F} . If \mathcal{F} is finite, say $\mathcal{F} = \{f_1, \dots, f_r\}$, then we can replace it by the product $f = f_1 \dots f_r$. Any splitting field for f_1, \dots, f_r over k is just a splitting field for f , and we are in the case considered in 3.2. As we noted after Prop. 2.4, an extension E/k is normal iff it is a minimal splitting field for some set of polynomials over k .

A field k is said to be *algebraically closed* (as in Vol. 1, p. 165) if every polynomial over k splits already in k . This can also be expressed by saying that k is its own splitting field for the set of all polynomials over it.

Every algebraically closed field is infinite. This follows by the argument used in Euclid's theorem to show the existence of an infinity of prime numbers (Vol. 1, p. 29):

THEOREM 3.1 *Every algebraically closed field is infinite.*

Proof. If k is a finite field, consider the polynomial

$$f = 1 + \prod_{a \in k} (x - a).$$

f has positive degree and $f(a) = 1$ for all $a \in k$, so f has no zeros in k ; hence k cannot be algebraically closed. ■

We shall see later (in 3.8) that the product $\prod(x - a)$ is $x^q - x$, where q is the number of elements of the finite field.

Let k be a field. By an *algebraic closure* of k we understand an extension E of k which is algebraic and which is algebraically closed. Our aim in this section will be to show that every field k has an algebraic closure and this is unique up to isomorphism over k . We begin by establishing the isomorphism property for splitting fields.

PROPOSITION 3.2 *Let k be any field and \mathcal{F} a set of polynomials over k . Then any two minimal splitting fields of \mathcal{F} over k are isomorphic.*

For the finite case this was proved in Th. 2.3. In the general case we shall give two proofs.

First proof. Denote by E, E' the two minimal splitting fields. Let \mathcal{F} be indexed by Λ and for any subset I of Λ write E_I for the subfield of E generated by the zeros of all $f_\lambda (\lambda \in I)$; in particular $E_\Lambda = E$. Now consider the set of all pairs (E_I, φ_I) , where $\varphi_I: E_I \rightarrow E'$ is a k -homomorphism; this set is partially ordered by the rule $(E_I, \varphi_I) \leqslant (E_J, \varphi_J)$ iff $E_I \subseteq E_J$, $\varphi_I = \varphi_J|_{E_I}$. Our set is clearly inductive: given any chain $\{(E_I, \varphi_I)\}$, we take as its upper bound the union $\bigcup E_I$ with the union of the φ_I as homomorphism. As upper bound of the empty chain we have $(k, 1)$. By Zorn's lemma there is a maximal element (E_J, φ_J) . If $J \neq \Lambda$, take $\lambda \in \Lambda \setminus J$ and let E_J be the minimal splitting field of f_λ in E over E_J . By Prop. 2.2 this exists and the homomorphism $\varphi_J: E_J \rightarrow E'$ can be extended to a homomorphism of E_J into E' . But this contradicts the maximality of (E_J, φ_J) , hence $J = \Lambda$ and we have a homomorphism $\varphi: E \rightarrow E'$. Now every $f \in \mathcal{F}$ splits over E and the zeros of all these polynomials are mapped by φ again to zeros of all the $f \in \mathcal{F}$ in E' ; but these zeros generate E' over k , hence φ is an isomorphism. ■

The second proof is shorter and uses tensor products; these are defined in 4.7, but for the proof below we note that if A, B are k -algebras, with bases $\{u_i\}, \{v_j\}$ then $A \otimes B$ is a k -algebra with a basis consisting of symbols $u_i \otimes v_j$ which multiply by the rule: $(x \otimes y)(x' \otimes y') = xx' \otimes yy'$.

Second proof. Form the tensor product $E \otimes E'$ over k . This is a non-zero commutative ring. Its quotient by a maximal ideal \mathfrak{p} is a field $F = (E \otimes E')/\mathfrak{p}$ and

we have homomorphisms of E , E' into F . Denote their images by E_1 , E'_1 respectively; each of E_1 , E'_1 as a minimal splitting field is generated by the zeros of all the polynomials in F ; hence $E_1 = E'_1$ and it follows that $E \cong E'$. ■

Now it is not hard to show the existence of a splitting field. Suppose first that the set of polynomials is countable: f_1, f_2, \dots . We put $E_0 = k$ and define E_n recursively as a minimal splitting field of f_n over E_{n-1} . Then $E_{n-1} \subseteq E_n$ and the union of all the E_n formed as in Prop. 2.2.14, is the required field. This method can be adapted to deal with the general case. Thus given a family of polynomials indexed by any set Λ , take any finite subset I of Λ and denote by f_I the product of the f_λ for $\lambda \in I$. By Th. 2.3 we have a minimal splitting field E_I of f_I over k and for $I \subseteq J$ there is an embedding $E_I \rightarrow E_J$. By making an obvious identification we can regard E_I as a subfield of E_J and we thus have a set $\{E_I\}$ of fields, partially ordered by inclusion. Moreover, this set is *directed*, i.e. given E_I, E_J there is a finite set K (viz. $K = I \cup J$) such that $E_I \subseteq E_K, E_J \subseteq E_K$. As a consequence we can define a field structure on the union E of all the E_I such that each E_I is a subfield. Given $x, y \in E$, say $x \in E_I, y \in E_J$, we can find E_K to contain E_I, E_J and in E_K we can form $x + y, xy, x^{-1}$ (if $x \neq 0$); the verification that E forms indeed a field is straightforward and may be left to the reader. We sum up the result as follows:

PROPOSITION 3.3 *For every set F of polynomials over a field k there is a minimal splitting field, unique up to isomorphism.* ■

Remark. We have shown that any two minimal splitting fields of a given set of polynomials are isomorphic, by an isomorphism which reduces to the identity on the ground field. However, this isomorphism is by no means unique; in fact, the study of the different possible isomorphisms constitutes the subject of Galois theory, which will be treated in 3.6 (and, in the infinite case, in Vol. 3).

We can now prove the existence of an algebraic closure.

THEOREM 3.4 *Let k be a field. Then a minimal splitting field Ω for the set of all polynomials over k is an algebraic closure of k .*

Proof. Since Ω is generated by algebraic elements over k , it is algebraic, by Cor. 1.5. Let f be a polynomial over Ω and denote by E the subfield of Ω generated over k by the coefficients of f . Since E is finitely generated over k , it has finite degree. Let E' be a minimal splitting field of f over E ; then $[E':k] = [E':E][E:k]$, and this is again finite. Hence the zeros of f in E' are zeros of some polynomial g over k ; but g splits over Ω , so these zeros lie in Ω and f splits over Ω , as we had shown. ■

Although Th. 3.4 provides algebraically closed fields in profusion, it does not

tell us whether a given field such as **C**, the complex numbers, is algebraically closed. As already mentioned, this is the case, and there are many proofs of this fact. A very simple one, using complex function theory, is based on Liouville's theorem: if a polynomial f has no complex zero, then $f(z)^{-1}$ is finite throughout the complex plane and bounded as $z \rightarrow \infty$, hence (by Liouville's theorem) a constant. Other more topological proofs are based on the notion of the degree of a mapping. If f does not vanish, then the map

$$z \mapsto f(z) \quad (1)$$

of the 2-sphere (Riemann sphere) into itself omits at least one point from the image, and hence can be deformed into a constant map. But if f has degree n (as a polynomial), this map can also be deformed into the map $z \mapsto z^n$. This would mean that the latter map can be deformed into the constant map, which contradicts the fact that the degree is preserved by deformation.

Whatever method is chosen, the completeness of the complex numbers has to be used at some stage. In Vol. 3 we shall present another proof, which is more algebraic and uses a minimum of topological properties of the real or complex numbers.

Exercises

- (1) Show that the algebraic closure of a countable field is again countable.
- (2) Let k be a field and F/k an algebraic extension. Show that if every finite algebraic extension of k admits a k -homomorphism into F , then F is an algebraic closure of k .
- (3) Show that over a finite field there are irreducible polynomials of arbitrarily high degree.

3.4 Separability

Let k be a field of prime characteristic p . The mapping

$$x \mapsto x^p \quad (1)$$

is an endomorphism, for we have

$$(xy)^p = x^p y^p, \quad 1^p = 1, \quad (x + y)^p = x^p + y^p. \quad (2)$$

Here the last equation holds because the binomial coefficient $\binom{p}{r}$ is divisible by p for $1 \leq r \leq p - 1$. The mapping (1) is sometimes called the *Frobenius endomorphism*; as endomorphism of a field it is necessarily injective. If it is also surjective, and hence an automorphism, the field k is said to be *perfect*. Thus k is perfect iff every element is a p th power, where $p = \text{char } k$; in addition, every

field of characteristic 0 is perfect, by definition. As an example of perfect fields of non-zero characteristic we have

THEOREM 4.1 *Every finite field is perfect.*

For if k is finite, then any injective mapping of k into itself must be a bijection (Lemma 3 of 1.3, p. 15, Vol. 1). ■

As an example of an imperfect field consider $k(x)$, the field of rational functions in an indeterminate x over a field of prime characteristic. Clearly the indeterminate x is not a p th power in $k(x)$. If we write k^p for the image of k under the Frobenius mapping, then

$$k(x)^p = k^p(x^p).$$

For example if k is a finite field, then $k^p = k$ and $k(x)^p$ consists of all rational functions in x^p .

In order to study extensions of imperfect fields we need to consider the multiplicities of zeros of polynomials more closely. Let f be a monic polynomial of degree n over a field k . In any minimal splitting field E of f we have

$$f = (x - \alpha_1)^{m_1} \dots (x - \alpha_t)^{m_t},$$

where $\alpha_1, \dots, \alpha_t$ are the distinct zeros of f and m_i their multiplicities. Let E' be another minimal splitting field of f ; this is isomorphic to E , by Th. 2.3, and if $\alpha_i \mapsto \alpha'_i$ under the isomorphism, then we have in E'

$$f = (x - \alpha'_1)^{m_1} \dots (x - \alpha'_{t'})^{m_{t'}}.$$

Thus the multiplicities are independent of the choice of the splitting field.

As we know from 6.8 of Vol. 1, we can test f for multiple zeros by finding its formal derivative f' . We recall that for $f = a_0 + a_1x + \dots + a_nx^n$ this is defined as $f' = a_1 + 2a_2x + \dots + na_nx^{n-1}$. We also recall the familiar rules (which suffice to define it uniquely): $(f + g)' = f' + g'$, $(af)' = af'(a \in k)$, $(fg)' = f'g + fg'$, $x' = 1$. In terms of the derivative we have the following test for multiple zeros, which can be carried out entirely within the ground field:

PROPOSITION 4.2 *Let k be any field and $f \in k[x]$, where $\deg f > 0$. Then the zeros of f (in some splitting field of f) are simple if and only if f is prime to its derivative f' .*

The proof (cf. Vol. 1, p. 169) rested on the observation that for any α in any extension of k ,

$$f = (x - \alpha)^2 g + (x - \alpha)f'(\alpha) + f(\alpha),$$

hence $(x - \alpha)^2 | f$ iff $f(\alpha) = f'(\alpha) = 0$, i.e. f and f' have $(x - \alpha)$ as common factor. ■

A polynomial is said to be *separable* if all its zeros (in some splitting field) are simple. If we apply Prop. 4.2 to an irreducible polynomial we find

COROLLARY 4.3 *An irreducible polynomial f over a field k is separable unless $f' = 0$. In particular, over a field of characteristic 0 every irreducible polynomial is separable.*

For assume that $f' \neq 0$ and put $d = (f, f')$. Then $\deg d \leq \deg f' < \deg f$; hence d , as a proper factor of an irreducible polynomial f , has degree 0, i.e. $d = 1$ and so f is prime to f' ; this ensures that f is separable, by Prop. 4.2. If $\text{char } k = 0$, then $f' = 0$ means that f is of degree 0, contradicting the irreducibility. This proves the second part. ■

This tells us all we need to know in characteristic 0, but the case of prime characteristic is more complicated.

LEMMA 4.4 *Let k be a field of prime characteristic p . For any polynomial f over k , $f' = 0$ if and only if $f(x) = g(x^p)$ for some polynomial g .*

Proof. Let $f = \sum a_i x^i$; if $f' = 0$, then $ia_i = 0$ and it follows that $a_i = 0$ for $p \nmid i$. Hence f must have the form

$$f = a_0 + a_p x^p + a_{2p} x^{2p} + \cdots + a_{sp} x^{sp} = g(x^p). \quad (3)$$

Conversely it is clear that any f of the form (3) satisfies $f' = 0$. ■

To illustrate the lemma we remark that $x^p - a$ has zero derivative, and in fact if $a \notin k^p$, then this polynomial has no zero in k , for over its splitting field it takes the form $(x - \alpha)^p$, where $a = \alpha^p$.

An element α of an algebraic extension F/k is called *separable* over k if its minimal polynomial over k is separable. If every element of F is separable over k we call F/k a *separable extension*. With the help of perfect fields we can describe separable extensions:

PROPOSITION 4.5 *Over a perfect field every algebraic extension is separable. Conversely, if every algebraic (or even every finite) extension of k is separable, then k is perfect.*

Proof. For characteristic 0 there is nothing to prove: in that case every field is perfect and every algebraic extension is separable. So we may take the characteristic p to be prime.

Let k be a perfect field and f any irreducible polynomial over k . If f is not separable, then $f' = 0$, hence $f(x) = g(x^p)$, say

$$f = a_0 + a_1 x^p + \cdots + a_r x^{pr}.$$

By hypothesis k is perfect, so $a_i = b_i^p$ for some $b_i \in k$, and now

$$\begin{aligned} f &= b_0^p + b_1^p x^p + \cdots + b_r^p x^{pr} \\ &= (b_0 + b_1 x + \cdots + b_r x^r)^p; \end{aligned}$$

but this contradicts the irreducibility of f . Hence every irreducible polynomial is separable and it follows that every algebraic extension is separable.

Conversely, if every finite extension is separable, take $a \in k$ and consider a splitting field of $x^p - a$. If b is a zero, then $x^p - a = (x - b)^p$; thus all zeros coincide, and $x^p - a$ is therefore reducible. An irreducible factor can only have one zero and so must be linear, hence $b \in k$, and $a = b^p \in k^p$; this shows k to be perfect. ■

Exercises

- (1) Show that every finite extension of a perfect field is perfect.
- (2) Show that if f is irreducible over k , then all its zeros have the same multiplicity.
- (3) Show that over a field of characteristic p , $x^p - a$ is either a p th power of a linear polynomial or irreducible.
- (4) Show that a field of characteristic p cannot have n distinct n th roots of 1 unless $p \nmid n$.
- (5) Let k be a field of prime characteristic p . Show that if an element α in an extension of k is separable over $k(\alpha^p)$, then $\alpha \in k(\alpha^p)$. Deduce another proof of the converse of Prop. 4.5.
- (6) Let f be a polynomial over a field of characteristic zero. Show that f has an m -fold zero iff $f, f', \dots, f^{(m-1)}$ have a common factor. Verify that this still holds for characteristic p if $p \geq m$; what happens if $p < m$?
- (7) Let F/k be any extension of prime characteristic p and write $k_0 = \{\alpha \in F \mid \alpha^{p^r} \in k \text{ for some } r \geq 0\}$. Show that k_0 is a subfield of F and if F is perfect, then so is k_0 , but no smaller field containing k is perfect. Show that any automorphism of F/k leaves k_0 elementwise fixed.

3.5 Automorphisms of field extensions

Galois theory may be described as the study of field extensions by means of their automorphisms. A basic aim is to show that under suitable conditions an extension of degree n has just n automorphisms. In this section we shall show that there cannot be more than n automorphisms, and in the next section we shall find the conditions for equality. Although Galois developed his theory to deal with roots of equations, the theory was expressed by Dedekind in terms of field extensions, by means of two key lemmas on automorphisms. The first gives an

upper bound on the number of k -automorphisms of a finite extension of k ; we shall state the result more generally for homomorphisms, since it is no harder to prove in this form.

Consider any family of homomorphisms $\sigma_i: E \rightarrow E'$ ($i = 1, \dots, r$) between fields, regarded as k -algebras. It will be convenient to write such homomorphisms as exponents. We obtain a k -linear mapping from E to E' by forming a linear combination of these homomorphisms with coefficients in E' :

$$\sum \sigma_i \alpha_i: x \mapsto \sum x^{\sigma_i} \alpha_i \quad (x \in E, \alpha_i \in E').$$

Of course this need not be a homomorphism. We claim that it is a non-zero mapping except in trivial cases, namely (i) if all the α_i vanish, or (ii) if $\sigma_1 = \sigma_j$ and $\alpha_i = -\alpha_j$.

LEMMA 5.1 (Dedekind's lemma) *Any set of distinct homomorphisms of a field E into another field E' is linearly independent over E' .*

Proof. Let $\{\sigma_i\}$ be a family of distinct homomorphisms. If they are linearly dependent, let us take a minimal linearly dependent subset, $\sigma_1, \dots, \sigma_r$, say, where σ_1 is linearly dependent on $\sigma_2, \dots, \sigma_r$:

$$x^{\sigma_1} = \sum_2^r x^{\sigma_i} \alpha_i \quad \text{for all } x \in E, \text{ where } \alpha_i \in E'. \quad (1)$$

Replace x by xy in (1):

$$x^{\sigma_1} y^{\sigma_1} = \sum_2^r x^{\sigma_i} y^{\sigma_i} \alpha_i. \quad (2)$$

Now multiply (1) by y^{σ_1} and subtract the result from (2):

$$\sum_2^r x^{\sigma_i} (y^{\sigma_i} - y^{\sigma_1}) \alpha_i = 0.$$

This holds for all $x \in E$, but by hypothesis, $\sigma_2, \dots, \sigma_r$ are linearly independent, so $(y^{\sigma_i} - y^{\sigma_1}) \alpha_i = 0$. Since $\sigma_i \neq \sigma_1$, it follows that $\alpha_i = 0$ for all $i \neq 1$, and (1) reduces to the form $x^{\sigma_1} = 0$; putting $x = 1$, we find $1 = 0$, a contradiction. ■

COROLLARY 5.2 *Given two extensions E, E' of k , if $[E:k] = n$, then there are at most n k -homomorphisms from E to E' .*

Proof. Let u_1, \dots, u_n be a k -basis of E , and suppose that there are $n + 1$ distinct k -homomorphisms $\sigma_0, \dots, \sigma_{n+1}$. Then the n equations in the $n + 1$ unknowns x_i :

$$\sum_i u_j^{\sigma_i} x_i = 0 \quad (j = 1, \dots, n)$$

have a non-zero solution $x_i = c_i$ in E' (Cor. 3 to Th. 1 of 5.1, Vol. 1). Any $a \in E$ has

the form $a = \sum \alpha_j u_j$, hence

$$\sum a^{\sigma_i} c_i = \sum (\sum \alpha_j u_j)^{\sigma_i} c_i = \sum \sum \alpha_j (u_j^{\sigma_i}) c_i = 0,$$

and this contradicts the lemma. ■

We note that when $E' = E$, any homomorphism is necessarily an automorphism (as injective endomorphism of a finite-dimensional vector space). We are specially interested in the case where the number of k -automorphisms of E is exactly n —by Cor. 5.2 it cannot be larger. Let us consider some examples to illustrate the situation.

Examples.

1. $[\mathbf{C}:\mathbf{R}] = 2$. Here there are two automorphisms, the identity and complex conjugation: $a \mapsto \bar{a}$.

2. $\mathbf{Q}(\alpha)/\mathbf{Q}$, where α is the real root of $x^3 = 2$. Here $[\mathbf{Q}(\alpha):\mathbf{Q}] = 3$, but there are no automorphisms other than the identity, for $\mathbf{Q}(\alpha)$ contains only one root of $x^3 = 2$, which is necessarily mapped to a root of the same equation by any automorphism, and must therefore stay fixed. The number of automorphisms falls short of the possible total of three because (as we shall see in a moment) the extension is not normal. If we go over to the normal closure $\mathbf{Q}(\alpha, \omega)$, where ω is a root of $x^3 = 1$, but $\omega \neq 1$, we find six automorphisms of $\mathbf{Q}(\alpha, \omega)$, and in fact $\mathbf{Q}(\alpha)$ has three \mathbf{Q} -homomorphisms into $\mathbf{Q}(\alpha, \omega)$.

3. Let k be a field of characteristic $p \neq 0$, $F = k(t)$ the field of rational functions in an indeterminate t and α a root of the equation $x^p = t$. Then $F(\alpha)/F$ has only one automorphism and in fact there is at most one F -homomorphism of $F(\alpha)$ into any field, because α must map to a root of $x^p = t$. This equation has at most one root, and the automorphism is determined once the image of α is fixed. Here the number of automorphisms falls short because the extension is inseparable.

These examples suggest that in order to get the maximum number of automorphisms of F/k we need to take F/k normal and separable, and this will be carried out in the next section. As a matter of fact, from any extension F/k we can form the normal closure, and this will be separable whenever F/k is; but when F/k is inseparable, there is nothing we can do to get ‘enough’ automorphisms. These ideas are made precise in Th. 5.4 below, which is preceded by a lemma on the extension of homomorphisms. Two extensions F/k and F'/k' will be called *isomorphic* if there is an isomorphism $F \cong F'$ and k maps to k' in this isomorphism.

LEMMA 5.3 *Let F/k be a finite extension and E/k a normal extension. If there is a k -homomorphism $\varphi: F \rightarrow E$, then any k -homomorphism of a subextension of F into E can be extended to a k -homomorphism of F into E .*

Proof. Let $k \subseteq D \subseteq F$ and let $\theta: D \rightarrow E$ be a k -homomorphism. If $D \neq F$, take

$\alpha \in F \setminus D$ and let f be the minimal polynomial of α over D ; then f is irreducible over D and the corresponding polynomial $f\theta$ is irreducible over $D' = D\theta$. Let g be the minimal polynomial of α over k ; since $g(\alpha) = 0$, we have $f|g$, hence (applying θ) we find that $f\theta|g$. Now g has a zero in E , viz. $\alpha\varphi$, therefore it splits completely and so $f\theta$ also splits over E . If α' is any zero of $f\theta$ in E , then by Prop. 3.2, there is an isomorphism from $D(\alpha)/D$ to $D'(\alpha')/D'$, in which α maps to α' . So θ has been extended to a homomorphism of $D(\alpha)$ into E , and after a finite number of steps we reach the required homomorphism $F \rightarrow E$. ■

We now come to the promised result, showing that any separable extension has enough homomorphisms into a normal closure:

THEOREM 5.4 *Let F/k be an extension of finite degree n and F'/k' an isomorphic extension; say $\varphi: F \cong F'$ is an isomorphism and*

$$\varphi_0:k \cong k', \quad (3)$$

where $\varphi_0 = \varphi|k$. If E'/k' is a normal extension containing F'/k' , then there are at most n homomorphisms $F \rightarrow E'$ which extend the isomorphism φ_0 in (3), and the normal closure of F'/k' in E'/k' is the field generated by the images of these homomorphisms.

Moreover, the following conditions on the extension F/k are equivalent:

- (a) *there are exactly $n = [F:k]$ homomorphisms from F to E' extending φ_0 ,*
- (b) *F/k is separable,*
- (c) *F/k is generated by separable elements.*

Proof. By Cor. 5.2 there are at most n homomorphisms $F \rightarrow E'$ extending φ_0 . Let E'' be the subfield of E' generated by the images of F . Then $E'' \supseteq F'$ and we have to show that E''/k' is normal. By construction, E'' is generated by the homomorphic images in E' of elements of F . Take any $\alpha \in F$, let f be its minimal polynomial over k , and denote by f' its image under φ_0 . Then f' is irreducible over k' and has a zero in E' , hence it splits completely over E' . If α' is any zero of f' in E' , then $k(\alpha)/k$ is isomorphic to $k'(\alpha')/k'$ and by Lemma 5.3 this isomorphism extends to a homomorphism of F into E' ; hence α' as the image of α lies in E'' . Moreover, E'' is generated by all such elements α' ; thus E'' is generated over k' by all the zeros of the polynomials corresponding to minimal polynomials of elements of F , and so E''/k' is normal, as claimed.

(a) \Rightarrow (b). To prove that F/k is separable we must show that each element of F is separable over k . Let $\alpha \in F$ have minimal polynomial f over k , of degree r . If α is not separable, then the number s of distinct zeros of f in a splitting field is less than r . In particular, there are only s homomorphisms of $k(\alpha)$ into E' , for each such homomorphism is determined by the image of α . Each of these homomorphisms extends in at most $[F:k(\alpha)] = n/r$ ways to a homomorphism of F into E' , by

Cor. 5.2. Hence there are $s.n/r < n$ homomorphisms in all, which contradicts (a).
 (b) \Rightarrow (c) is clear.

To prove (c) \Rightarrow (a), let $F = k(\alpha_1, \dots, \alpha_r)$, where each α_i is separable over k . If α_1 has the minimal polynomial f_1 of degree n_1 over k , then the corresponding polynomial over k' has n_1 zeros and we obtain n_1 homomorphisms $k(\alpha_1) \rightarrow E'$ by mapping α_1 to one of these zeros. Now $F/k(\alpha_1)$ is separably generated and by induction on the degree, any homomorphism of $k(\alpha_1)$ into E' extends in $n/n_1 = [F:k(\alpha_1)]$ ways to a homomorphism of F into E' . Hence the isomorphism φ_0 extends in $n_1 \cdot n/n_1 = n$ ways to a homomorphism of F into E' , as claimed. ■

In the next section we shall study the case of separable normal extensions in more detail; for the moment we shall prove a result which ensures the existence of a sufficient supply of automorphisms under rather different conditions. This is known as Artin's theorem, although it also goes back to Dedekind, cf. Dedekind (1964 p. 50).

THEOREM 5.5 (Artin's theorem) *Let E be a field, G a group of automorphisms of E and k the set of elements of E fixed by G . Then k is a subfield of E and E has finite degree over k if and only if G is finite. In this case*

$$[E:k] = |G|. \quad (4)$$

Proof. That k is a field is easily checked. By Cor. 5.2, $|G| \leq [E:k]$, and it remains to establish equality. When G is infinite, this is clear, so let $|G| = r$ and assume that we have $r+1$ elements of E that are linearly independent over k : u_0, \dots, u_r . Consider the r equations in the $r+1$ unknowns x_j :

$$\sum u_j^\sigma x_j = 0 \quad \text{for all } \sigma \in G. \quad (5)$$

They have a non-trivial solution $x_j = a_j$ in E . We pick a solution with the fewest non-zero terms; if $a_0 \neq 0$, say, we can solve for the term in u_0 :

$$u_0^\sigma = \sum_1^r u_j^\sigma b_j \quad \text{for some } b_j \in E \text{ and all } \sigma \in G. \quad (6)$$

For $\sigma = 1$, (6) reads $u_0 = \sum u_j b_j$, so not all the b_j lie in k , by the linear independence of the u 's over k , say $b_1 \notin k$. By the definition of k there exists $\tau \in G$ such that $b_1^\tau \neq b_1$. Now replace σ in (6) by $\sigma\tau^{-1}$, apply τ and note that σ runs over G as $\sigma\tau^{-1}$ does:

$$u_0^\sigma = \sum_1^r u_j^\sigma b_j^\tau \quad \text{for all } \sigma \in G. \quad (7)$$

Subtracting (7) from (6), we obtain

$$\sum_1^r u_j^\sigma (b_j - b_j^\tau) = 0.$$

This is a shorter relation than (6), and is non-trivial, because $b_1^r \neq b_1$, so we have a contradiction. It follows that $[E:k] \leq r$, and this proves equality in (4). ■

Exercises

- (1) Show that Dedekind's lemma holds for any set of homomorphisms of a group G into the multiplicative group of a field, i.e. verify that the additive structure of E is not involved.
- (2) Use Artin's theorem to show that for any field E with n distinct automorphisms, if k is the fixed field of this set of automorphisms, then $[E:k] \geq n$.
- (3) Find the fixed field F of the rational function field $k(x)$ under the automorphisms $x \mapsto 1-x$, $x \mapsto 1/x$; show that the degree is 6. Verify that $(x^2 - x + 1)^3 / (x^2 - x)^2$ lies in F and use this fact to find an equation for x over F .
- (4) Let F be a perfect field; show that the set of elements fixed under all automorphisms of F is a perfect subfield.
- (5) Let E be a field, G a group of automorphisms and k the fixed field. Show that any $\alpha \in E$ is algebraic over k iff it lies in a finite G -orbit.
- (6) Let E/k be a normal extension. Show that an element of E , of degree r over k , has at most r conjugates over k , with equality iff it is separable.

3.6 The fundamental theorem of Galois theory

Let E be any field and G a group of automorphisms of E . With each subgroup H of G we associate a subset of E :

$$H^* = \{x \in E \mid x^\sigma = x \text{ for all } \sigma \in H\}.$$

Thus H^* is the set of elements fixed by H . It is easily seen to be a subfield of E , called the *fixed field* of the subgroup H . For example, $1^* = E$, and for any subgroup H of G , H^* clearly contains G^* .

Similarly, with each subfield F of E we associate a subset of G , consisting of all the elements leaving F elementwise fixed:

$$F^* = \{\sigma \in G \mid x^\sigma = x \text{ for all } x \in F\}.$$

Again it is easily verified that F^* is a subgroup of G , called the *group of F -automorphisms* in G . Thus, writing $\mathcal{F}(E)$ for the set of all subfields of E and $\mathcal{B}(G)$ for the set of all subgroups of G , we have mappings $\mathcal{B}(G) \rightarrow \mathcal{F}(E)$ and $\mathcal{F}(E) \rightarrow \mathcal{B}(G)$. These mappings satisfying the rules

- $\Gamma.1 \quad F_1 \subseteq F_2 \Rightarrow F_1^* \supseteq F_2^*, \quad (F_i \in \mathcal{F}(E)),$
- $H_1 \subseteq H_2 \Rightarrow H_1^* \supseteq H_2^* \quad (H_i \in \mathcal{B}(G)),$
- $\Gamma.2 \quad F \subseteq F^{**}, \quad H \subseteq H^{**},$
- $\Gamma.3 \quad F^{***} = F^*, \quad H^{***} = H^*.$

$\Gamma.1$ states that the mappings are order-inverting, and follows almost immediately from the definitions, as does $\Gamma.2$. To prove $\Gamma.3$ we need only use $\Gamma.1, 2$: if we replace H by F^* in $\Gamma.2$ we get $F^* \subseteq F^{***}$, and if we operate with $*$ on $\Gamma.2$ and use $\Gamma.1$ we find $F^{***} \subseteq F^*$, hence $F^{***} = F^*$; similarly $H^{***} = H^*$.

Any correspondence between two partially ordered sets satisfying $\Gamma.1, 2$ and hence $\Gamma.3$ is called a *Galois connexion*. We are particularly interested to know how the subfield F and the subgroup H have to be restricted so that the correspondence becomes a bijection. By $\Gamma.3$ we see that we must consider the subfields F satisfying $F^{**} = F$ and the subgroups H satisfying $H^{**} = H$. For the present we shall limit ourselves to the case of finite groups of automorphisms and leave the general case to Vol. 3. We shall see that every subgroup H of G then satisfies $H^{**} = H$, and every field F satisfies $F^{**} = F$, provided that we start as above with a field E and a group G of automorphisms which singles out certain subfields of E . However, it is not true that every finite extension is of this form; a finite extension E/F will be called a *Galois extension* if F is the fixed field of a group of automorphisms of E . The group of all F -automorphisms of E is then called the *Galois group* of the extension and is written $\text{Gal}(E/F)$. With the help of the results in 3.5 it is not hard to describe all Galois extensions:

PROPOSITION 6.1 *Let E/F be a finite field extension. Then (i) E/F is a Galois extension if and only if it is normal and separable; (ii) E/F is contained in a Galois extension if and only if it is separable.*

Proof. Let E/F be a Galois extension with group G and let L be its normal closure. Then E/F has at most $[E:F]$ F -homomorphisms into L , but by Artin's theorem it has that many F -automorphisms; hence by Th. 5.4 E/F is separable and $L = E$, i.e. E/F is normal.

Conversely, suppose that E/F is normal and separable. Then by Th. 5.4, E/F has $[E:F]$ automorphisms. Let $G = F^*$ be the group of all F -automorphisms and put $F_1 = F^{**}$. Then $F_1 \supseteq F$ and by Artin's theorem, $|G| = [E:F_1]$; but by Th. 5.4, $|G| \geq [E:F]$; hence $[E:F_1] \geq [E:F] = [E:F_1][F_1:F]$. It follows that $F_1 = F$ and so E/F is Galois.

To prove (ii) we note that the condition is necessary because every subextension of a separable extension is separable. Conversely, assume that E/F is separable and let L be a normal closure of E over F . Then L/F is normal and separable, as minimal splitting field of a separable polynomial, and it contains E/F , so by (i), E/F is embedded in a Galois extension. ■

We now come to the main theorem of Galois theory, establishing the Galois connexion between fields and automorphism groups.

THEOREM 6.2 *Let E be a field, G a finite group of automorphisms of E and k the fixed field of G . For any field F between k and E define the group of F -*

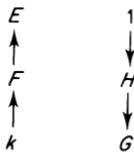
automorphisms in G as

$$F^* = \{\sigma \in G \mid x^\sigma = x \text{ for all } x \in F\},$$

and for any subgroup H of G define the fixed field of H as

$$H^* = \{x \in E \mid x^\sigma = x \text{ for all } \sigma \in H\}.$$

Then the mappings $F \mapsto F^*$, $H \mapsto H^*$ define a Galois connexion between the set of all subgroups of G and the set of all fields between k and E , which is an order-inverting bijection.



Proof. We have to show that $H^{**} = H$ and $F^{**} = F$. Let $H^* = F$; then E/F is Galois, by definition, and $H^{**} \supseteq H$. Moreover, $H^{***} = H^* = F$ and hence, by Artin's theorem, $|H| = [E:F] = |H^{**}|$; therefore $H^{**} = H$.

Secondly, given F , since E/k is normal separable, by Prop. 6.1, so is E/F ; therefore it is Galois, i.e. $F = H^*$ for some $H \subseteq G$. It follows that $F^{**} = H^{***} = H^* = F$. ■

We note the correspondence between group orders and field degrees:

COROLLARY 6.3 *Let E/k be a Galois extension with group G , and let the subfield F correspond to the subgroup H . Then*

$$|H| = [E:F], \quad (G:H) = [F:k].$$

Proof. By Artin's theorem, $[E:k] = |G|$, $[E:F] = |H|$; now the result follows by division. ■

Next we note that conjugate extensions correspond to conjugate subgroups, and normal extensions correspond to normal subgroups.

COROLLARY 6.4 *Let E/k be Galois with group G and let H_1, H_2 be any subgroups, corresponding to subfields F_1, F_2 . Then F_1 and F_2 are conjugate under an automorphism $\sigma \in G$ if and only if H_1 and H_2 are conjugate subgroups of G : $H_2 = \sigma^{-1}H_1\sigma$. In particular, F/k is normal if and only if $H = F^*$ is a normal subgroup of G and when this is so, then $\text{Gal}(F/k) \cong G/H$.*

Proof. By definition, $\tau \in H_1 \Leftrightarrow x^\tau = x$ for all $x \in F_1$. Let $F_2 = F_1^\sigma$; then for any $y \in F_2$, $y^{\sigma^{-1}} \in F_1$, hence $y^{\sigma^{-1}\tau} = y^{\sigma^{-1}}$, i.e. $y^{\sigma^{-1}\tau\sigma} = y$, which means that $\sigma^{-1}\tau\sigma \in H_2$. Thus $\sigma^{-1}H_1\sigma \subseteq H_2$ and since $F_2^{\sigma^{-1}} = F_1$, it follows that $\sigma H_2 \sigma^{-1} \subseteq H_1$, i.e. $\sigma^{-1}H_1\sigma = H_2$. Retracing our steps, we obtain the converse.

If F corresponds to H , then F/k is normal iff it coincides with all its conjugates, i.e. $F^\sigma = F$ for all $\sigma \in G$, and this holds precisely when $\sigma^{-1}H\sigma = H$ for all $\sigma \in G$, i.e. $H \triangleleft G$. In this case we can define a homomorphism from G to $N = \text{Gal}(F/k)$ by restricting each automorphism of E to F . Each $\sigma \in G$ maps F into itself and so defines an element, $\bar{\sigma}$ say, of N ; clearly the mapping $\sigma \mapsto \bar{\sigma}$ is a homomorphism. By Cor. 2.5, each automorphism of F/k extends to an automorphism of E/k , so every element of N is in the image. To determine the kernel, we note that $\bar{\sigma} = 1$ iff σ fixes F , i.e. $\sigma \in H$. Thus $N \cong G/H$, as asserted. ■

This result in particular makes it clear why not every extension is Galois: to a given separable extension F/k there correspond in general several conjugate extensions, and the relation between them is described by a family of conjugate subgroups of $\text{Gal}(E/k)$, where E is the normal closure of F/k . In fact it was this situation which first led to an understanding of the concept of a normal subgroup (which is due to Galois). We also record a condition for the conjugacy of elements; we recall that two elements of an extension are conjugate if they lie in the same G -orbit.

COROLLARY 6.5 *Two elements of a Galois extension E/k are conjugate if and only if they have the same minimal polynomial over k .*

This is an immediate consequence of Prop. 2.2 and Cor. 2.5. ■

We give some examples to illustrate these results.

Examples

1. Quadratic extensions. An extension F/k is said to be *quadratic* if $[F:k] = 2$. In this case F is generated by an element of degree 2 over k (in fact any element of $F \setminus k$ will do), and any polynomial of degree 2 with a zero in F splits in F , so the extension must be normal. If $\alpha \in F \setminus k$ has the minimal polynomial $x^2 + ax + b$, then over F we have $x^2 + ax + b = (x - \alpha)(x - \beta)$, where $\beta = -a - \alpha$. We see that the extension is separable unless $\text{char } k = 2$ and $a = 0$.

2. The symmetric group as Galois group. Let k be any field and write $E = k(x_1, \dots, x_n)$, where the x_i are independent indeterminates. Then E admits the symmetric group G of all permutations of x_1, \dots, x_n as automorphism group. Its fixed field F consists of all symmetric functions in the x 's over k . Let e_1, \dots, e_n be the elementary symmetric functions in the x 's; then clearly $F \cong k(e_1, \dots, e_n)$ and x_1, \dots, x_n are roots of the equation

$$x^n - e_1 x^{n-1} + \cdots + (-1)^n e_n = 0. \quad (1)$$

Hence E is a minimal splitting field of this equation over $k(e_1, \dots, e_n)$; as such its degree is at most $n!$ by Th. 2.3. But we saw that there are $n!$ automorphisms of $E/k(e_1, \dots, e_n)$, namely G . Hence $F = k(e_1, \dots, e_n)$ and $[E:F] = n!$. In particular, it follows that every symmetric function in x_1, \dots, x_n is a rational function in

e_1, \dots, e_n . This is almost the fundamental theorem on symmetric functions (which asserts that every symmetric *polynomial* in the x 's is a *polynomial* in the e 's, cf. 6.9 of Vol. 1).

This example also shows that every finite group occurs as the Galois group of some extension. For any finite group is isomorphic to a subgroup of some symmetric group G , by Cayley's theorem (Vol. 1, p. 62). Take E/F as above with $\text{Gal}(E/F) = G$; then each subgroup H of G is the Galois group of an extension E/K , where K is the fixed field of H . Here we were able to prescribe E , but not K , so the following question remains: Given a field k and a group G , does there exist a Galois extension E/k with group G ? In general the answer is 'no', e.g. for $k = \mathbf{C}$, but for \mathbf{Q} we can always find a Galois extension with the symmetric group, as we shall see in 3.11 below. However, many unsolved questions remain, even when $k = \mathbf{Q}$.

We can now answer a question raised in 3.1, by giving an example of an extension of prescribed degree, without proper subextensions. All we need is to find, for any given n , a finite group with a maximal subgroup of index n . This is provided by S_n , the symmetric group of degree n , and T , the stabilizer of a symbol. Clearly $(S_n : T) = n$; if H is a subgroup such that $T \subset H \subseteq S_n$, then H contains T as well as a permutation moving the symbol fixed by T ; hence it acts transitively and so, by the orbit formula (p. 54, Vol. 1), $n = (H : T)$. It follows that $H = S_n$ and this shows T to be a maximal subgroup. If as in the above example we take $F = k(e_1, \dots, e_n)$ with indeterminates e_i and adjoin a single root of the equation (1), we obtain an extension of degree n without a proper subextension.

3. In general, if we adjoin a root of an equation $f = 0$, we do not get a normal extension, i.e. adjoining one root will usually not be enough to split f into linear factors. If it does, i.e. if we get a splitting field of f by adjoining a single (suitably chosen) zero of f , the equation $f = 0$ is said to be *normal* over k . In example 1 we saw that every quadratic equation is normal, and earlier we noted $x^3 - 2 = 0$ as an example of an equation which is not normal over \mathbf{Q} .

As an example of a normal equation, consider

$$x^3 - 3x - 1 = 0.$$

It may be verified that if θ is any root, then so is $-(\theta + 1)^{-1}$. Taken over \mathbf{Q} , the equation is irreducible, hence $\sigma: \theta \mapsto -(\theta + 1)^{-1}$ defines an automorphism of $\mathbf{Q}(\theta)/\mathbf{Q}$. It is easily checked that $\sigma^2: \theta \mapsto -1 - \theta^{-1}$ and $\sigma^3 = 1$. Thus σ has order 3 and we have found three automorphisms of $\mathbf{Q}(\theta)/\mathbf{Q}$. Since the degree is 3, this is the most we can have and $\mathbf{Q}(\theta)/\mathbf{Q}$ is a Galois extension.

4. So far all our extensions have been finite; to end with, here is an example of what can happen in the infinite case. Let k be any field of characteristic 0, and $F = k(t)$ the field of rational functions in an indeterminate t . On F we have two automorphisms, each of order 2:

$$\begin{aligned} \sigma: t &\mapsto -t, & \text{fixed field: } k(t^2), \\ \tau: t &\mapsto 1-t, & \text{fixed field: } k(t^2 - t). \end{aligned}$$

The least field containing both $k(t^2)$ and $k(t^2 - t)$ is clearly $k(t)$, while $k(t^2) \cap k(t^2 - t) = k$. To see this, let us take f in the intersection, say $f = g/h$, where g, h are coprime polynomials in t . By hypothesis, $f^{\sigma\tau} = f$, hence $g^{\sigma\tau}h = h^{\sigma\tau}g$. On comparing degrees we see that $g^{\sigma\tau} = \lambda g$, where $\lambda \in k$ (because g and h are coprime), and a comparison of leading terms shows that $\lambda = 1$. Thus $g^{\sigma\tau} = g$ and $h^{\sigma\tau} = h$. Now $\sigma\tau: t \mapsto t - 1$; hence if α is a root of $g(x) = 0$, then so is $\alpha - 1$, and by induction, so is $\alpha - n$, for all $n \in \mathbb{N}$. Therefore if g has positive degree (and hence a zero in some extension field), it has infinitely many zeros, a contradiction. Thus $\deg g = 0$; similarly $\deg h = 0$, and it follows that k is the fixed field of the group generated by σ and τ . But in fact k is already the fixed field of the cyclic group generated by $\sigma\tau$, which is clearly smaller, so we no longer have a bijection between subgroups and subfields. Nevertheless there is a Galois connexion for infinite extensions; this will be dealt with in Vol. 3.

Galois' main results in the theory of equations were published in a memoir in 1846 (some 14 years after his death, at the age of 20, as the result of a duel). Even after this long delay, his results appeared very novel and it took many years before they were properly assimilated and further developed. The treatment of Galois theory by Dedekind (in the famous XIth supplement to the 1894 edition of Dirichlet's *Vorlesungen über Zahlentheorie*) emphasized the field rather than the equation and gave the theory its modern 'linear' form. It is also the basis of the account by Artin (1948) which has served as a model for most subsequent expositions of the theory, including this one.

Exercises

- (1) Let k be a field of characteristic not 2. Show that any extension of degree 2 over k is separable and can be generated by a root of an equation $x^2 = a$, where a is an element of k not a square. Conversely, any such equation is separable and has roots $\pm \alpha$; show that the k -automorphisms are 1 and $b + \alpha c \mapsto b - \alpha c$ ($b, c \in k$).
- (2) Let k be of characteristic 2. Show that any separable extension of degree 2 over k can be generated by a root of an equation $x^2 + x + a = 0$, where $a \in k$; conversely, any such equation is separable over k and has roots $\alpha, \alpha + 1$. Show that the automorphisms are 1 and $b + \alpha c \mapsto b + c + \alpha c$ ($b, c \in k$).
- (3) Let k be of characteristic 2. Show that any inseparable element α of degree 2 over k satisfies an equation $x^2 + a = 0$, where a is an element in k not a square, and conversely, any such equation is inseparable over k with a single root.
- (4) Let E/k be a Galois extension, F a field between k and E , and G the subgroup of $\text{Gal}(E/k)$ mapping F into itself. Show that G is the normalizer of $\text{Gal}(E/F)$ in $\text{Gal}(E/k)$ and describe $G/\text{Gal}(E/F)$.

- (5) Let E/k be a Galois extension with group G and N_i a subextension with group $G_i = \text{Gal}(E/N_i)$ ($i = 1, 2$). Show that $G = G_1 \times G_2$ iff N_i/k is Galois and $N_1 \cap N_2 = k$, $N_1 N_2 = E$.
- (6) Let x_1, \dots, x_n be independent indeterminates and let F be the fixed field in $k(x_1, \dots, x_n)$ under the group of all permutations of all the x 's. Show that x_i has degree i over $F(x_{i+1}, \dots, x_n)$, and deduce that the monomials $x_1^{r_1} \cdots x_n^{r_n}$ ($0 \leq r_i \leq i - 1$) form a basis of $k(x_1, \dots, x_n)$ over F .
- (7) Let k be a field of prime characteristic p and $F = k(t)$ the rational function field. Show that the group of automorphisms generated by $\sigma: t \mapsto -t$, $\tau: t \mapsto 1 - t$ is finite. Find the fixed field F_0 and the minimal equation of t over F_0 .
- (8) Let f be a polynomial over a field k and let E be a minimal splitting field for f . Show that the automorphisms of E/k permute the zeros of f transitively iff f is a power of an irreducible polynomial.

3.7 Roots of unity

In factorizing polynomials we shall need Gauss's lemma; this was proved in Vol. 1 (p. 165), but we shall recall it below, in a slightly different form.

Let A be an integral domain. An element p of A is called a *prime* if it is not zero or a unit and if $p|ab$ implies $p|a$ or $p|b$, for any $a, b \in A$. A domain in which every non-zero element is either a unit or a product of primes is called a *unique factorization domain* (UFD); in such a ring any prime factorization of a given element is unique up to the order of the factors and associates (cf. Vol. 1, p. 158 or 9.3 below). Let A be a UFD and consider an element f of the polynomial ring $A[x]$. By taking out common factors of the coefficients we can write $f = af_1$, where $a \in A$ and f_1 is a polynomial whose coefficients have no common factor; such f_1 is called *primitive*. Now Gauss's lemma, as stated in Vol. 1 (p. 165) asserted that over a UFD the product of two primitive polynomials is primitive. We shall prove the form below, from which this easily follows:

LEMMA 7.1 (Gauss's lemma) *Let A be an integral domain. Then any prime of A stays prime in $A[x]$.*

Proof. Let p be a prime in A ; we have to show that $p|fg \Rightarrow p|f$ or $p|g$ for $f, g \in A[x]$. Write $\bar{A} = A/(p)$ and denote the residue class mapping on $A[x]$ by $f \mapsto \bar{f}$. Since p is prime, \bar{A} is again an integral domain, hence so is $\bar{A}[x]$. If $p|fg$, then $\bar{f}\bar{g} = 0$, hence $\bar{f} = 0$ or $\bar{g} = 0$, because $\bar{A}[x]$ is an integral domain. But this means that $p|f$ or $p|g$, as we had to show. ■

Given a ring A and a ring B containing A , we shall say that A is *inert* in B if for $c \in A$, such that $c = ab$ for some $a, b \in B$, there exists a unit u in B such that

$au, u^{-1}b \in A$. This just means that every factorization of c in B can be ‘pulled down’ to A ; in particular, if c is unfactorable in A , it remains so over B .

THEOREM 7.2 *Let A be a UFD with field of fractions K . Then $A[x]$ is inert in $K[x]$.*

Proof. Let $f \in A[x]$ and suppose that $f = gh$ for some $g, h \in K[x]$. By taking a common denominator of the coefficients of g we can find $\alpha_0 \in A$ such that $\alpha_0 g$ has coefficients in A . On dividing by any common factor we find $\alpha \in K$ such that $g_1 = \alpha g$ is a primitive polynomial over A . Similarly there exists $\beta \in K$ such that $h_1 = \beta h$ is primitive. Write $\alpha\beta = a/b$, where a, b are coprime elements of A . Then

$$\frac{a}{b}f = \alpha\beta gh = g_1h_1,$$

hence $af = bg_1h_1$. We claim that a must be a unit. For if not, then there is a prime p dividing a . Hence $p \mid bg_1h_1$ but $p \nmid b$, because a, b are coprime, and $p \nmid g_1, p \nmid h_1$ because g_1, h_1 are primitive. This contradicts Gauss’s lemma, hence a is a unit. Now $\alpha\beta b = a$, and $\alpha g = g_1, \alpha^{-1}h = a^{-1}b\beta h = a^{-1}bh_1$ have coefficients in A , and this is what we had to show. ■

Applied to \mathbf{Z} this result shows that any polynomial over \mathbf{Z} which can be factorized over \mathbf{Q} can also be factorized over \mathbf{Z} . In particular, if a monic polynomial over \mathbf{Z} can be factorized over \mathbf{Q} , we can take all the factors to be monic with coefficients in \mathbf{Z} . Another consequence is the familiar fact that any rational root of an equation with integer coefficients

$$a_0x^n + a_1x^{n-1} + \cdots + a_n = 0 \quad (a_i \in \mathbf{Z}) \quad (1)$$

has a denominator dividing a_0 and a numerator dividing a_n .

By an n th root of 1 (or of unity) we understand any root of the equation

$$x^n = 1. \quad (2)$$

The basic fact is that this equation, like any equation of degree n , cannot have more than n roots in any field (Th. 3, Cor. 1 of **6.6**, Vol. 1). We ask: When does it have precisely n roots? In a splitting field (2) has distinct roots provided that $x^n - 1$ is prime to its derivative, nx^{n-1} , and clearly this is the case iff

$$nx^{n-1} \neq 0.$$

Thus (2) has distinct roots whenever the characteristic is prime to n . In the case when $\text{char } k = p$ and $n = n'p^r$, $(n', p) = 1$, we have $x^n - 1 = (x^{n'} - 1)^{p^r}$, so (2) is then equivalent to the equation

$$x^{n'} = 1.$$

We shall therefore assume in what follows that the characteristic of the field is

prime to n . We also recall the basis theorem for finite abelian groups (9.7 in Vol. 1), stated in multiplicative form for convenience:

Every finite abelian group can be written as a direct product of cyclic groups,

$$A = C_1 \times \cdots \times C_r, \quad (3)$$

and if $|C_i| = \gamma_i$, then $\gamma_i | \gamma_{i+1}$.

For any finite group G the *exponent* is defined as the least positive integer v such that $x^v = 1$ for all $x \in G$. Then it is clear that the abelian group A in (3) has exponent γ_r , and the representation (3) shows that A has an element of order γ_r , viz. any generator of C_r .

THEOREM 7.3 *In any field k , the roots of the equation $x^n = 1$ form a cyclic group under multiplication, whose order is a divisor of n .*

Proof. Clearly the roots form a group, G say, namely a subgroup of the multiplicative group k^\times of k . If G has exponent v , then $v|n$ and G has at most v elements, because the equation $x^v = 1$ has at most v roots in k . By what has been said, G contains an element of order v and therefore G is cyclic. ■

By a *primitive* n th root of 1 one understands a field element whose multiplicative order is precisely n . In the above proof the elements which form generators of G are precisely the primitive n th roots of 1.

Alternative proof. We now give a second proof of Th. 7.3, more elementary and more explicit. It will be enough to show that any finite subgroup G of k^\times is cyclic. Let $|G| = n$; then the elements of G all satisfy (2). Thus k contains n n th roots of 1 and we have to find a primitive n th root. For any $d|n$ denote by $\psi(d)$ the number of primitive d th roots of 1 in k . If there is a primitive d th root, say ζ , then the roots of $x^d = 1$ are the different powers of ζ , and just $\phi(d)$ of these are primitive d th roots. Here $\phi(d)$, Euler's function, is the number of integers in the range $1, 2, \dots, d$ that are prime to d . Thus we see that $\psi(d)$ is either 0 or $\phi(d)$. Now we observe the following:

(i) For each $d|n$ there are just $\phi(n/d)$ integers x in the range $1 \leq x \leq n$ such that $(x, n) = d$, and as d ranges over all divisors of n , every number from 1 to n is counted exactly once. Hence

$$n = \sum \phi(n/d) = \sum \phi(d),$$

where the summation is over all the divisors d of n .

(ii) Every n th root of 1 is a primitive d th root for some $d|n$. Since there are altogether n n th roots of 1 in k , by hypothesis, we find

$$n = \sum \psi(d),$$

where the summation is again over all the divisors of n . Thus

$$\sum_{d|n} [\varphi(d) - \psi(d)] = 0,$$

and all the summands are non-negative; hence they must all be zero, in particular, $\psi(n) = \varphi(n) > 0$. This shows that any group of order n in k^\times contains a primitive n th root of 1; in fact the number is $\varphi(n)$ and Th. 7.3 follows. ■

If ξ_1, \dots, ξ_r ($r = \varphi(n)$) are all the primitive n th roots of 1, then the polynomial

$$\Phi_n(x) = \prod_{i=1}^r (x - \xi_i)$$

is called the n th cyclotomic polynomial; its degree is $\varphi(n)$. The word ‘cyclotomic’ (= circle-cutting) is used because in the complex field \mathbf{C} the n th roots are $e^{2\pi i k/n}$ ($k = 1, 2, \dots, n$) and are obtained geometrically by dividing the unit circle about the origin into n equal parts. To obtain an explicit expression for Φ_n we observe that every n th root of 1 is a primitive d th root of 1, for a unique divisor d of n . Hence

$$x^n - 1 = \prod_{d|n} \Phi_d(x). \quad (4)$$

By Th. 7.2, $\Phi_d(x)$ is again monic, with integer coefficients. Using the Möbius function introduced in 2.4, we can solve for Φ_n and obtain the formula

$$\Phi_n(x) = \prod_{d|n} (x^d - 1)^{\mu(n/d)}.$$

For example, $\Phi_1 = x - 1$, $\Phi_2 = x + 1$, $\Phi_3 = x^2 + x + 1$, $\Phi_4 = x^2 + 1$, $\Phi_5 = x^4 + x^3 + x^2 + x + 1$, $\Phi_6 = x^2 - x + 1$, $\Phi_{12} = x^4 - x^2 + 1$. It is a curious fact that the coefficients of Φ_n are 0 or ± 1 for $n < 105$, but when $n = 105$, there is a coefficient -2 , and as Schur has shown, the absolute values of the coefficients of Φ_n are unbounded as n increases.

Let k be a field of characteristic prime to n , and ζ a primitive n th root of 1 over k . Then $k(\zeta)$ is a minimal splitting field of $x^n = 1$ over k ; since $nx^{n-1} \neq 0$, the roots are distinct and so $k(\zeta)/k$ is a Galois extension; it is also called a cyclotomic extension. Let f be the minimal polynomial of ζ over k , of degree s say; then $[k(\zeta):k] = s$. Clearly $f|\Phi_n$, hence $s \leq \varphi(n)$, with equality iff Φ_n is irreducible over k . Whether this is so naturally depends on k ; we now show that it is the case for $k = \mathbf{Q}$.

THEOREM 7.4 *The cyclotomic polynomial Φ_n is irreducible over \mathbf{Q} .*

Proof. Let ζ be a primitive n th root of 1 over \mathbf{Q} , and let f be its minimal polynomial over \mathbf{Q} . Then $f|\Phi_n$, say $\Phi_n = ff_1$ and by Th. 7.2, f may be taken to be monic and over \mathbf{Z} .

If p is any prime not dividing n , then ζ^p is also primitive, with minimal

polynomial g , say. Thus $f(x)$ and $g(x^p)$ have a common zero ζ , and so

$$g(x^p) = f(x)h(x). \quad (6)$$

We claim that f and g have a common zero. For if not, then the zeros of f and g are distinct zeros of $x^n - 1$, and so

$$x^n - 1 = f(x)g(x)k(x). \quad (7)$$

Now both (6) and (7) have integer coefficients, so we can reduce them mod p . As a polynomial over \mathbf{F}_p , the right-hand side of (7) has distinct zeros, because its derivative is $nx^{n-1} \not\equiv 0 \pmod{p}$. But (6), taken mod p , becomes $f(x)h(x) \equiv g(x^p) \equiv g(x)^p$, so that f and g have a common zero mod p , a contradiction.

Thus we have shown that f and g have a common zero, and by irreducibility, taking them both to be monic, we find that $f = g$. This means that $f(\zeta) = 0$ implies $f(\zeta^p) = 0$. This holds for any prime p not dividing n , and by induction, for any m prime to n . Since $f(\zeta) = 0$, it follows that $f(\zeta^m) = 0$ for all m prime to n . Hence f has $\varphi(n)$ distinct zeros, and so $f = \Phi_n$, i.e. Φ_n is irreducible, as claimed. ■

We note that over \mathbf{F}_p , Φ_n may become reducible. For example, if $n = 12$ and $p = 11$, then $\Phi_{12} = x^4 - x^2 + 1 = (x^2 - 5x + 1)(x^2 + 5x + 1)$, while for $p = 13$, $x^4 - x^2 + 1 = (x - 2)(x + 2)(x - 6)(x + 6)$. This is not surprising, for as we shall see below, Φ_{12} becomes reducible over $\mathbf{Q}(\sqrt{3})$ and $3 \equiv 5^2 \pmod{11}$, $3 \equiv 4^2 \pmod{13}$.

Next we examine the Galois group of a cyclotomic extension. For any integer n , let us write $\mathbf{U}(n)$ for the multiplicative group of the prime residue classes mod n , i.e. the group of units of \mathbf{Z}/n . By the definition of $\varphi(n)$ we have $|\mathbf{U}(n)| = \varphi(n)$. Now if ζ is a primitive n th root of 1 over \mathbf{Q} , then the other primitive n th roots are of the form ζ^m , where m runs over the prime residue classes mod n . Thus the automorphisms of $\mathbf{Q}(\zeta)/\mathbf{Q}$ are given by

$$\alpha_m: \zeta \mapsto \zeta^m \quad (m \in \mathbf{U}(n)).$$

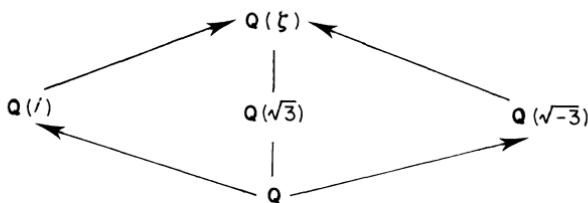
Clearly $\alpha_r \alpha_s = \alpha_{rs}$, so the mapping $m \mapsto \alpha_m$ is a homomorphism of $\mathbf{U}(n)$ onto $\text{Gal}(\mathbf{Q}(\zeta)/\mathbf{Q})$, and since both groups have the same order, they are isomorphic. Thus we have

THEOREM 7.5 *The Galois group of the cyclotomic extension defined by Φ_n over \mathbf{Q} is isomorphic to $\mathbf{U}(n)$, the group of units of \mathbf{Z}/n .* ■

This shows in particular that the Galois group of a cyclotomic extension is abelian. Kronecker has shown that, conversely, every abelian extension of \mathbf{Q} is contained in a cyclotomic extension (Weber 1963, Vol. II, p. 762).

As an example of a cyclotomic extension, let us take $n = 12$, $\varphi(12) = 4$. The prime residue classes are 1, 5, 7, 11, and the group $\text{Gal}(\mathbf{Q}(\zeta)/\mathbf{Q})$ is generated by the set $\{\alpha, \beta\}$, where $\alpha: \zeta \mapsto \zeta^5$, $\beta: \zeta \mapsto \zeta^7$. We note that $\alpha\beta: \zeta \mapsto \zeta^{35} = \zeta^{11}$, thus G is the

Klein 4-group. There are three subgroups of order 2 corresponding to three quadratic extensions. We find them as follows: $\zeta^3 = i$ is left fixed by α and $\zeta^4 = (-1 + \sqrt{-3})/2$ is left fixed by β , hence we have the between-fields $\mathbf{Q}(i)$, $\mathbf{Q}(\sqrt{-3})$, and $\mathbf{Q}(\sqrt{3})$. Hence $\mathbf{Q}(i)$ is generated by i , a primitive fourth root of 1, hence a root of $x^{12} = 1$; similarly for $\mathbf{Q}(\sqrt{-3})$. However, the field $\mathbf{Q}(\sqrt{3})$ is not generated by a twelfth root of 1, for the only such roots contained in it are ± 1 .



Let us consider Φ_n over \mathbf{F}_p where p is a prime not dividing n . Then $x^n - 1$ is separable over \mathbf{F}_p and the primitive n th roots of 1 in \mathbf{F}_p are roots of $\Phi_n(x) = 0$, where this equation may now be reducible. We first determine when Φ_n has a linear factor over \mathbf{F}_p , or, what is the same, when \mathbf{F}_p contains a primitive n th root of 1.

PROPOSITION 7.6 *The finite field \mathbf{F}_p contains a primitive n th root of 1 if and only if $p \equiv 1 \pmod{n}$.*

Proof. Any element c of \mathbf{F}_p satisfies $c^{p-1} = 1$, so if c has order n , then $n|p-1$. Conversely, when $n|p-1$, then since \mathbf{F}_p^\times is cyclic of order $p-1$, by Th. 7.3, it follows that \mathbf{F}_p contains an element of order n . ■

From this remark it is easy to deduce a special case of Dirichlet's theorem on primes in arithmetic progressions. This states that for any coprime integers m, n there are infinitely many primes of the form $nr + m$ ($r \in \mathbb{N}$). Most proofs require the theory of the ζ -function or other analytical methods, but for $m=1$ there is a simple direct proof:

THEOREM 7.7 *For any positive integer n there are infinitely many primes congruent to 1 $(\bmod n)$.*

Proof. Assume that the number of such primes is finite, say p_1, \dots, p_r . Put $a = np_1 \cdots p_r$ and let $t \in \mathbb{Z}$; then $\Phi_n(at) \equiv \Phi_n(0) \equiv \pm 1 \pmod{a}$, hence $a = np_1 \cdots p_r |\Phi_n(at) \mp 1$. As $t \rightarrow \infty$, $\Phi_n(at) \rightarrow \infty$, so we have $\Phi_n(at) \neq \pm 1$ for large enough t . Therefore $\Phi_n(at)$ is then divisible by a prime p , and since $n|a$ and

$a|\Phi_n(at) \mp 1$, n is prime to p . This means that at is a primitive n th root of 1 in \mathbf{F}_p , i.e. it has multiplicative order $n \pmod p$, hence $p \equiv 1 \pmod n$, by Prop. 7.6, and $p \neq p_i$ because $\Phi_n(at) \equiv 0 \pmod p$, $\Phi_n(at) \equiv \pm 1 \pmod {p_i}$, $i = 1, \dots, r$. Thus p is another prime $\equiv 1 \pmod n$. ■

Exercises

- (1) Let G be a finite group in which there are at most n elements of order dividing n , for any n . Show that G is cyclic.
- (2) Find Φ_n for $n = 18, 24, 50$.
- (3) Let p be a prime, v a positive integer and $q = p^v$, $r = p^{v-1}$. Show that $\Phi_q(x) = 1 + x^r + x^{2r} + \dots + x^{(p-1)r}$.
- (4) Show that Φ_m is irreducible over a minimal splitting field of Φ_n over \mathbf{Q} iff the highest common factor of m and n is 1 or 2.
- (5) Prove that $\Phi_n(1) = \prod d^{\mu(n/d)}$, where d ranges over all divisors of n . Deduce that $\Phi_n(1)$ is 0, p or 1 according as n is 1, p^v (p a prime) or divisible by at least two primes.
- (6) Solve $x^2 = 2$ over \mathbf{F}_5 .
- (7) Let p, q be distinct primes. Show that $x^q - 1$ splits into linear factors over \mathbf{F}_p iff $p \equiv 1 \pmod q$.
- (8) Show that $\sum \mu(d) = \delta_{n1}$, $\sum |\mu(d)| = 2^k$, where each time the sum is over all divisors of n , and k is the number of distinct prime factors of n .
- (9) Show that for given M the number of integers n such that $\varphi(n) < M$ is finite. Deduce that any finitely generated field extension of \mathbf{Q} contains only finitely many roots of 1.

3.8 Finite fields

A field with a finite number of elements is called a *finite field*, or also a *Galois field*, after its discoverer. The alternative name is used in some books to avoid confusion with finite extensions.

Let V be an n -dimensional vector space over \mathbf{F}_p , the field of p elements. If u_1, \dots, u_n is a basis of V , then each element of V is unique of the form $\sum \alpha_i u_i$, where $\alpha_i \in \mathbf{F}_p$. Since each coefficient can assume p values independently, we obtain p^n elements in all:

LEMMA 8.1 *An n -dimensional vector space over \mathbf{F}_p has p^n elements.* ■

Clearly this does not depend on p being prime, so a corresponding result holds for any finite field.

Any finite field F clearly has prime characteristic p and its prime subfield is \mathbf{F}_p . Hence F is a finite-dimensional vector space over \mathbf{F}_p and it follows that the number of elements in F is a prime power p^n , where $p = \text{char } F$.

If F is a field of q elements, then the multiplicative group F^\times of F has $q - 1$ elements, hence every non-zero element of F satisfies the equation

$$x^{q-1} = 1. \quad (1)$$

By Th. 7.3 it follows that the group F^\times is cyclic. Any generator of F^\times is called a *primitive element* of F . If ζ is a primitive element, then every element of F^\times can be written as a power ζ^a , and these powers can then be used as logarithms; we shall meet an example later.

Since (1) holds identically in F^\times , it follows that every element of F , zero or not, satisfies

$$x^q = x. \quad (2)$$

This equation has at most q roots in any field; therefore its roots are precisely the elements of F . Since $|F| = q$, it follows that (2) has distinct roots in F . This can also be checked directly by taking derivatives and using Prop. 4.2: $(x^q - x)' = qx^{q-1} - 1 = -1 \neq 0$. We can therefore write

$$x^q - x = \prod_{a \in F} (x - a), \quad (3)$$

and this shows that F is a minimal splitting field of $x^q - x$ over its prime subfield \mathbf{F}_p . This description of F shows that it is determined up to isomorphism by q . Thus for any integer q there can be at most one field of q elements, and only when q is a prime power. Conversely, when $q = p^n$, where p is a prime, then there is a field of q elements, namely the minimal splitting field of (2) over \mathbf{F}_p . To show that this splitting field has exactly q elements we observe that the roots of (2) already form a field: if $a^q = a$ and $b^q = b$, then $(a - b)^q = a^q - b^q = a - b$, $(ab)^q = a^q b^q = ab$ and if $b \neq 0$, then $(b^{-1})^q = (b^q)^{-1} = b^{-1}$. We sum up the result as follows.

THEOREM 8.2 (E. H. Moore 1893) *For each prime p and each $n \geq 1$ there is exactly one field of $q = p^n$ elements (up to isomorphism), namely the minimal splitting field of $x^q - x$ over \mathbf{F}_p , and these are the only finite fields.* ■

The field of q elements is denoted by \mathbf{F}_q or sometimes by $\text{GF}(q)$ (for ‘Galois field’). For $q = p$ this agrees with the notation \mathbf{F}_p introduced earlier.

As an example consider \mathbf{F}_9 ; this is the minimal splitting field of $x^9 - x$ over \mathbf{F}_3 . Its degree over \mathbf{F}_3 is 2, so we need to find an irreducible factor of $x^9 - x$ of degree 2. Since \mathbf{F}_9 consists of all the zeros of this polynomial, its irreducible factors must all be of degree 1 or 2 over \mathbf{F}_3 . We have

$$x^9 - x = x(x - 1)(x + 1)(x^2 + 1)(x^2 + x - 1)(x^2 - x - 1).$$

Let α be a root of $x^2 + 1 = 0$. Then the elements of \mathbf{F}_9 can be written as $a + \alpha b$,

where $a, b = 0, 1, 2$; in fact since $\alpha^2 = -1$, we can regard the elements of \mathbf{F}_9 as ‘complex numbers’ over \mathbf{F}_3 .

In this example α is not primitive, because $\alpha^4 = 1$, but $\alpha + 1$ is primitive. Writing $\zeta = \alpha + 1$, we have the following index table, where the second row lists the power of ζ representing the entry in the first row, i.e. the logarithm to base ζ :

x	1	2	α	2α	$\alpha + 1$	$\alpha + 2$	$2\alpha + 1$	$2\alpha + 2$
$\log x$	0	4	6	2	1	7	3	5

For example, $(\alpha + 1)(2\alpha + 1) = \zeta \cdot \zeta^3 = \zeta^4 = 2$. Operating with powers of ζ simplifies multiplication, but addition is now more work. For example, the equation $(\alpha + 2) + (\alpha + 1) = 2\alpha$ becomes $\zeta^7 + \zeta = \zeta^2$. To facilitate addition one defines a function $Z(n)$ (the ‘Zech logarithm’) on the range $\{0, 1, \dots, q-1\}$ by $\zeta^{Z(n)} = \zeta^n + 1$ and uses the formula

$$\zeta^a + \zeta^b = \zeta^{Z(a-b)+b}.$$

For example, since $Z(6) = 1$, we have $\zeta^7 + \zeta = \zeta^{1+Z(6)} = \zeta^2$.

Let us determine the automorphisms of \mathbf{F}_q , where $q = p^n$. As minimal splitting field of $x^q = x$, \mathbf{F}_q is normal and separable over \mathbf{F}_p , hence $\mathbf{F}_q/\mathbf{F}_p$ is a Galois extension, and from 3.6 we know that there are $[\mathbf{F}_q:\mathbf{F}_p] = n$ automorphisms. In this special case we can describe them explicitly. Firstly we have the Frobenius mapping

$$\alpha: a \mapsto a^p, \tag{4}$$

which is an automorphism, because \mathbf{F}_q is finite (Th. 4.1). The fixed field of α is the set of solutions of $x^p = x$, i.e. \mathbf{F}_p . Iterating α , we obtain

$$\alpha^r: a \mapsto a^{p^r},$$

and this is the identity on \mathbf{F}_q iff $n|r$. Thus α has order n , and its powers are the n automorphisms of $\mathbf{F}_q/\mathbf{F}_p$; there can be no more since $|\text{Gal}(\mathbf{F}_q/\mathbf{F}_p)| = [\mathbf{F}_q:\mathbf{F}_p] = n$.

THEOREM 8.3 *Let $q = p^n$, where p is a prime and $n \geq 1$. Then \mathbf{F}_q is a Galois extension of \mathbf{F}_p with cyclic Galois group of order n , generated by the Frobenius mapping (4).* ■

For any $d \geq 1$, every finite field \mathbf{F}_q has an extension of degree d , namely the minimal splitting field of $x^{q^d} - x$, and as in the proof of Th. 8.2 we see that this extension is unique up to isomorphism. Since each extension is generated by a primitive element, its minimal polynomial has degree d over \mathbf{F}_q ; thus \mathbf{F}_q has an irreducible polynomial in each degree d . All these extensions are separable, by Prop. 4.5. Moreover, every finite extension is normal, for its normal closure has a cyclic Galois group, whose subgroups are therefore all normal. Thus every finite extension of a finite field is Galois.

To find the subextensions of \mathbf{F}_q , where $q = p^n$, we need only look for the

subgroups of \mathbf{C}_n , the cyclic group of order n . As we know (3.4 of Vol. 1), they correspond to the factors of n : if $n = dm$, then there is a unique subgroup of order d , generated by α^m , with corresponding subfield \mathbf{F}_{p^m} . Thus we obtain

THEOREM 8.4 *For any prime power $q = p^m$ and any integer d the field \mathbf{F}_q has an irreducible polynomial of degree d , and so has an extension of degree d , namely \mathbf{F}_{q^d} , and any two extensions of degree d are \mathbf{F}_q -isomorphic. Moreover, \mathbf{F}_{p^m} is embeddable in \mathbf{F}_{p^n} if and only if $m|n$; thus the subfields of \mathbf{F}_{p^n} correspond to the factors of n . When $m|n$, \mathbf{F}_{p^m} is the unique subfield of order p^m in \mathbf{F}_{p^n} and $\mathbf{F}_{p^n}/\mathbf{F}_{p^m}$ is a Galois extension with cyclic Galois group of order n/m , generated by α^m , where α is the Frobenius mapping (4).* ■

One of the major applications of the theory of finite fields is coding theory, which forms the subject of Ch. 10.

Finite fields share with the 2-element Boolean algebra the property of being functionally complete (Th. 2.3.3), thus every function can be represented by a polynomial:

PROPOSITION 8.5 *Every finite field is functionally complete. More precisely, any function on \mathbf{F}_q can be represented by a unique polynomial of degree at most $q - 1$ in each variable.*

Proof. We have to show that every mapping $f: \mathbf{F}_q^n \rightarrow \mathbf{F}_q$ is a polynomial. For $n = 1$ this follows by the Lagrange interpolation formula (Vol. 1, p. 185). In fact we have the point function $1 - (x - a)^{q-1}$, which is 1 at $a \in \mathbf{F}_q$ and 0 elsewhere, and which leads to the explicit formula

$$f(x) = \sum_{a \in \mathbf{F}_q} f(a)[1 - (x - a)^{q-1}]. \quad (5)$$

For general n we have similarly,

$$f(x_1, \dots, x_n) = \sum f(a_1, \dots, a_n) \prod_{i=1}^n [1 - (x_i - a_i)^{q-1}], \quad (6)$$

where the sum is taken over all $(a_1, \dots, a_n) \in \mathbf{F}_q^n$. For the product on the right of (6) is 1 when $x_i = a_i$ ($i = 1, \dots, n$) and 0 otherwise. Further, (6) is of degree $< q$ in each x_i ; if we had two such polynomials for f , their difference would be a polynomial of degree $< q$ vanishing on all of \mathbf{F}_q , i.e. on q values, which is only possible when this difference is zero. ■

We end this section with the remarkable result obtained by Wedderburn in 1905 that all finite division rings are commutative. The proof given here is due to Witt (1931). We recall the result of an exercise from Ch. 2 of Vol. 1: If q, m, n are positive integers such that $q > 1$ and $q^m - 1 | q^n - 1$, then $m|n$. For we can write

$n = ma + b$, where $0 \leq b \leq m$, by the division algorithm. Then $(\bmod q^m - 1)$ we have $q^n \equiv q^{ma}q^b \equiv q^b$, and this can only be 1 if $b = 0$.

THEOREM 8.6 (Wedderburn) *Every finite division ring is commutative.*

Proof. Let E be a finite division ring. Its centre k must be a finite field, with q elements, say, and taking a basis for E over k , we see that $|E| = q^n$, where $n = [E:k]$. Let us consider the multiplicative group E^\times of E ; it has order $q^n - 1$. If $\alpha \in E \setminus k$, then the centralizer of α is a proper division subalgebra C of E , of order q^d say. The group C^\times has order $q^d - 1$ and since it is a subgroup of E^\times , we must have $q^d - 1 | q^n - 1$. By the remark made earlier this is possible only if $d | n$. Now the conjugates (in the group sense) of α in E^\times correspond to the cosets of C^\times in E^\times ; hence the number of conjugates is $(q^n - 1)/(q^d - 1)$. This accounts for all elements of E^\times outside the centre, while the centre has order $q - 1$. Since the conjugacy classes form a partition of E^\times , we obtain the equation

$$q^n - 1 = q - 1 + \sum \frac{q^n - 1}{q^d - 1}, \quad (7)$$

where the sum on the right is taken over various proper divisors d of n (possibly repeated). Now consider the cyclotomic polynomial $\Phi_n(x)$. It is a factor of $(x^n - 1)/(x^d - 1)$ for every proper factor d of n , hence the integer $r = \Phi_n(q)$ divides each term in the sum on the right of (7), as well as the left-hand side. Therefore

$$r | q - 1. \quad (8)$$

But $r = \Phi_n(q) = \prod (q - \zeta)$, where ζ runs over the primitive n th roots of 1. Taking the roots to be complex numbers, we see that for $n > 1$, $|q - \zeta| > q - 1$, and hence $r > q - 1$. This contradicts (8), hence $n = 1$, and so $E = k$ is commutative. ■

Exercises

Unless otherwise specified, p is a prime and q a prime power.

- (1) How many elements of \mathbf{F}_q are generators? How many are primitive?
- (2) Show that for $q = p^n$, $x^q - x$ considered over \mathbf{F}_p has irreducible factors of degree n but of no higher degree.
- (3) Show that an irreducible polynomial of degree r over \mathbf{F}_q is a factor of $x^{q^n} - x$ iff $r | n$. Deduce that $x^{q^n} - x = \prod f_i(x)$, where f_i runs over all monic irreducible polynomials whose degrees divide n . Show that if t_r is the number of such polynomials, then $\sum r t_r = q^n$, and deduce a formula for t_r in terms of q, r and the Möbius function.
- (4) Show that $x^p - x - a$ ($a \neq 0$) is irreducible over \mathbf{F}_p . Over \mathbf{F}_q , where $q = p^n$, show that $x^p - x - a$ is irreducible iff it has no linear factor. (Hint. If α is a zero, then so is $\alpha + 1$.)

(5) Let \mathfrak{p} be a non-zero prime ideal in the ring $\mathbf{Z}[i]$ of Gaussian integers (Vol. 1, p. 184 or 9.5 below). Show that $\mathbf{Z}[i]/\mathfrak{p}$ is a finite field; further show that the only fields that can occur are of p or p^2 elements, where p is prime. Describe the quotient rings by the ideals (7) , $(2+i)$, (5) ; which are fields?

(6) For any $r > 0$, denote by S_r the sum of the r th powers of the elements of \mathbf{F}_q . If $q - 1|r$, find $y \in \mathbf{F}_q^\times$ such that $y^r \neq 1$, and deduce that $S_r = 0$. Prove the formula

$$S_r = \begin{cases} -1 & \text{if } q - 1|r, \\ 0 & \text{otherwise.} \end{cases}$$

(7) Put $k = \mathbf{F}_q$, where $q = p^s$, take $f_1, \dots, f_r \in k[x_1, \dots, x_n]$ such that $\sum \deg f_i < n$ and let V be the subset of k^n consisting of the common zeros of all the f 's. Verify that $P = \prod(1 - f_i^{q-1})$ is the characteristic function for V , i.e. for any $\xi = (\xi_1, \dots, \xi_n) \in k^n$,

$$P(\xi) = \begin{cases} 1 & \text{if } \xi \in V, \\ 0 & \text{if } \xi \notin V. \end{cases}$$

For any $f \in k[x_1, \dots, x_n]$ write $S(f)$ for the sum $\sum f(\xi)$ taken over all $\xi \in k^n$. Deduce that $|V| \equiv S(P) \pmod{p}$, and by expressing P as a linear combination of monomials $x_1^{r_1} \dots x_n^{r_n}$ with $\sum r_i < n(q-1)$ and using Ex. (6), show that $S(P) = 0$. Deduce that the number of points on V is divisible by p . If the f_i are homogeneous, they have at least one common solution 0, hence V is then non-empty (Chevalley–Warning).

(8) Show that every quadratic form in at least three variables over \mathbf{F}_q vanishes at a point other than 0 (in particular, ‘every conic over a finite field has a rational point’).

(9) Let ω be a primitive $(p-1)$ th root of 1 in \mathbf{F}_p . Show that the affine group $G: x \mapsto ax + b$, $a, b \in \mathbf{F}_p$, $a \neq 0$, is generated by the mappings $xt = x + 1$ and $xp = \omega x$, and the subgroup generated by τ is normal in G . Prove that every automorphism of G is inner.

3.9 Primitive elements; norm and trace

We have seen that a finite field extension F/k , regarded as a k -algebra, may be completely described in terms of k and the basis products $u_i u_j$. It is therefore in our interest to choose the basis in as simple a form as possible. In particular we shall examine the case when F is *simple*, i.e. generated by a single element over k . Such a generating element for F is called a *primitive element* over k . We have seen in 3.1 that in this case F has a k -basis of the form $1, \alpha, \dots, \alpha^{n-1}$. We shall find that every finite separable extension is simple, but some preparation is necessary.

Firstly we recall the density property from Vol. 1, p. 163: For any non-zero polynomial f in x_1, \dots, x_r over an infinite field k there exist $c_1, \dots, c_r \in k$ such that $f(c_1, \dots, c_r) \neq 0$. For $r = 1$ this follows because a polynomial of degree n cannot vanish at more than n points; for $r > 1$ the result follows by looking at the coefficients of the powers x_1^i and using induction on r . Secondly we need an almost obvious remark on the Galois correspondence:

LEMMA 9.1 *Let E/k be a Galois extension with group G and consider a subgroup H of G with corresponding subfield F . Given $\alpha_1, \dots, \alpha_r \in F$, we have $k(\alpha_1, \dots, \alpha_r) = F$ if and only if the group leaving each of $\alpha_1, \dots, \alpha_r$ fixed is equal to H .*

For if H_1 is the subgroup fixing $\alpha_1, \dots, \alpha_r$, then H_1 corresponds to $k(\alpha_1, \dots, \alpha_r)$ and so $H_1 = H$ iff $k(\alpha_1, \dots, \alpha_r) = F$. ■

We can now show that finite separable extensions are simple:

THEOREM 9.2 (theorem of the primitive element) *If F is a finite separable extension of k then F has a primitive element, i.e. there exists $\beta \in F$ such that $F = k(\beta)$.*

Proof. When k is finite, then so is F and the conclusion certainly holds by Th. 7.3. We may therefore take k to be infinite.

Let E be a normal closure of F over k , with Galois group G , and let H be the subgroup corresponding to F . If $F = k(\alpha_1, \dots, \alpha_r)$, then H is the fixed group of $\alpha_1, \dots, \alpha_r$, and we have to find $\beta \in F$ with fixed group exactly H . Since any $\beta \in F$ has fixed group containing H , this means that every element of $G \setminus H$ must move β .

Adjoin indeterminates t_1, \dots, t_r and consider $\lambda = \sum \alpha_i t_i$. Any $\sigma \in G$ acts by $\lambda^\sigma = \sum \alpha_i^\sigma t_i$, so σ moves λ iff $\sigma \notin H$. Hence

$$\varphi(t_1, \dots, t_r) = \prod_{\sigma \notin H} (\lambda^\sigma - \lambda) \neq 0.$$

By density we can specialize t_i to $c_i \in k$ such that $\varphi(c_1, \dots, c_r) \neq 0$. So if $\beta = \sum \alpha_i c_i$, then $\beta^\sigma \neq \beta$ for $\sigma \notin H$, and it follows that $F = k(\beta)$. ■

A closer analysis shows that any finite extension $k(\alpha_1, \dots, \alpha_r)/k$, where all but at most one of the α_i is separable, is still simple (Ex. (3)), but there are examples of non-simple extensions generated by two elements (Ex. (5)). The situation is clarified by the following criterion for an extension to be simple.

THEOREM 9.3 (Steinitz) *A finite extension F/k is simple if and only if the number of fields between F and k is finite.*

Proof. Assume that $F = k(\alpha)$ and let f be the minimal polynomial for α over k . For any field E between F and k denote by p_E the minimal polynomial for α over E . It follows that $p_E | f$ and we have a correspondence $E \mapsto p_E$ between subfields and factors of f . Since a polynomial of degree n has at most 2^n monic factors (as we see by looking at a complete factorization in a splitting field), we need only show that the correspondence $E \mapsto p_E$ is injective, i.e. that E is determined by p_E .

Given E and p_E as above, let E' be the field obtained by adjoining the coefficients of p_E to k . Then $E' \subseteq E$ and p_E is irreducible over E' , hence $[F:E] = [F:E'] = \deg p_E$, and it follows that $E' = E$. Thus E is determined by p_E and we

have shown that when F/k is simple, of degree n , there are at most 2^n intermediate fields.

Conversely, assume that there are only finitely many fields between F and k . The case where k is finite has been dealt with in Th. 9.2, so we may take k to be infinite.

Given $\alpha, \beta \in F$ and $a \in k$, write $\gamma_a = \alpha + a\beta$. As a runs through k , we get an infinite family $\{\gamma_a\}$, but by hypothesis there are only finitely many fields between k and F , so there exist $a, b \in k$, $a \neq b$ such that $k(\gamma_a) = k(\gamma_b) = E$, say. Then $\gamma_a, \gamma_b \in E$, and solving the equations

$$\alpha + a\beta = \gamma_a,$$

$$\alpha + b\beta = \gamma_b,$$

for α, β we find $\alpha, \beta \in E$ and hence

$$E = k(\alpha, \beta) = k(\gamma), \quad \text{where } \gamma = \gamma_a.$$

Here α, β were arbitrary in F , so we see that any extension $k(\alpha, \beta)$ of k is simple. Now choose $\alpha \in F$ such that its degree $[k(\alpha):k]$ is maximal. If $k(\alpha) \neq F$, take $\beta \notin F \setminus k(\alpha)$. By what has been shown, we have $k(\alpha) \subset k(\alpha, \beta) = k(\gamma)$ for some $\gamma \in F$, and this contradicts the maximality of $k(\alpha)$. Hence $k(\alpha) = F$, as we wished to show. ■

This result shows again that every separable extension is simple, for it is contained in a Galois extension, by Prop. 6.1, and the intermediate fields of the latter correspond to the subgroups of the Galois group.

With every element α of an algebraic extension we can associate two elements of the ground field, its trace and norm, analogous to the trace and determinant of an endomorphism, and in fact equal to the latter for the endomorphism α_R of right multiplication. For the present we shall limit ourselves to separable extensions and leave the general case to 5.6 below. Thus let F/k be a separable extension of degree n , L any Galois extension of k containing F , and $\sigma_1 = 1, \sigma_2, \dots, \sigma_n$ the n k -homomorphisms of F into L . If $G = \text{Gal}(L/k)$ and H is the subgroup corresponding to F , then $(G:H) = n$ and $\sigma_1, \dots, \sigma_n$ is a right transversal of H in G , i.e. G has the coset decomposition $G = \bigcup H\sigma_i$. We recall that for any $\alpha \in F$, the elements $\alpha^{\sigma_1} = \alpha, \alpha^{\sigma_2}, \dots, \alpha^{\sigma_n}$ are the *conjugates* of α ; they need not lie in F but do so whenever F/k is normal. We define two mappings from F to k as follows:

$$N_{F/k}(\alpha) = N(\alpha) = \prod \alpha^{\sigma_i}, \tag{1}$$

$$T_{F/k}(\alpha) = T(\alpha) = \sum \alpha^{\sigma_i}. \tag{2}$$

$N(\alpha)$ is called the *norm* and $T(\alpha)$ the *trace* of α . Both lie in k , for any $\tau \in G$ permutes the n cosets: $H\sigma_i \mapsto H\sigma_i\tau$. Hence the elements $\alpha^{\sigma_i\tau}$ ($i = 1, \dots, n$) agree with the elements α^{σ_i} except for order, and so $N(\alpha^\tau) = N(\alpha)$, $T(\alpha^\tau) = T(\alpha)$. The following formulae are an immediate consequence of the fact that the σ_i in (1), (2) are

automorphisms:

$$\begin{aligned} N(\alpha\beta) &= N(\alpha)N(\beta), & T(\alpha + \beta) &= T(\alpha) + T(\beta), \\ N(\lambda\alpha) &= \lambda^n N(\alpha), & T(\lambda\alpha) &= \lambda T(\alpha), \quad (\lambda \in k), \\ N(1) &= 1, & T(1) &= n. \end{aligned}$$

We note that although we needed a Galois extension L to define the norm and trace, actual values are independent of the choice of L : the norm and trace can be defined in terms of any Galois extension containing F and the outcome will be the same. This is easily verified directly but also follows from Prop. 9.4 below. We note the important

TRANSITIVITY FORMULAE *If $E \supseteq F \supseteq k$ are separable extensions, then for any $\alpha \in E$,*

$$N_{E/k}(\alpha) = N_{F/k}(N_{E/F}(\alpha)), \quad T_{E/k}(\alpha) = T_{F/k}(T_{E/F}(\alpha)). \quad (3)$$

Proof of (3). Denote by L/k any Galois extension containing E/k . Let $\sigma_1, \dots, \sigma_n$ be the k -homomorphisms of F into L and extend σ_i to an automorphism of L , again denoted by σ_i (Cor. 2.5). Further, let τ_1, \dots, τ_m be the F -homomorphisms of E into L . Every k -homomorphism φ of E into L , when restricted to F , agrees with some σ_i , hence $\varphi\sigma_i^{-1}$ leaves F pointwise fixed and so $\varphi\sigma_i^{-1} = \tau_j$ for some j . Thus $\varphi = \tau_j\sigma_i$ for a unique pair of indices (j, i) . Now

$$T_{E/k}(\alpha) = \sum \alpha^{\tau_j \sigma_i} = \sum_j T_{F/k}(\alpha^{\tau_j}) = T_{F/k}(T_{E/F}(\alpha)),$$

and similarly for $N(\alpha)$. ■

PROPOSITION 9.4 *If F/k is a separable extension of degree n and $\alpha \in F$ has the minimal polynomial $x^r + a_1x^{r-1} + \dots + a_r$, then*

$$N_{F/k}(\alpha) = [(-1)^r a_r]^{n/r}, \quad T_{F/k}(\alpha) = -(n/r)a_1. \quad (4)$$

This follows by applying the transitivity formula to the tower $k \subseteq k(\alpha) \subseteq F$; the latter also shows that $r|n$. ■

The trace can be used to define an inner product on E which is often useful. Let us put

$$T(x, y) = T_{F/k}(xy). \quad (5)$$

By the properties of the trace this is bilinear and symmetric; thus we have an inner product defined on F with values in k . Its usefulness derives from the fact that it is non-singular:

PROPOSITION 9.5 *For any separable extension F/k , the inner product defined by (5) is non-singular.*

Proof. We need only show that the matrix $T(u_i, u_j)$ of the form (5) relative to a basis u_1, \dots, u_n of F over k is non-singular (cf. 8.2 in Vol. 1 or 6.1 below). Let $\sigma_1, \dots, \sigma_n$ be the k -homomorphisms of F into a normal closure E ; then

$$T(u_i u_j) = \sum_v u_i^{\sigma_v} u_j^{\sigma_v}.$$

Hence the matrix $(T(u_i u_j))$ has the form

$$(T(u_i u_j)) = DD^T, \quad \text{where } D = (d_{iv}), \quad d_{iv} = u_i^{\sigma_v} \quad (6)$$

Now by Dedekind's lemma (Lemma 5.1), the system of equations

$$\sum_v u_i^{\sigma_v} x_v = 0$$

has only the trivial solution $x_v = 0$ over E , hence D and with it $(T(u_i u_j))$ is non-singular, as claimed. ■

The value of $\det(T(u_i u_j))$ is called the *discriminant* of the extension F/k . It is not quite independent of the choice of basis; when we change the basis, the discriminant is multiplied by the square of an element of the ground field. So strictly speaking, the discriminant is a coset of the subgroup of squares in k^\times . If every element in k is a square, this tells us nothing; but for example, over a real field, if the discriminant for a given basis is positive, then it is positive for all bases.

In conclusion we record a useful property of norms and traces of finite fields.

PROPOSITION 9.6 *Let k be a finite field. Then for any finite field extension F/k the trace and norm are surjective.*

Proof. The trace is a k -linear function from F to k , and it is non-zero, by Prop. 9.5, hence it is surjective (because the image is one-dimensional). Here we have not used the fact that the field is finite, but merely that we have a Galois extension.

To prove that every element of k is a norm we observe that F/k is a Galois extension, with cyclic group generated by $\tau: x \mapsto x^q$, where $q = |k|$. Let $[F:k] = n$, and consider the homomorphism $N: F^\times \rightarrow k^\times$. We have $N(a) = a^t$, where $t = 1 + q + \dots + q^{n-1} = (q^n - 1)/(q - 1)$; hence $\ker N$ consists of the solutions in F of $a^t = 1$. Thus $|\ker N| \leq t$, and so $|\operatorname{im} N| \geq (q^n - 1)/t = q - 1$. Since $|k^\times| = q - 1$, we must have $\operatorname{im} N = k^\times$, and this shows N to be surjective. ■

Exercises

(1) Find a primitive element for $\mathbb{Q}(\sqrt[2]{2}, \sqrt[3]{3}, \sqrt[5]{5})$.

(2) Let F/k be a separable extension of degree n . Show that an element of F is primitive iff it has n conjugates in a normal closure of E/k .

- (3) By examining the proof of Th. 9.3 show that if α is algebraic and β is separable over k , then $k(\alpha, \beta)/k$ is simple. Deduce that $k(\alpha_1, \dots, \alpha_r)/k$ is simple provided that all but at most one of the α 's are separable over k .
- (4) Use Ex. (3) to show that if a non-simple finite extension exists, then there is one generated by two elements.
- (5) Let k be a field of prime characteristic p , $E = k(x, y)$ the field of rational functions in two indeterminates x, y and write $F = k(x^p, y^p)$. Show that every element of E has degree 1 or p over F ; deduce that E/F is not simple. Find infinitely many fields between E and F .
- (6) Let F/k be a separable extension of degree n . Show that the n elements a_1, \dots, a_n of F form a basis iff $(T(a_i a_j))$ is non-singular.
- (7) Show that if F/k is a Galois extension, then a_1, \dots, a_n is a basis iff the matrix (a_i^σ) is non-singular, where the columns correspond to the different elements σ of the group $\text{Gal}(F/k)$.
- (8) Let k be a finite field and F/k a finite extension. Show that every k -linear map from F to k has the form $\lambda_x : x \mapsto T(\alpha x)$, for a unique $\alpha \in F$.
- (9) Let f be a separable polynomial of degree n , irreducible over k , and let α be a zero in a splitting field. By taking a basis of the form $1, \alpha, \dots, \alpha^{n-1}$ for $k(\alpha)$ over k , show that the discriminant of $k(\alpha)$ over k can be written as a product of two Vandermonde determinants (cf. p. 191 ff, Vol. 1). Deduce that the discriminant is the product of the squares of the differences of the roots of $f = 0$.

3.10 Galois theory of equations

We have seen that Galois theory may be described as the analysis of field extensions by means of automorphism groups. However, originally it was associated with the solution of equations, and in this section we shall describe the connexion. It is based on the fact that every polynomial equation $f(x) = 0$ over a field k defines a minimal splitting field E and if f is separable, E/k is a Galois extension. In particular, in characteristic 0 (the only case considered classically) all finite extensions are separable. Even in finite characteristic the separable case is the most important. For example, in algebraic number theory the residue class fields for prime ideals are finite fields, and hence perfect. The inseparable case arises mostly in algebraic geometry.

Consider a separable polynomial f over k , and let E be a minimal splitting field. Then

$$E = k(\alpha_1, \dots, \alpha_n),$$

where $\alpha_1, \dots, \alpha_n$ are the roots of $f = 0$ in E . The Galois group $G = \text{Gal}(E/k)$ is also called the *group* of the equation $f = 0$ over k . Any automorphism of E/k is completely determined by its effect on $\Sigma = \{\alpha_1, \dots, \alpha_n\}$; moreover it must send

any root to another root, so it maps Σ into itself, and being invertible, it therefore defines a permutation of Σ . Thus any automorphism of E/k can be specified completely by the permutation of Σ it induces, and so $\text{Gal}(E/k)$ is isomorphic to a subgroup of Sym_n , the symmetric group of degree n . Of course it will not in general be the whole of Sym_n , e.g. any α_i which lies in k must stay fixed.

As an example consider the equation $x^3 - 2 = 0$ over \mathbf{Q} . If α is a root and ω is a primitive cube root of 1, then $E = \mathbf{Q}(\alpha, \omega)$ is a minimal splitting field. Any permutation of the roots defines an automorphism of E , hence the group is the full symmetric group, in agreement with the fact that $[E:\mathbf{Q}] = 6$. Over $\mathbf{Q}(\omega)$ the same equation has the cyclic group of order 3: the roots $\alpha, \alpha\omega, \alpha\omega^2$ can be permuted cyclically, but not in any other way, while over $\mathbf{Q}(\alpha)$ we have the cyclic group of order 2: we can only interchange $\alpha\omega$ and $\alpha\omega^2$, while α remains fixed.

We first note a criterion for irreducibility in terms of the Galois group.

THEOREM 10.1 *Let f be a separable polynomial over k . Then f is irreducible over k if and only if its group acts transitively on the roots.*

Proof. Let $\alpha_1, \dots, \alpha_n$ be the roots of $f = 0$ in some splitting field E . Its group G acts by permutations on the set of roots. Assume first that f is irreducible over k ; then for each $i = 2, \dots, n$, $k(\alpha_1) \cong k(\alpha_i)$ under a k -isomorphism which maps α_1 to α_i , by Prop. 2.2. This k -isomorphism can be extended to a k -automorphism of E , by Cor. 2.5. Thus we obtain an element of G mapping α_1 to α_i and so G is transitive.

Conversely, if G is transitive, and p is the minimal polynomial for α_1 over k , then $p(\alpha_1) = 0$. By transitivity there exists for each $i, 2 \leq i \leq n$, a σ in G such that $\alpha_1^\sigma = \alpha_i$. Hence $p(\alpha_i) = p(\alpha_1^\sigma) = [p(\alpha_1)]^\sigma = 0$, so α_i is also a root of $p = 0$. This holds for all i , so all the α 's are roots of $p = 0$, and since the roots are distinct, p must agree with f except for a constant factor. Thus f is irreducible over k , as claimed. ■

To explore the connexion between the group and the equation we need to know how the group changes when the field is enlarged. We begin by translating the second isomorphism theorem of group theory to fields. Let us recall the statement (Th. 4 of 9.1, p. 252, Vol. 1 or 4.1 below): Given a group G with subgroups H, N , where $N \triangleleft G$, we have $H \cap N \triangleleft H$ and

$$HN/N \cong H/(H \cap N).$$

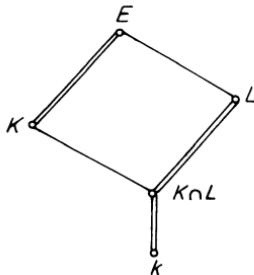
PROPOSITION 10.2 *Given a Galois extension E/k , let K and L be fields between E and k such that L/k is normal (and hence Galois). Write $\text{Gal}(E/k) = G$ and let H, N be the subgroups corresponding to K and L , respectively. Then KL/K is normal with group $\text{Gal}(KL/K) \cong HN/N$.*

Here KL stands for the subfield of E generated by K and L . The proof follows because E/KL has group $H \cap N$, which is normal in H , so that KL/K is normal with group $H/(H \cap N) \cong HN/N$. ■

In this result we were limited to extensions K/k which are finite and separable; in fact a more general result holds:

THEOREM 10.3 (on natural irrationalities) *Let f be a separable polynomial over a field k , let K be any field containing k and E a minimal splitting field of f over K . If L denotes the minimal splitting field of f over k contained in E , then E/K and L/k are normal and $\text{Gal}(E/K)$ is isomorphic to the subgroup of $\text{Gal}(L/k)$ corresponding to the field $K \cap L$.*

Briefly, the assertion is that ‘extending the field reduces the Galois group to a subgroup’. An element of L is called a *natural irrationality* for f .



Proof. Since f is separable over k , it is separable over K and so E/K and L/k are separable, hence Galois. Let $\alpha_1, \dots, \alpha_n$ be the zeros of f in E ; then $E = K(\alpha_1, \dots, \alpha_n)$, $L = k(\alpha_1, \dots, \alpha_n)$ and $\text{Gal}(E/K)$, $\text{Gal}(L/k)$ may be regarded as groups of permutations of $\{\alpha_1, \dots, \alpha_n\}$. Take $\sigma \in \text{Gal}(E/K)$; its restriction $\sigma_0 = \sigma|L$ is a homomorphism of L fixing $K \cap L \supseteq k$, hence a k -homomorphism of L and so an automorphism because L/k is normal. The mapping $\sigma \mapsto \sigma_0$ is a homomorphism from $\text{Gal}(E/K)$ to $\text{Gal}(L/k)$, as is easily checked. Its image is the subgroup corresponding to $K \cap L$, for every automorphism of L fixing $K \cap L$ extends to an automorphism of E , and it is injective, because if $\sigma_0 = 1$, then σ leaves each α_i fixed, but that can happen only when $\sigma = 1$. ■

COROLLARY 10.4 *If f is a separable polynomial over a field k , with group G of prime order, then the group of f over any extension of k is either G or 1.*

For the group must be a subgroup of G , and there are only the two possibilities stated, by Lagrange’s theorem. ■

As another application of this theorem, consider the extension $k(\omega)/k$, where k is a field of characteristic 0 and ω is a primitive n th root of 1. As we have seen,

$\mathbb{Q}(\omega)/\mathbb{Q}$ is abelian, i.e. a Galois extension with abelian group, and k may be regarded as an extension of \mathbb{Q} ; hence $k(\omega)/k$ is an abelian extension. This is still true in prime characteristic p as long as p/n .

When we have a permutation group G , we can always form the subgroup G_+ of all even permutations; this is a subgroup of index 1 or 2 in G . Correspondingly we have, for every equation over k , an extension of degree 1 or 2 which we now identify.

THEOREM 10.5 *Let k be a field of characteristic not 2 and f a separable polynomial over k . Denote the roots of $f = 0$ in a splitting field by $\alpha_1, \dots, \alpha_n$ and form*

$$\delta = \prod_{i < j} (\alpha_i - \alpha_j).$$

If G is the group of f , regarded as permutation group of the roots, then the group G_+ of even permutations corresponds to the field $k(\delta)$.

Proof. Let D be the fixed field of G_+ . Clearly $D \supseteq k$ and $\delta \in D$, hence $k(\delta) \subseteq D$. Further, $[D:k] = (G:G_+) = 1$ or 2; if this index is 1, then $k(\delta) = k$ and there is nothing to prove, so assume that $[D:k] = 2$ and take $\beta \in D \setminus k$. Then there is a permutation σ , necessarily odd, such that $\beta^\sigma \neq \beta$. But then $\delta^\sigma = -\delta$, hence $k \subset k(\delta) \subseteq D$, and since $[D:k] = 2$, it follows that $D = k(\delta)$, as we wished to prove. ■

We remark that the discriminant Δ of f equals

$$(-1)^{\binom{n}{2}} \prod_{i \neq j} (\alpha_i - \alpha_j) = \delta^2.$$

Hence we obtain

COROLLARY 10.6 *The discriminant Δ of f is a square in k if and only if the group of f over k consists of even permutations only.* ■

Examples.

1. The quadratic equation $x^2 + px + q = 0$ (in characteristic not 2) has discriminant $\Delta = p^2 - 4q$. The group over k has order 1 or 2 according as Δ is or is not a square in k .

2. Consider the cubic $f = x^3 + px^2 + qx + r$. If f is reducible, it must have a linear factor: $f = (x - \alpha)g$, where $\alpha \in k$ and g is a quadratic, to which we can apply (1). If f is irreducible, the group is transitive, and so is either Sym_3 or Alt_3 . Which it is depends on the discriminant $\Delta = -4p^3r + p^2q^2 + 18pqr - 4q^3 - 27r^2$.

We conclude this section by examining a special case going back to Lagrange, the cyclic extensions. An extension F/k is said to be *cyclic* if it is Galois, with cyclic Galois group. For example, every finite extension of a finite field is cyclic, as we

have seen in 3.8. By a *radical extension* we understand a simple extension $k(\alpha)/k$, where α is a root of a binomial equation irreducible over k :

$$x^n = a, \quad \text{where } a \in k,$$

and n is prime to $\text{char } k$. Clearly radical extensions will be of importance in the explicit solution of equations; our first observation is that they are closely related to cyclic extensions.

PROPOSITION 10.7 *Let k be a field containing a primitive n th root of 1 (and hence of characteristic prime to n). Then an extension F/k of degree n is radical if and only if F/k is cyclic.*

Proof. Let ω be a primitive n th root of 1 in k . If $F = k(\alpha)$, where $\alpha^n = a \in k$, then the equation $x^n = a$ has distinct roots $\alpha, \alpha\omega, \dots, \alpha\omega^{n-1}$ in F , and any automorphism of F/k has the form

$$\sigma: \alpha \mapsto \alpha\omega^{i(\sigma)},$$

where $i(\sigma)$ is some integer prime to n . It is easily checked that the mapping $\sigma \mapsto i(\sigma)$ is a homomorphism from $\text{Gal}(F/k)$ to \mathbf{C}_n , in fact an isomorphism, since it is injective and both sides have the same order n . This shows F/k to be a cyclic extension.

Conversely, if F/k is cyclic of degree n , with group generated by σ , let us take $\alpha \in F$ and form the ‘Lagrange resolvent’

$$(\omega, \alpha) = \alpha + \omega^{-1}\alpha^\sigma + \omega^{-2}\alpha^{\sigma^2} + \cdots + \omega^{1-n}\alpha^{\sigma^{n-1}}.$$

By Dedekind’s lemma this is non-zero for some $\alpha \in F$, and

$$(\omega, \alpha)^\sigma = \omega(\omega, \alpha). \tag{1}$$

Hence $(\omega, \alpha)^n \in k$, and by (1), (ω, α) has distinct conjugates, so there are n distinct automorphisms of $k((\omega, \alpha))/k$, the minimal equation of (ω, α) has degree n over k and $k((\omega, \alpha)) = F$. ■

In case n is a prime number p , we need not assume that ω lies in the ground field and we can say rather more:

THEOREM 10.8 *Let k be a field, p a prime number and consider the equation*

$$x^p = a, \quad \text{where } a \in k. \tag{2}$$

Either (2) has a linear factor or it is irreducible over k , according as a is or is not a p th power in k .

Proof. If a is a p th power, say $a = b^p$, then $x^p - a$ has the linear factor $x - b$. Thus if $x^p - a$ is irreducible over k , a cannot be a p th power in k .

Conversely, suppose that $x^p - a$ is reducible over k . A splitting field of (2) contains a p th root of a , say α . If $\text{char } k = p$, then $x^p - a = (x - \alpha)^p$ and by hypothesis some factor $(x - \alpha)^r$ lies in $k[x]$, where $0 < r < p$. Here the coefficient of x^{r-1} is $-r\alpha$, hence $\alpha \in k$ and so $a = \alpha^p$ is a p th power in k . If $\text{char } k \neq p$, then any splitting field will also contain a primitive p th root of 1, say ω . Now by hypothesis $x^p - a$ has a non-trivial factorization over k :

$$x^p - a = fg,$$

where f, g are monic of positive degrees r, s respectively and $r + s = p$. The constant term of f is a product of r roots $\omega\alpha^v$ and hence is of the form $b = \alpha^r\omega_1$, where $\omega_1^p = 1$. Thus $b^p = \alpha^{pr} = \alpha^r$, and since $0 < r < p$, there exist $u, v \in \mathbb{Z}$ such that $ru + pv = 1$. It follows that $a = \alpha^{ru}\alpha^{pv} = b^{pu}\alpha^{pv} = (b^u\alpha^v)^p$, which shows a to be a p th power in k , as claimed. ■

If k contains a primitive p th root of 1, then (2) is either irreducible or it splits completely into linear factors, by Prop. 10.7, but in general this need not be so. For example, if $\alpha = 2^{1/3}$, then

$$x^3 - 2 = (x - \alpha)(x^2 + \alpha x + \alpha^2),$$

and here the second factor on the right is irreducible over $\mathbf{Q}(\alpha)$, for its zeros are not real.

We remark that in the case of Prop. 10.7, we have $(\omega^i, \alpha)^\sigma = \omega^i(\omega^i, \alpha)$, hence the n th powers of the resolvents are in the ground field. For an equation of prime degree p , the p th powers of the resolvents formed from the roots satisfy an equation of degree $p - 1$ whose coefficients are rational functions of a root of an equation of degree $(p - 2)!$ over the ground field. This is at the basis of Lagrange's method of solving cubics (cf. Ex. (8), 3.11 below). However, when $p = 5$, the resolvent has degree $3! = 6$, so the method cannot be used for $p \geq 5$, and as we shall see in 3.11 below, the general equation of degree 5 or higher cannot be solved by radicals.

Exercises

- (1) Let f be an irreducible polynomial over k . Show that in a normal extension of k , f splits (if at all) into factors that are all of the same degree and are conjugate over k .
- (2) Let F/k be a Galois extension with group G . Show that if the adjunction of α to k reduces G to a subgroup H , then the degree of α is a multiple of $(G:H)$.
- (3) Show that an equation is normal iff all its roots can be expressed rationally in terms of a single one.
- (4) Show that if α satisfies the equation $x^3 - 3x + 1 = 0$, then so does $\alpha^2 - 2$. Hence find its group and solve the equation over \mathbf{Q} in terms of radicals. (Hint. Put $x = u + v$.)

- (5) Show that an abelian transitive permutation group acts regularly (i.e. each permutation either moves all symbols or none). Deduce that an irreducible equation with abelian group is normal.
- (6) Suppose that $x^4 - ax^2 + b = 0$ is irreducible over \mathbf{Q} . If k is an extension containing no root, show that over k the group is C_4 if $a^2/b - 1$ is a square in k , the Klein 4-group if b is a square, C_2 if both hold and the dihedral group of order 8 if neither holds. Find the group of $x^4 - 2$ over \mathbf{Q} .
- (7) Let k be a field and p a prime different from $\text{char } k$. Show that for any $a \in k$, the equation $x^p = a$ splits into linear factors over k iff a is a p th power and k contains a primitive p th root of 1.
- (8) Let k be a field of prime characteristic p and F/k a cyclic extension of degree p . Show that if σ is a generating automorphism, then the linear transformation $S: \alpha \mapsto \alpha - \alpha^\sigma$ is nilpotent, and hence find an element α in the kernel of S^2 but not of S . Show that $\beta = \alpha/(\alpha^\sigma - \alpha)$ satisfies $\beta^\sigma = \beta + 1$; deduce that β satisfies an equation $x^p - x - a = 0$.
- (9) Let $V = (\alpha_i^{j-1})$ be the Vandermonde matrix formed from the roots α_i of a separable equation of degree n . By evaluating $\det(V^T V)$ show that the discriminant is the determinant of (s_{i+j-2}) , where s_r is the r th power sum of the roots.

3.11 The solution of equations by radicals

The familiar method of solving quadratic equations by completing the square was known in antiquity. Examples of cubics were solved by Diophantos (AD 300); the complete solution is due to Scipio Ferro (1515) and Niccolo Tartaglia, published by Cardano in his *Ars Magna* (1545) (cf. Ore 1953). This was soon followed by the general solution of the quartic equation (due to Ferrari, later also Gregory) and attempts continued during the 17th century to obtain the solution of the general quintic by radicals. At the end of the 18th century P. Ruffini published what was claimed to be a proof of the impossibility of solving the general quintic by radicals; the proof was incomplete, although it contained some of the ideas utilized later. This result was finally proved rigorously by N. H. Abel in 1826. His proof becomes much easier to understand with the help of Galois theory, which was developed partly with the object of giving a precise classification of the equations soluble by radicals.

An equation $f = 0$ is said to be *soluble by radicals* over a given field k , if there exists a finite tower of radical extensions

$$k = k_0 \subset k_1 \subset \cdots \subset k_r, \quad (1)$$

such that f splits completely in k_r ; (1) is then called a *root tower* for f over k .

We can now establish the connexion with soluble groups; we recall from Vol. 1

(Th. 3, **9.5**, p. 268) that a finite group is soluble iff it has a composition series with factors of prime order, or equivalently, one with abelian factors.

THEOREM 11.1 (Galois) *An equation $f = 0$ over a field k of characteristic 0 is soluble by radicals if and only if the group of f over k is soluble.*

Proof. Assume that f has a soluble group, of order n say; we shall use induction on n to show that f is soluble by radicals. In the first place, $x^n - 1$ has an abelian, hence soluble, group of order $\varphi(n) < n$, by Th. 7.5. By induction we can find a root tower (1) such that k_r contains a primitive n th root of 1. Now f has a soluble group over k_r , and by Th. 10.3, the group of f over k_r is a subgroup, which is still soluble, and so has a normal chain of subgroups with cyclic factors. This corresponds to a tower of fields

$$k_r \subset k_{r+1} \subset \cdots \subset k_m$$

where k_i/k_{i-1} is cyclic of order dividing n and hence is radical, by Prop. 10.7. Thus $f = 0$ is soluble by radicals.

Conversely, let $f = 0$ be soluble by radicals and consider a root tower (1) for f , where $[k_i:k_{i-1}]$ is prime for $i = 1, \dots, r$. If $[k_r:k] = n$, let ω be a primitive n th root of 1 and replace the tower (1) by

$$k \subseteq K_0 \subseteq K_1 \subseteq \cdots \subseteq K_r, \quad \text{where } K_i = k_i(\omega). \quad (2)$$

By Th. 10.8, the steps which have not become trivial are still radical of prime degree. We now extend this tower to a Galois extension as follows. We know that K_1/k is Galois. If K_2/K_1 is a minimal splitting field of $x^p - a$, $a \in K_1$, we adjoin all the roots of all equations $x^p = a^\tau$, where τ ranges over $\text{Gal}(K_1/k)$. This gives a Galois extension containing K_2 which can be reached by a root tower. Continuing in this way we get another root tower (2) in which K_r/k is Galois. Again by Prop. 10.7, K_i/k_{i-1} is cyclic while K_0/k is abelian; hence $\text{Gal}(K_r/k)$ has a normal chain with abelian factors and so is soluble. Moreover, K_r contains a minimal splitting field E of f over k ; now $\text{Gal}(E/k)$ is a homomorphic image of $\text{Gal}(K_r/k)$ and so is soluble. ■

We have seen in Vol. 1 (p. 269f.) that the symmetric groups of degree 3 and 4 are soluble, hence every equation of degree 3 or 4 is soluble. On the other hand the symmetric group Sym_5 is insoluble, since Alt_5 is simple. In Th. 4, **9.5** of Vol. 1 (p. 269) we proved more generally that Alt_n is simple for $n \geq 5$; for the case $n = 5$ this may be seen directly as follows. We first list the 120 elements of Sym_5 by conjugacy classes— C_1 : 1; C_2 : $\binom{5}{2} = 10$ elements of type (1 2); C_3 : $2\binom{5}{2} = 20$ elements of type (1 2 3); C_4 : 30 elements of type (1 2 3 4); C_5 : 15 elements of type (1 2)(3 4); C_6 : 20 elements of type (1 2 3)(4 5); C_7 : 24 elements of type (1 2 3 4 5). The alternating group contains C_1, C_3, C_5, C_7 and here C_1, C_3, C_5 remain conjugacy

classes, while C_7 splits into two classes of 12 each. We thus obtain as class equation for Alt_5 :

$$60 = 1 + 20 + 15 + 12 + 12. \quad (3)$$

For a normal subgroup we have to find a union of conjugacy classes including C_1 , and thus some of the numbers on the right of (3), including 1, have to add up to a factor of 60; this is easily seen to be impossible except for the trivial cases 1, 60. This shows Alt_5 to be simple, and it shows incidentally that no Sym_n for $n \geq 5$ can be soluble, because every subgroup of a soluble group is again soluble.

As we saw in 3.6, the general quintic

$$x^5 - e_1 x^4 + \cdots - e_5 = 0 \quad (4)$$

has the symmetric group over $k(e_1, \dots, e_5)$, where the e 's are independent indeterminates; it is therefore insoluble. To be precise, we can reduce the group to Alt_5 by adjoining the square root of the discriminant, but no further reduction is possible, because Alt_5 is simple.

It is also possible to construct equations with symmetric group over \mathbf{Q} . To do so we first show that when an equation over \mathbf{Z} is reduced mod p , its group is replaced by a subgroup.

THEOREM 11.2 *Let A be a UFD with field of fractions K , and let \mathfrak{p} be a prime ideal of A , so that $\bar{A} = A/\mathfrak{p}$ is an integral domain, whose field of fractions is denoted by k . If f is a monic polynomial over A , \bar{f} the corresponding polynomial over \bar{A} , where \bar{f} is separable, then f is separable and if its group over K is G , then the group Γ of \bar{f} over k is a subgroup of G , as a permutation group of the roots.*

Proof. It is clear that if f and f' have a common factor, then so do \bar{f} and \bar{f}' , hence f must be separable. Let us denote its zeros in a splitting field E over K by $\alpha_1, \dots, \alpha_n$ and with indeterminates t_1, \dots, t_n put $\lambda = \sum t_i \alpha_i$. For any permutation σ of $1, 2, \dots, n$ we define the action of σ on λ by

$$\lambda^\sigma = \sum t_{i\sigma}^{-1} \alpha_i = \sum t_i \alpha_{i\sigma}. \quad (5)$$

Next form the polynomial

$$\varphi = \prod_{\sigma} (x - \lambda^\sigma),$$

where σ ranges over all $n!$ permutations of $1, 2, \dots, n$. Clearly its coefficients are symmetric functions in the α 's and so belong to $K(t_1, \dots, t_n)$. In fact it is clear from the construction that the coefficients are polynomials in the t 's with coefficients in A . Let us factorize φ over $A[t_1, \dots, t_n]$:

$$\varphi = \varphi_1 \varphi_2 \dots \varphi_r. \quad (6)$$

By inertia (Th. 7.2) the φ_i are irreducible over $K(t_1, \dots, t_n)$. Since λ is a zero of φ , it must be a zero of some φ_i , say $\varphi_1(\lambda) = 0$. We claim that the group G of f over K is

precisely the group of all permutations which map φ_1 into itself. For let $\sigma \in G$; then σ maps λ to λ^σ , satisfying the same irreducible equation. Hence φ_1 and φ_1^σ have a common factor and so coincide. Conversely, if $\varphi_1^\sigma = \varphi_1$, then λ^σ is again a zero of φ_1 , hence $\sigma \in G$.

We now pass to the residue class ring \bar{A} and obtain a factorization

$$\bar{\varphi} = \bar{\varphi}_1 \dots \bar{\varphi}_r;$$

of course the factors $\bar{\varphi}_i$ on the right may well be reducible over k . The permutations of G are precisely the permutations transforming $\bar{\varphi}_1$ into itself while any other permutation maps φ_1 to φ_i ($i \neq 1$). The permutations of Γ transform an irreducible factor of $\bar{\varphi}_1$, viz. that containing $x - \lambda$, into itself, and so must transform $\bar{\varphi}_1$ into itself. Hence Γ is a subgroup of G . ■

We shall apply this result with $A = \mathbf{Z}$. Given $f \in \mathbf{Z}[x]$ and a prime p , let us factorize $f \pmod{p}$. If $f \equiv f_1 \dots f_r \pmod{p}$, where f_i is irreducible of degree d_i , then the group of f contains a permutation whose cycle structure is d_1, d_2, \dots, d_r . For the group Γ of $f \pmod{p}$ is cyclic, as Galois group of an extension of finite fields. The orbits of Γ correspond to the irreducible factors of f , so the generating permutation of Γ consists of a d_1 -cycle, a d_2 -cycle, ..., a d_r -cycle, as claimed. It follows that if f has a factorization into irreducible factors of degrees $d_1, d_2, \dots, d_r \pmod{p}$, then the group of f contains a permutation of type d_1, d_2, \dots, d_r .

We shall also need a result on generating sets for Sym_n .

LEMMA 11.3 *Any transitive subgroup of Sym_n containing a 2-cycle and an $(n-1)$ -cycle is the whole group.*

Proof. Let H be the subgroup; by suitable numbering we can write the $(n-1)$ -cycle as $\rho = (2 3 \dots n)$. Since H is transitive, we can take the 2-cycle to include 1, say $(1 i)$. By conjugating with ρ we obtain $(1 2), (1 3), \dots, (1 n)$ and these transpositions generate the whole group. ■

We can now construct equations with symmetric group over \mathbf{Q} .

THEOREM 11.4 *For every $n \geq 1$ there is an irreducible equation of degree n over \mathbf{Q} whose group is the symmetric group of degree n .*

Proof. For any prime p and any $n \geq 1$ there are irreducible polynomials of degree n over \mathbf{F}_p , by Th. 8.4. We choose three polynomials over \mathbf{Z} as follows. f_1 is a product of an irreducible factor of degree $n-1$ by a linear factor $\pmod{2}$, and f_2 a product of an irreducible quadratic factor and $n-2$ linear factors $\pmod{3}$. Both f_1, f_2 may be taken monic; then $3f_1 - 2f_2$ is again monic and is congruent to $f_1 \pmod{2}$ and to $f_2 \pmod{3}$. The same is true of

$$f = 3f_1 - 2f_2 + 6f_3,$$

for any polynomial f_3 of degree $n - 1$ over \mathbf{Z} . We now choose f_3 so that all coefficients of f after the first are divisible by 5, but the absolute term is not divisible by 25. Then f is irreducible, by Eisenstein's criterion, hence its group is transitive. By Th. 11.2 and the remark following it, the group contains an $(n - 1)$ -cycle and a transposition, and therefore, by Lemma 11.3, it must be the whole of Sym_n . ■

By these methods it can also be shown that for any $n \geq 1$ there is an equation over \mathbf{Q} whose group is the alternating group of degree n . But it is still an open question whether every finite group occurs as the Galois group of an extension of \mathbf{Q} .

Returning to ruler-and-compass constructions (3.1), we see that the elements of a Galois extension are constructible whenever the Galois group is a 2-group, for since any 2-group is soluble (Vol. 1, p. 291), there is then a root tower in which all steps have degree 2. This has an application to the construction of a regular n -gon.

We need to solve the equation $\Phi_n(x) = 0$, for if ζ is a root, then $\cos(2\pi/n) = \frac{1}{2}(\zeta + \zeta^{-1})$. By Th. 7.5, its group is abelian of order $\varphi(n)$. Writing $n = \prod p_i^{v_i}$, we have $\varphi(n) = \prod p_i^{v_i-1}(p_i - 1)$, and this is a power of 2 precisely when each odd prime divisor occurs at most once and has the form $p = 2^r + 1$. If r has an odd factor, say $r = cd$, where d is odd, then $2^r + 1$ is divisible by $2^c + 1$, hence for a prime, r has to be a power of 2. The primes of the form $F_m = 2^{2^m} + 1$ are called *Fermat primes*; the first few are 3, 5, 17, 257, 65 537, but it is not known whether there are others (F_5 was proved composite by Euler in 1732).

THEOREM 11.5 *A regular n -gon can be constructed by ruler and compasses precisely when each odd prime factor of n occurs only once and is a Fermat prime.* ■

This result was essentially known to Gauss, who gave an explicit construction of the regular heptadecagon in 1796 (at the age of 19).

For irreducible equations of prime degree there is another criterion for solubility, also due to Galois. In these cases the group has a rather remarkable form, whose description depends on the following result.

PROPOSITION 11.6 *Let G be a transitive permutation group of prime degree p . Then the following conditions are equivalent:*

- (a) G is soluble;
- (b) G has a transitive normal subgroup T of order p which is its own centralizer;
- (c) G can be written as a group of affine transformations over \mathbf{F}_p :

$$x^\sigma \equiv ax + b \pmod{p}, \quad (a \not\equiv 0 \pmod{p}), \quad (7)$$

which includes the subgroup T of all translations, $x \mapsto x + b$;

(d) every element $\neq 1$ of G fixes at most one symbol.

When this is so, G has a cyclic subgroup M such that $MT = G$, $M \cap T = 1$, and the order of M divides $p - 1$. Every non-trivial normal subgroup of G has the form NT , where N is a subgroup of M .

Proof. We remark that as a transitive group of prime degree p , G has order divisible by p . Further, it may be regarded as a subgroup of Sym_p , which has order $p!$, so the highest power of p dividing $|G|$ is the first. Hence any Sylow p -subgroup of G has order p .

We begin by proving the equivalence of (a), (b) and (c).

(a) \Rightarrow (b). Let $N \triangleleft G$; since G is transitive, it permutes the orbits under the action of N transitively, so they must all have the same size. But the total number of symbols permuted is a prime, hence each orbit of N has either 1 symbol or p symbols, i.e. N is either the trivial group or it acts transitively. Now G is soluble, hence any minimal normal subgroup is elementary abelian, by Th. 5, 9.6, p. 276 of Vol. 1. Such a subgroup is transitive and of degree p , so it must be the Sylow p -subgroup T of G , which is therefore normal and contained in every non-trivial normal subgroup N of G . The subgroup T is of order p and hence cyclic. Let τ be a generator; as a permutation this is a p -cycle. If $\sigma \in G$ commutes with τ , then σ is a permutation of p symbols which commutes with the p -cycle τ . But the only such permutations are the powers of τ , therefore T is its own centralizer.

(b) \Rightarrow (c). By hypothesis T is of order p , hence cyclic. Let τ be a generator of T ; then for any $\sigma \in G$ there exists $a \in \mathbb{Z}$ such that

$$\sigma^{-1}\tau\sigma = \tau^a. \quad (8)$$

Moreover, a is determined uniquely $(\bmod p)$. We denote the unique residue class determined by a $(\bmod p)$ in (8) by a_σ ; then

$$\sigma \mapsto a_\sigma \quad (9)$$

is a homomorphism from G to \mathbf{F}_p^\times . If $a_\sigma = 1$, then $\sigma^{-1}\tau\sigma = \tau$ and so $\sigma \in T$ by hypothesis; conversely, if $\sigma \in T$, then $a_\sigma = 1$, so the kernel of (9) is T . Let us number the symbols permuted as $0, 1, \dots, p - 1$ in such a way that τ^v maps 0 to v . Then $x^\tau = x + 1$ and $x^{\tau\sigma} = x^{\sigma\tau\sigma} = x^{\sigma\tau}$, i.e. $(x + 1)^\sigma = x^\sigma + a$. Taking $x = 0$, we find $1^\sigma = 0^\sigma + a$, hence by induction on x ,

$$x^\sigma = ax + b, \quad \text{where } b = 0^\sigma. \quad (10)$$

(c) \Rightarrow (a) is clear: the translations form a cyclic normal subgroup T in G with quotient isomorphic to a subgroup of \mathbf{F}_p^\times .

(c) \Rightarrow (d). The affine transformation (7), where $\sigma \neq 1$, leaves at most one symbol fixed, because the congruence $ax + b = x$ does not hold identically and so has at most one solution.

(d) \Rightarrow (b). By the orbit formula (Th. 1 of 3.6, p. 63, Vol. 1), the ‘average number’ of symbols fixed by a permutation is the number of orbits, which is 1. Since the

identity fixes p symbols, and the other permutations fix at most one symbol, there are exactly $p - 1$ permutations fixing no symbol. Each is necessarily a cycle of length p ; if one of them is denoted by τ , the others are powers of τ and the subgroup T generated by τ has order p . Since there are no other p -cycles in G , T is normal in G , and if $\sigma \in G$ satisfies $\sigma\tau = \tau\sigma$, then either σ fixes no symbol and so lies in T , or σ fixes a ; then σ also fixes $a^{\tau} \neq a$, and so $\sigma = 1$. Thus T is its own centralizer in G .

Finally let G be as in (c) and let M be the subgroup of multiplications in G . Then $M \cap T = 1$ and $MT = G$, as is easily seen, and M is isomorphic to a subgroup of \mathbf{F}_p^{\times} and therefore has order dividing $p - 1$. Every non-trivial normal subgroup of G contains T and so these normal subgroups correspond to the subgroups of $G/T \cong M$. ■

We remark that having prime degree is a strong condition on G . The group G is the whole affine group $\text{Aff}_1(\mathbf{F}_p)$ precisely when G is doubly transitive (i.e. transitive on ordered pairs of symbols). To apply the above result, suppose that we have an irreducible equation of prime degree p . Its group is transitive of degree p , by Th. 10.1, and condition (d) means that any permutation fixing two roots fixes all, so by the fundamental theorem (Th. 6.2), all roots can be expressed rationally in terms of any two, whenever the equation is soluble. Thus we obtain from Prop. 11.6 Galois' criterion for the solubility of equations of prime degree:

THEOREM 11.7 *Let*

$$f = 0 \quad (11)$$

be an irreducible equation of prime degree p over a field k of characteristic 0, and let E be a minimal splitting field of f . Then (11) is soluble by radicals if and only if E can be generated over k by any two roots of (11). When this is so, its group is a group of affine transformations mod p including all translations, there is a cyclic extension F of k over which E has degree p , and every proper normal subextension of E/k is contained in F . ■

To apply the above results we construct some equations with group Sym_5 over \mathbf{Q} .

Examples

1. The equation

$$x^5 + 10x^3 - 15 = 0$$

is irreducible by Eisenstein's criterion; it is $x^5 - 1 \pmod{2}$, which is $x - 1$ times an irreducible factor of degree 4, and it is $x^5 + x^3 \pmod{3}$, which is x^3 times an irreducible factor of degree 2. By Th. 11.2 its group must be Sym_5 .

2. Consider the equation

$$f(x) = x^5 - 5x + 1 = 0. \quad (12)$$

This is again irreducible by Eisenstein's criterion, as we see by putting $x = y - 1$. By Descartes' rule (Vol. 1, p. 172), (12) has at most three real roots, and in fact $f(-2), f(-1), f(1), f(2)$ have signs $-$, $+$, $-$, $+$ respectively, hence there are exactly three real and two complex roots. If (12) were soluble, all its roots could be expressed rationally in terms of any two, by Th. 11.7; choosing two real roots, we obtain a contradiction. Thus the group is Alt_5 or Sym_5 and the former can be ruled out because complex conjugation interchanges the complex roots and so is an odd permutation.

Exercises

- (1) Show that a real algebraic number admits quadrature iff the Galois group of its minimal equation over \mathbf{Q} is a 2-group.
- (2) Carry out the construction of a regular n -gon for $n = 3, 4, 5, 6$.
- (3) Show that the binomial equation $x^n = a$ is soluble over \mathbf{Q} but not necessarily abelian. Describe the group when n is prime, using Th. 11.7.
- (4) Show that the equation $x^p - x - t = 0$ is not soluble by radicals over $\mathbf{F}_p(t)$, though its group is cyclic.
- (5) Show that if a separable equation over a field of prime characteristic p is soluble by radicals, then its group G is soluble (of order at most $p(p - 1)$) and the converse holds when all prime factors of $|G|$ are less than p .
- (6) Show that a transitive permutation group G of prime degree p contains a cycle of length p . If G also contains a transposition, then it must be the full symmetric group. (*Hint.* G contains an element of order p .)
- (7) Show that $x^4 - 5 = 0$ has the dihedral group of order 8 over \mathbf{Q} .
- (8) Let $f = 0$ be a cubic with roots $\alpha_1, \alpha_2, \alpha_3$ over a field k containing a primitive cube root ω of 1. Verify that $\theta = (\alpha_1 + \alpha_2\omega + \alpha_3\omega^2)^3$ is of degree 1 or 2 over k , and show that $k \subseteq k(\theta) \subseteq k(\theta^{1/3}) = k(\alpha_1, \alpha_2)$ is a root tower for the equation. What is the analogue for quartics? Show that for a quartic, $(\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4)^2$ has degree 1 or 3.
- (9) Let $f = 0$ be a quartic over a field k with a primitive sixth root of 1, and denote its roots in some splitting field by $\alpha_1, \dots, \alpha_4$. Write $\gamma = \alpha_1 + \alpha_2$, $\beta = \alpha_1\alpha_2 + \alpha_3\alpha_4$ and let δ be the square root of the discriminant. Show that $k \subseteq k(\delta) \subseteq k(\delta, \beta) \subseteq k(\delta, \beta, \gamma) \subseteq k(\alpha_1, \alpha_2, \alpha_3)$ is a root tower.
- (10) Let $f = 0$ be a quintic with roots $\alpha_1, \dots, \alpha_5$ and denote by ω a primitive fifth root of 1. Show that in general $\theta = (\alpha_1 + \alpha_2\omega + \dots + \alpha_5\omega^4)^5$ has degree dividing 24, but when f is irreducible and soluble, then θ has degree dividing 4.

Further exercises on Chapter 3

- (1) Let F_1 and F_2 be subfields of a field E . Show that if $F_1 \subseteq F_2$ and $[E:F_1] < \infty$, then $[E:F_1] \geq [E:F_2]$, with equality iff $F_1 = F_2$.
- (2) Show that if $k \subseteq F_i \subseteq E$ and F_i is finite over k ($i = 1, 2$), then the subfield L generated by F_1 and F_2 is finite over k .
- (3) Show that $\mathbf{Q}(\sqrt{2}, \sqrt{-1})$ is a normal extension of \mathbf{Q} and that $x^4 - 2x^2 + 9 = 0$ is a normal equation for it (i.e. the minimal equation of a primitive element; this is also called the *resolvent*).
- (4) Prove that $k[x]$ is a UFD by using splitting fields.
- (5) Let F/k be a finite extension and F perfect. Show that k is perfect.
- (6) Let k be a field of prime characteristic p and $\alpha: x \mapsto x^p$ the Frobenius mapping. Show that there is a strictly descending tower of subfields $k \supseteq \text{im } \alpha \supseteq \text{im } \alpha^2 \supseteq \dots$, whose intersection k_0 is a perfect subfield of k , and show that every perfect subfield of k is contained in k_0 .
- (7) (König–Rados) Consider the equation

$$\alpha_0 + \alpha_1 x + \cdots + \alpha_{q-2} x^{q-2} = 0, \quad (\text{i})$$
 over \mathbf{F}_q . By examining the product of the Vandermonde matrix with rows $1, \beta, \dots, \beta^{q-2}$ ($\beta \in \mathbf{F}_q^\times$) and the circulant matrix whose rows are cyclic permutations of the coefficients $\alpha_0, \alpha_1, \dots, \alpha_{q-2}$, show that the number of non-zero roots of (i) in \mathbf{F}_q is $q - 1 - r$, where r is the rank of the circulant matrix.
- (8) Show that in a Galois extension of odd order the discriminant is a square. Conversely, show that in a cyclic Galois extension of even order the discriminant is not a square.
- (9) Let k be a finite field and $n \geq 1$. Show that there is an irreducible polynomial of degree n over k in which the coefficient of x is non-zero. (Hint. Consider the equation for x^{-1} .)
- (10) Let F/k be a finite extension and $|k| = q$. Show that $T(\alpha^q) = T(\alpha)$ for all $\alpha \in F$. Deduce that $T(\alpha) = 0$ iff the equation $x^q - x - \alpha = 0$ has a root in F .
- (11) (Kronecker) Let $f = 0$ be an irreducible equation over \mathbf{Q} of odd prime degree p . Show that if f is soluble by radicals, then the number of its real roots is either 1 or p . Moreover, when $p \equiv 3 \pmod{4}$, the number of real roots is 1 or p according as the discriminant is negative or positive.
- (12) Let F/k be a Galois extension. If $\sigma \in \text{Gal}(F/k)$, $\sigma \neq 1$, find $a_1, \dots, a_n, b_1, \dots, b_n \in F$ such that $\sum a_i(b_i - b_i^\sigma) = 1$. Show that for $b_0 = 1$ and suitably chosen a_0 ,

$$\sum_0^n a_i b_i = 1, \quad \sum_0^n a_i b_i^\sigma = 0.$$

By doing this for each $\sigma \neq 1$ and multiplying the results together, obtain $u_1, \dots, u_r, v_1, \dots, v_r \in F$ such that $\sum u_i v_i^\sigma = \delta_{10}$. (This property can be used to characterize a notion of Galois extension of commutative rings, cf. Chase *et al.* (1965).)

(13) Let p be a prime and n an integer prime to p . Show that $\Phi_n(x^p) = \Phi_{pn}(x)\Phi_n(x)$. What happens if $p|n$?

(14) (Vahlen–Capelli) Let F/k be a separable extension, say $F = k(\alpha)$, where α has the minimal polynomial f over k , and for any polynomial g over k let a complete factorization of $g(x) - \alpha$ over F be

$$g(x) - \alpha = \prod g_i(x).$$

Using the formula $f(x) = N(x - \alpha)$, prove that

$$f(g(x)) = \prod N(g_i(x)), \quad (\text{ii})$$

where the norm is relative to the extension $F(x)/k(x)$. If $p(x)$ is an irreducible factor of $f(g(x))$ over k , show that $g_i(x)|p(x)$ for some i , and hence by passing to a splitting field of f over k , deduce that $N(g_i(x))|p(x)$. Use this fact to prove that (ii) is a complete factorization of $f(g(x))$ over k ; in particular, show that $f(g(x))$ is irreducible over k iff $g(x) - \alpha$ is irreducible over F .

(15) (Vahlen–Capelli) Show that $x^n = a$ (where $n > 1$, $a \neq 0$) is reducible over a field k iff either (i) $a = b^d$ for some $b \in k$ and $d|n$, $d > 1$, or (ii) $4|n$ and $a = -4c^4$. (For case (ii) use the identity $x^{4m} + 4c^4 = (x^{2m} - 2cx^m + 2c^2)(x^{2m} + 2cx^m + 2c^2)$. In the other direction use induction on n and Ex. (14); note also the identity $u^4 + v^4 + (u - v)^4 = 2[u^3 + v^3]/(u + v)^2$.)

(16) (E. Artin) Let k be a field of characteristic 0 and K/k a finite extension. Show that if K is algebraically closed then $[K:k] \leq 2$ and $K = k(\sqrt{-1})$. (The fact that $\text{char } k = 0$ can actually be proved from the other assumptions. Hint. Use Ex. (15).)

(17) An equation is said to be *reciprocal* if with α , $1/\alpha$ is also a root. Show that if an irreducible equation over \mathbb{Q} has a complex root of absolute value 1, then the equation is reciprocal of even degree.

(18) Show that the group of a reciprocal equation of degree $2m$ or $2m + 1$ is of order at most $2^m \cdot m!$. Determine the possible groups for a reciprocal quartic.

(19) Let $f = 0$ be an irreducible cubic over \mathbb{Q} with three real roots. Show that there is no root tower for $f = 0$ consisting entirely of real fields. (This is the ‘casus irreducibilis’, which expresses the fact that in this case none of the roots can be expressed in terms of real radicals alone, in contrast to the case of a cubic with only one real root.)

(20) Let E/k be separable of degree r . If $E = k(\alpha)$ and $\beta \in E$, find a polynomial f over k of degree less than r such that $\beta = f(\alpha)$. (Hint. Use Lagrange interpolation, given that $f(\alpha^\sigma) = \beta^\sigma$.)

- (21) Let p be a prime and a an integer. The Legendre symbol $\left(\frac{a}{p}\right)$ is defined as 0 if $p|a$, 1 if $x^2 \equiv a \pmod{p}$ has a solution and -1 otherwise; in the last two cases a is a *quadratic residue* (resp. *non-residue*) mod p . Taking a to be an integer mod p , verify that the map of \mathbf{F}_p^\times into \mathbf{C}_2 given by $a \mapsto \left(\frac{a}{p}\right)$ is a homomorphism. Let p be an odd prime and z a generator of the multiplicative group \mathbf{F}_p^\times . Show that $z^{(p-1)/2} = -1$ and hence deduce *Euler's criterion*: For any d prime to p , $d^{(p-1)/2} \equiv \left(\frac{d}{p}\right) \pmod{p}$.

- (22) Let p, q be distinct odd primes and denote by w a primitive p th root of 1 in an extension of \mathbf{F}_q . For any $a \in \mathbf{F}_p^\times$ define the *Gaussian sum* (in an extension of \mathbf{F}_q) as

$$\tau(a) = \sum_{x \in \mathbf{F}_p^\times} \left(\frac{x}{p}\right) w^{ax}.$$

Prove (i) $\tau(a) = \left(\frac{a}{p}\right) \tau(1)$, (ii) $\tau(1)^q = \tau(q)$, (iii) $\tau(1)^2 = (-1)^{(p-1)/2} p$. (*Hint.* Evaluate $\sum w^{ax}$ and use Euler's criterion.)

- (23) For any odd n , put $e(n) \equiv (n-1)/2 \pmod{4}$. Use Ex. (22) to show that $\tau(1)^{q-1} = \left(\frac{q}{p}\right)$.

By evaluating $\tau(1)^{q-1} = [\tau(1)^2]^{(q-1)/2}$ in two ways, prove the *law of quadratic reciprocity*:

$$\left(\frac{p}{q}\right) \left(\frac{q}{p}\right) = (-1)^{e(p)e(q)}.$$

- (24) For any odd n , put $\omega(n) \equiv (n^2 - 1)/2 \pmod{8}$. Let α be a primitive eighth root of 1 in an extension of \mathbf{F}_p and put $\beta = \alpha + \alpha^{-1}$. Show that $\beta^2 = 2$, and using the Frobenius endomorphism $x \mapsto x^p$ and Euler's criterion to evaluate β^{p-1} in two ways, prove that $\left(\frac{2}{p}\right) = (-1)^{\omega(p)}$. (This formula, together with the law of quadratic reciprocity and Euler's criterion, enables one to evaluate Legendre symbols, e.g.

$$\left(\frac{29}{43}\right) = \left(\frac{43}{29}\right) = \left(\frac{14}{29}\right) = \left(\frac{2}{29}\right) \left(\frac{7}{29}\right) = -\left(\frac{7}{29}\right) = -\left(\frac{1}{7}\right) = -1,$$

thus $x^2 \equiv 29 \pmod{43}$ has no solution.)

4

Modules

Modules are to a general ring what vector spaces are to a field, but whereas a vector space over a given field is determined up to isomorphism by its dimension, there is a much greater multiplicity of modules. Another way of regarding modules is as abelian groups (written additively) with operators. This means that much of general group theory applies, and after recalling the isomorphism theorems (proved in Ch. 9 of Vol. 1) we treat a number of special situations. Semisimple modules come closest to vector spaces; they are direct sums of simple modules, but we have to bear in mind that over a given ring there may be more than one type of simple module. In the free modules we have another generalization of vector spaces. The homological treatment of module theory requires the notions of projective and injective module, and they can usefully be introduced here, as they will occur again later in Ch. 9, although their main use will be in Vol. 3. Other important notions introduced here are those of matrix ring and tensor product.

4.1 The category of modules over a ring

Let R be any ring (as always, with 1). We recall from Ch. 10, Vol. 1 that a *right R -module* is an abelian group M with a mapping from $M \times R$ to M , $(x, r) \mapsto xr$, such that

- M.1** $(x + y)r = xr + yr, \quad x, y \in M,$
- M.2** $x(r + s) = xr + xs, \quad r, s \in R,$
- M.3** $x(rs) = (xr)s,$
- M.4** $x1 = x.$

Of these rules **M.1** states that each $r \in R$ defines an endomorphism of the abelian group M , so that we have a mapping from R to the endomorphism ring of M , $\varphi: R \rightarrow \text{End}(M)$, and **M.2–4** just state that φ is a ring homomorphism. Thus a module is essentially an abelian group with operators, and the notions of submodule and generating set are defined as in that case and will not be repeated.

If in place of **M.3** we have $x(rs) = (xs)r$, we shall usually write $r.x$ instead of xr , so that our module now satisfies, instead of **M.3**,

$$\mathbf{M.3}^\circ \quad (rs).x = r.(s.x).$$

This is called a *left R-module*, in contrast to the right *R*-modules satisfying **M.3**. If we define the *opposite ring* of *R* as the ring R° whose additive group is the same as that of *R*, but with multiplication

$$x.y = yx,$$

then we can say: a left *R*-module is a right R° -module and a right *R*-module is a left R° -module. In other words, changing the side from which a ring operates on a module corresponds to a reversal of the order of multiplication in the ring. This means that instead of using an R° -module we can take the coefficients on the other side, so as to get an *R*-module. Of course when the ring *R* is commutative, then $R^\circ = R$; in this case the difference between left and right *R*-modules is purely notational. We sometimes write $_R M$ resp. M_R to indicate that *M* is a left resp. right *R*-module.

If *M* and *N* are right *R*-modules, an *R*-homomorphism $f: M \rightarrow N$, also called an *R-linear map*, is a map *f* satisfying

$$f(x + y) = fx + fy, \quad f(xr) = (fx)r, \quad x \in M, r \in R.$$

Here we have written *f* on the left, on the opposite side from the ring coefficients. We shall usually follow this practice; thus a homomorphism of left modules will be written on the right. This convention is chiefly of use for endomorphisms, which again form a ring, but it is not always possible to adhere to it rigidly (e.g. in the case of bimodules). We shall frequently use left modules, so as to be able to put maps on the right.

If *M*, *N* are left *R*-modules and *f*, *g* are homomorphisms from *M* to *N*, then so is the map *f* + *g* defined by

$$x(f + g) = xf + xg;$$

in this way the set $\text{Hom}_R(M, N)$ of all *R*-homomorphisms from *M* to *N* becomes an abelian group. When *R* is commutative, this is actually an *R*-module, taking *rf* to be defined by

$$x.(rf) = rx.f = r(xf).$$

For general rings this is no longer so; we shall meet the appropriate generalization in a moment. The set $\text{Hom}_R(M, M)$ of all endomorphisms of *M* is denoted by $\text{End}_R(M)$. We recall that it is a ring, using the composition of endomorphisms as multiplication.

Let *R*, *S* be rings and let *M* be an abelian group which is both a left *R*-module and a right *S*-module, such that

$$(rx)s = r(xs) \quad \text{for all } x \in M, r \in R, s \in S. \tag{1}$$

Then *M* is said to be an (R, S) -bimodule, and we indicate this by writing $_R M_S$. In

the case $S = R$ we speak of an R -bimodule. This is then a left and right R -module M such that $(ax)b = a(xb)$ for $x \in M$ and $a, b \in R$. For example, the ring R itself is an R -bimodule, as the associative law shows. Further, when R is commutative, any left R -module may be considered as right R -module and hence as R -bimodule.

We remark that any left R -module may be defined as (R, S) -bimodule by means of a homomorphism $f: S \rightarrow \text{End}_R(M)$. Likewise a right S -module may be defined as an (R, S) -bimodule by an antihomomorphism $g: R \rightarrow \text{End}_S(M)$; that we need an antihomomorphism is so because R acts on M from the left.

Given an (R, S) -bimodule M and an (R, T) -bimodule N , we can define $\text{Hom}_R(M, N)$ in a natural way as (S, T) -bimodule by the rules

$$x(sf) = (xs)f, \quad x(ft) = (xf)t, \quad \text{where } f: M \rightarrow N, x \in M, s \in S, t \in T.$$

The bimodule property follows because

$$x[(sf)t] = [x(sf)]t = ((xs)f)t, \quad x[s(ft)] = (xs)(ft) = ((xs)f)t.$$

This rule: $(_RM_S, {}_RN_T) \mapsto ({}_{\text{Hom}_R}(M, N)_T)$ is easily remembered if it is borne in mind that Hom is contravariant in the first and covariant in the second argument, so the order in S is reversed while that in T is preserved. In particular, when R is commutative, any R -modules may be regarded as R -bimodules and $\text{Hom}_R(M, N)$ is again an R -bimodule.

We recall from Vol. 1, **10.3** that for any homomorphism $f: M \rightarrow N$, its *kernel*, defined as $\ker f = \{x \in M \mid xf = 0\}$, is a submodule of M , its *image* $\text{im } f = \{y \in N \mid y = xf \text{ for some } x \in M\}$ is a submodule of N , and its *cokernel* and *coimage* are respectively $\text{coker } f = N/\text{im } f$, $\text{coim } f = M/\ker f$. Their relations are summarized in the following commutative diagram with exact row:

$$\begin{array}{ccccccc} 0 & \longrightarrow & \ker f & \longrightarrow & M & \xrightarrow{f} & N & \longrightarrow & \text{coker } f & \longrightarrow & 0 \\ & & \alpha \downarrow & & f' & & \beta \uparrow & & & & & \\ & & & & \text{coim } f & \longrightarrow & \text{im } f & & & & \end{array}$$

Here α is surjective, β is injective and f' is an isomorphism. The existence of the factorization $f = \alpha f' \beta$ (or $f = \beta f' \alpha$ in the case of right modules, when maps are on the left), is known as the first isomorphism theorem. We also recall the other isomorphism theorems (**10.2**, Vol. 1):

SECOND ISOMORPHISM THEOREM (parallelogram law) *For any submodules A, B of a module M we have the isomorphism*

$$(A + B)/B \cong A/(A \cap B).$$

THIRD ISOMORPHISM THEOREM *Given a module M and a submodule M' , there is a natural (order-preserving) bijection between the submodules of M containing M' and the submodules of M/M' , $N \leftrightarrow N/M'$, with an isomorphism of*

corresponding quotients:

$$(M/M')/(N/M') \cong M/N.$$

$$\begin{array}{ccccc} M & \xrightarrow{\text{nat}} & M/M' & \xrightarrow{\text{nat}} & (M/M')/(N/M') \\ & \searrow \text{nat} & \downarrow f & \nearrow f' & \\ & & M/N & & \end{array}$$

Here 'nat' denotes the natural homomorphism to the quotient module, and f' is the isomorphism whose existence is being asserted.

Let $f: R \rightarrow S$ be a homomorphism of rings. Given a left S -module M , we can define a left R -module ${}^f M$ by taking the additive group of M with the R -action defined by

$$a.x = (af)x, \quad x \in M, a \in R.$$

It is clear that we obtain an R -module in this way, said to be obtained from the S -module M by *pullback* along f . It is also clear that a subgroup of M is an R -submodule of ${}^f M$ iff it is an S -submodule of M .

Suppose now that $f: R \rightarrow S$ is a surjective homomorphism and U is a left R -module such that $aU = 0$ for all $a \in R$ such that $af = 0$. Then we can define an S -module structure on U by the rule

$$c.x = ax, \quad \text{where } x \in U, c \in S, \quad (2)$$

and $a \in R$ is such that $af = c$. If also $bf = c$, then $(a - b)f = 0$, hence $(a - b)x = 0$ and so $ax = bx$; this shows that the definition (2) is unambiguous. It is easily checked that this defines an S -module structure on U ; we shall not use a special symbol for it, but note that the original R -module structure on U can be obtained by pullback along f .

For any ring R we have a category \mathcal{M}_R whose objects are the right R -modules while the morphisms are the R -linear maps. Similarly we can form the category ${}_R\mathcal{M}$ of left R -modules and the category ${}_S\mathcal{M}_R$ of (S, R) -bimodules.

Given a left R -module X and any sequence of left R -modules and homomorphisms $Y' \xrightarrow{\alpha} Y \xrightarrow{\beta} Y''$, we have for each homomorphism $f: X \rightarrow Y'$ a commutative diagram

$$\begin{array}{ccccc} & & X & & \\ & \swarrow f & \downarrow f_\alpha & \searrow f\alpha\beta & \\ Y' & \xrightarrow{a} & Y & \xrightarrow{\beta} & Y'' \end{array}$$

and hence a sequence of homomorphisms of abelian groups

$$\text{Hom}_R(X, Y') \xrightarrow{\alpha_*} \text{Hom}_R(X, Y) \xrightarrow{\beta_*} \text{Hom}_R(X, Y''),$$

where α_* maps f to $f\alpha$ and β_* maps g to $g\beta$. Similarly $(\alpha\beta)_*$ maps f to $f(\alpha\beta)$ $= (f\alpha)\beta$, so that $(\alpha\beta)_* = \alpha_*\beta_*$. If $Y' = Y$ and $\alpha = 1_Y$ then clearly $1_* = 1$ and this shows $\text{Hom}_R(X, -)$: $Y \mapsto \text{Hom}_R(X, Y)$ to be a covariant functor from $R\mathcal{M}$ to the category Ab of abelian groups.

If in $\text{Hom}_R(X, Y)$ we keep Y fixed and vary X we obtain a contravariant functor in the same way. The contravariance is expressed by the reversal of the arrows in passing from $X' \xrightarrow{\alpha} X \xrightarrow{\beta} X''$ to

$$\text{Hom}_R(X'', Y) \xrightarrow{\beta^*} \text{Hom}_R(X, Y) \xrightarrow{\alpha^*} \text{Hom}_R(X', Y).$$

$$\begin{array}{ccccc} & & & & \\ & & & & \\ X' & \xrightarrow{\alpha} & X & \xrightarrow{\beta} & X'' \\ & \searrow \alpha\beta g & \downarrow \beta g & \nearrow g & \\ & & Y & & \end{array}$$

Thus $\text{Hom}_R(-, -)$ is a functor of two arguments, also called a *bifunctor*.

We recall from 10.3 of Vol. 1 that for any family $(M_i)_{i \in I}$ of left R -modules, their *direct product* $P = \prod_I M_i$ is defined as the Cartesian product of the M_i , consisting of all families (x_i) ($x_i \in M_i$), in which the operations are defined componentwise: $(x_i) + (y_i) = (x_i + y_i)$, $r(x_i) = (rx_i)$ (or for right modules, $(x_i)r = (x_ir)$). If $M_i = M$ for all i , we have a *direct power* of M , written M^I . The subset S of P consisting of all families (x_i) with at most finitely many non-zero terms is again a module, called the *direct sum* of the M_i and written $S = \coprod_I M_i$. For each $i \in I$ we have the canonical surjection $\pi_i: P \rightarrow M_i$ which picks out the i th coordinate, and the canonical injection $\mu_i: M_i \rightarrow S$ which maps $x \in M_i$ to the family (x_j) , where $x_j = \delta_{ij}x$. When $M_i = M$ for all $i \in I$, we have a direct sum of copies of M , written $'M$. For a finite index set I the direct sum and direct product coincide as modules; nevertheless it is often convenient to distinguish between them; their significance will become clearer with the homological development of module theory in Vol. 3.

A module M is called the *sum* of a family of submodules (M_i) , $M = \sum_I M_i$, if every element of M can be written as $\sum x_i$ ($x_i \in M_i$), where almost all x_i vanish, so that the sum is well-defined. If the sum is unique, we write $M = \bigoplus_I M_i$ and call it the *direct sum*. This is sometimes called the *internal* direct sum, to distinguish it from the *external* direct sum or *coproduct* $\coprod_I M_i$. It is clear that $\coprod_I M_i$ is in fact the internal direct sum of the submodules $\text{im } \mu_i$, where the μ_i are the canonical injections.

Finally we recall the universal properties of $P = \prod M_i$ and $S = \coprod M_i$.

(i) Given a family M_i ($i \in I$) of R -modules, there exists an R -module P with homomorphisms $\pi_i: P \rightarrow M_i$ such that for any family of homomorphisms $f_i: A \rightarrow M_i$ from an R -module A there is a unique homomorphism $f: A \rightarrow P$ such that $f_i = f\pi_i$ for all $i \in I$.

(ii) Given a family M_i ($i \in I$) of R -modules, there exists an R -module S with homomorphisms $\mu_i: M_i \rightarrow S$ such that for any family of homomorphisms $g_i: M_i \rightarrow B$ to an R -module B there exists a unique homomorphism $g: S \rightarrow B$ such that $g_i = \mu_i g$ for all $i \in I$.

$$\begin{array}{ccc} & \pi_i & \\ P & \xrightarrow{\quad} & M_i \\ f \uparrow & \nearrow f_i & \\ A & & \end{array} \qquad \begin{array}{ccc} & \mu_i & \\ M_i & \xrightarrow{\quad} & S \\ & \searrow g_i & \downarrow g \\ & & B \end{array}$$

We remark that these universal properties can be expressed by the equations
 $\text{Hom}_R(A, \prod M_i) \cong \prod \text{Hom}_R(A, M_i), \quad \text{Hom}_R(\coprod M_i, B) \cong \prod \text{Hom}_R(M_i, B).$

Exercises

- (1) Verify that $\text{End}_R(M)$ is a ring.
- (2) Let M, N be left R -modules and $H = \text{Hom}(M, N)$ the abelian group of all homomorphisms (not necessarily R -linear) from M to N , *qua* abelian groups. Verify that $\text{Hom}_R(M, N)$ is a subgroup of H . Taking $N = M$, show that the R -module action on M defines a ring homomorphism $\varphi: R \rightarrow \text{End}(M)$ and show that $\text{End}_R(M)$ is the centralizer in $\text{End}(M)$ of the image of φ .
- (3) Let M, N be right R -modules and $H = \text{Hom}(M, N)$ as in Ex. (2). Show that H can be regarded (i) as a right R -module, by defining $x(fr) = (xf)r$, (ii) as a left R -module, by defining $x(rf) = (xr)f$, where $x \in M, r \in R, f \in H$. Verify that H is an R -bimodule, and show that this argument breaks down if we replace $\text{Hom}(M, N)$ by $\text{Hom}_R(M, N)$.
- (4) Let R be any ring. Show that $R^\circ \cong R$ precisely when R has an antiautomorphism. Show also that $R^{\circ\circ} = R$; when is $R^\circ = R$?
- (5) Let A be an abelian group (written additively) and n a positive integer such that $nx = 0$ for all $x \in A$. Show how to define A as a (\mathbb{Z}/n) -module in a natural way.
- (6) Let M be a left R -module. Show that for each $u \in M$, the annihilator of u in R , $\text{Ann}(u) = \{x \in R \mid xu = 0\}$ is a left ideal and that $\text{Ann } M = \{x \in R \mid xM = 0\}$ is an ideal in R .
- (7) Let R be a ring and \mathfrak{a} an ideal in R . Show that every (R/\mathfrak{a}) -module can be defined as an R -module by pullback along the natural homomorphism $R \rightarrow R/\mathfrak{a}$, and conversely, every right R -module M such that $M\mathfrak{a} = 0$ can be defined as an (R/\mathfrak{a}) -module.

- (8) Let R be a ring, \mathfrak{a} an ideal and $\mathcal{M}_{(R,\mathfrak{a})}$ the category of right R -modules M such that $M\mathfrak{a} = 0$. Show that $\mathcal{M}_{(R,\mathfrak{a})}$ is a full subcategory of \mathcal{M}_R which is equivalent to $\mathcal{M}_{R/\mathfrak{a}}$.

4.2 Semisimple modules

Let R be any ring. An R -module M is called *simple* if $M \neq 0$ and M has no submodules other than 0 and M . For example, over a field the only simple module is a one-dimensional vector space. When $R = \mathbf{Z}$, the \mathbf{Z} -modules are just abelian groups, and the simple \mathbf{Z} -modules are the cyclic groups of prime order \mathbf{C}_p .

A module which can be expressed as a direct sum of simple modules is said to be *semisimple*. For example, \mathbf{C}_6 is semisimple since it can be written $\mathbf{C}_6 = \mathbf{C}_2 \oplus \mathbf{C}_3$, but \mathbf{C}_4 is not semisimple. To show that a module M is semisimple we need to find a family of simple submodules S_λ such that the sum $\sum S_\lambda$ is direct. Here it is not necessary for the family to be finite; in fact we have the following condition for the sum to be finite.

PROPOSITION 2.1 *Let M be a semisimple module, say*

$$M = \bigoplus_I S_i, \quad (1)$$

where each S_i is simple. Then the number of summands in (1) is finite if and only if M is finitely generated.

Proof. If the index set I is finite, then since each S_i , being simple, is cyclic, M is finitely generated. Conversely, let M be finitely generated, by u_1, \dots, u_r , say. For each u_j we can find finitely many terms S_i whose sum contains u_j , hence all the u_j are contained in the sum of a finite subfamily of the S_i , and this family generates M , so that I must be finite. ■

Below we shall describe semisimple modules in some alternative ways which are often used, but first we single out a property needed in the proof.

LEMMA 2.2 *Let M be a module of the form $M = \sum_I S_i$, where S_i is simple. If N is any submodule of M , then there is a subset J of I such that*

$$M = N \oplus \left(\bigoplus_{i \in J} S_i \right).$$

In particular, $M = \bigoplus_L S_i$ for some $L \subseteq I$.

Proof. Let \mathcal{F} be the family of all subsets J of I such that the sum $N + \sum_J S_i$ is direct. The sum fails to be direct iff there is a finite subset J_0 of J and $x_j \in S_i$ ($j \in J_0$), $u \in N$, such that $x_j \neq 0$ and $u + \sum_{J_0} x_j = 0$. This shows \mathcal{F} to be of finite character and hence inductive. By Zorn's lemma there is a maximal subset J in \mathcal{F} ;

with this set J the sum $P = N + \sum_j S_j$ is direct. For any S_k , $P \cap S_k$ is either S_k or 0; if it is 0, then the sum $N + \sum_j S_j + S_k$ is still direct, contradicting the maximality of J . Hence $P \cap S_k = S_k$ for all k , so $P \supseteq \sum S_k = M$, and hence $P = M$, as claimed. The final assertion is the case $N = 0$. ■

Let us recall that a submodule N of a module M is said to be *complemented* in M if there is a submodule N' such that $M = N \oplus N'$. Here N' is called a *complement* of N in M ; generally it will not be unique, but it is unique up to isomorphism, because $N' \cong M/N$, by the second isomorphism theorem. We remark that when $M = N \oplus N'$, then for any submodule P containing N we have $P = M \cap P = (N + N') \cap P = N + (N' \cap P)$, by the modular law, and $N \cap (N' \cap P) = 0$, hence $P = N \oplus (N' \cap P)$; this shows that N is also complemented in P .

THEOREM 2.3 *Let R be any ring. For any R -module M the following are equivalent:*

- (a) *M is semisimple,*
- (b) *M is a sum of simple modules,*
- (c) *every submodule of M is complemented.*

Proof. (a) \Rightarrow (b) is clear. To prove (b) \Rightarrow (c), let $M = \sum_i S_i$, where each S_i is simple. For any submodule N of M we have, by Lemma 2.2, a subset J of I such that $M = N \oplus (\bigoplus_j S_j)$; hence $\bigoplus_j S_j$ is a complement of N in M .

(c) \Rightarrow (a). Let S be the sum of all simple modules in M ; we first show that $M = S$. For if not, then $M = S \oplus T$, where $T \neq 0$. Let N be a non-zero cyclic submodule of T ; by Prop. 2.2.10 there exists a maximal proper submodule N' of N : N' has a complement in M , and by the remark preceding the theorem, N' also has a complement in N , say $N = N' \oplus P$. Since $P \cong N/N'$, P is simple and it is contained in T ; thus $P \subseteq S \cap T$, which is a contradiction. This shows that $T = 0$ and so $S = M$. If we now apply Lemma 2.2 with $N = 0$ to the sum S of simple modules, we obtain M as a direct sum of simple modules. ■

COROLLARY 2.4 *If $M = \bigoplus_i S_i$, where S_i is simple, then any submodule N of M has a complement of the form $\bigoplus_{K'} S_i$ for some subset K' of I , and*

$$N \cong \bigoplus_{K'} S_i, \quad \text{where } K' = I \setminus K. \tag{2}$$

In particular, if N is simple, then $N \cong S_i$ for some $i \in I$.

Proof. By Lemma 2.2 we have $M = N \oplus (\bigoplus_K S_i)$ for some $K \subseteq I$; thus $N' = \bigoplus_{K'} S_i$ is a complement of N . Since we also have $M = (\bigoplus_{K'} S_i) \oplus (\bigoplus_{I \setminus K} S_i)$, it follows that $N \cong M/N' \cong \bigoplus_{K'} S_i$ and (2) follows. When N is simple, K' reduces to a single element and so $N \cong S_i$ in this case. ■

As an example that the isomorphism (2) cannot always be strengthened to equality, take a simple module P and in P^2 define the submodules $P_1 = \{(x, 0) | x \in P\}$, $P_2 = \{(0, x) | x \in P\}$, $P_3 = \{(x, x) | x \in P\}$. Clearly $P^2 = P_1 \oplus P_2 = P_1 \oplus P_3 = P_2 \oplus P_3$, and $P_1 \cong P_2 \cong P_3$, but these submodules are all different. This example also illustrates the point that to prove the directness of a sum $\sum P_i$ of more than two terms it is not enough to have $P_i \cap P_j = 0$ for $i \neq j$, we must have $P_i \cap (\sum_{j \neq i} P_j) = 0$, or at least $P_i \cap (\sum_{j=1}^{i-1} P_j) = 0$.

The form of (2), together with the fact that every quotient of M is isomorphic to a submodule, shows the truth of

COROLLARY 2.5 *If M is semisimple, then so is every submodule and every quotient of M .* ■

We observe that all of (a)–(c) in Th. 2.3 are lattice conditions, but they are not equivalent in all lattices. Thus (a) \Leftrightarrow (b) in any lattice and (b) \Leftrightarrow (c) in any modular lattice of finite length, but in general neither of (b), (c) implies the other. Thus (c) but not (b) is self-dual, and in the lattice of ideals of \mathbb{Z} the dual of (b) but not of (c) holds. To find an example satisfying (c) but not (b), let B be the Boolean algebra of all subsets of an infinite set S , and \mathcal{F} the ideal of all finite subsets. Then B/\mathcal{F} is again a Boolean algebra, hence a complemented distributive lattice, but it has no atoms, so (c) holds, but not (b).

Condition (c) of Th. 2.3 may also be expressed by saying that any exact sequence $0 \rightarrow N \rightarrow M \rightarrow M/N \rightarrow 0$ splits; in other words, every short exact sequence with middle term M splits. We recall that the short exact sequence

$$0 \rightarrow M' \xrightarrow{f} M \xrightarrow{g} M'' \rightarrow 0$$

is said to *split*, or be *split exact*, if $\text{im } f = \ker g$, and this module $\text{im } f$ is complemented in M . This is equivalent to either of the following conditions:

(a) There is an R -homomorphism $f': M \rightarrow M'$ such that $ff' = 1_M$; this mapping f' is called a *right inverse* or *retraction* for f .

(b) There is an R -homomorphism $g': M'' \rightarrow M$ such that $g'g = 1_{M''}$; the mapping g' is called a *left inverse* or a *section* for g .

Clearly (a), (b) hold when $\text{im } f$ is complemented; conversely, in case (a) we have $M = \text{im } f \oplus \ker f'$ and in case (b) $M = \ker g \oplus \text{im } g'$, as is easily verified.

Over a field (even skew) the simple modules are just the one-dimensional vector spaces; since every vector space can be written as a sum of one-dimensional spaces, it follows that over a field every module (= vector space) is semisimple. In particular, this proves the existence of a basis for any vector space, even infinite-dimensional. For if $V = \bigoplus_{i \in I} S_i$ and u_i is a generator of the simple module S_i , then $\{u_i\}_{i \in I}$ is a basis of V , as is easily checked.

For an arbitrary ring R the theory of semisimple modules is quite similar to the theory of vector spaces over a field. The main difference is that there may be more

than one type of simple module. We shall say that two simple R -modules have the same *type* if they are isomorphic. A semisimple R -module is said to be *isotypic* if it can be written as a sum of simple modules all of the same type. In any R -module M the sum of all simple submodules is called the *socle* of M ; thus the semisimple modules are those that coincide with their socle. The sum of all simple submodules of a given isomorphism type α is called the α -socle of M or also a *type component* in M .

Let M be any R -module; a submodule N of M is said to be *fully invariant* in M if it admits (i.e. is mapped into itself by) all R -endomorphisms of M . We note that for an R -module M , the set $S = \text{End}_R(M)$ is just the centralizer in $\text{End}(M)$ of the image of R defining the R -action on M , so a subgroup N of M is a fully invariant submodule iff it admits R and S . In a semisimple module the fully invariant submodules are easily described: they are sums of type components.

THEOREM 2.6 *Let R be a ring and M an R -module. Then*

- (i) *for any type α the α -socle is an isotypic submodule of M containing all simple submodules of type α , and the socle is the direct sum of the α -socles, for different α ;*
- (ii) *any sum of type components in M is fully invariant in M ; when M is semisimple, the converse holds: a fully invariant submodule is a sum of type components;*
- (iii) *if M is semisimple and S is the centralizer of the R -action, then M is also semisimple as S -module.*

Proof. (i) Denote the α -socle of M by H_α . By Th. 2.3, H_α is semisimple, and it contains all simple submodules of type α , by definition. By Cor. 2.4, any submodule is a sum of modules of type α , hence $H_\alpha \cap (\sum_{\beta \neq \alpha} H_\beta) = 0$, so the sum $\sum H_\alpha$ is direct.

(ii) Let H_α again be the α -socle of M , say $H_\alpha = \sum P_i$, where the P_i are all the simple modules of type α . For any R -endomorphism f of M we have $H_\alpha f = \sum P_i f$; each $P_i f$ is a homomorphic image of P_i , hence either 0 or simple of type α , so in any case $P_i f \subseteq H_\alpha$. Hence $H_\alpha f \subseteq H_\alpha$, i.e. H_α is fully invariant. It follows that any sum of type components is fully invariant.

Now let N be a fully invariant submodule of a semisimple module M . To show that N is a sum of type components of M we need only show: if P is a simple submodule of N and Q a simple submodule of M such that $Q \cong P$, then $Q \subseteq N$. If $Q = P$, this is clear; otherwise $P \cap Q = 0$, because the intersection must be a proper submodule of both; hence $P + Q$ is direct and $M = P \oplus Q \oplus U$, for some U , by Th. 2.3(c). Let $f: P \rightarrow Q$ be the given isomorphism and define a mapping $\theta: M \rightarrow M$ by

$$\theta: x + y + z \mapsto xf + yf^{-1} + z, \quad \text{where } x \in P, y \in Q, z \in U.$$

Clearly θ is an R -endomorphism, hence $N\theta \subseteq N$, and $P \subseteq N$, therefore $Q = P\theta \subseteq N$. Thus N contains with P the whole α -socle of M and so is a sum of type components of M .

(iii) Let us write M as left R -, right S -module. We have $M = \sum xS$, where x runs over all elements of all simple R -submodules, for every element of M is a sum of such elements. Now the result will follow if we show that xS is simple or 0. Let $y \in xS$, $y \neq 0$, say $y = xs$. Then for any $a \in R$, the mapping $ax \mapsto axs = ay$ is a homomorphism $Rx \rightarrow Ry$, which is surjective and $Ry \neq 0$. Hence it is an isomorphism, so there is also an R -homomorphism which maps y to x . As in (ii) we can find an R -endomorphism of M which maps y to x . Hence yS contains x , therefore $yS = xS$ and so xS is indeed simple. ■

COROLLARY 2.7 *Let M be a semisimple R -module and express M as the sum of its α -socles: $H = \bigoplus_{\alpha} H_{\alpha}$. Then*

$$\text{End}_R(M) = \prod_{\alpha} \text{End}_R(H_{\alpha}). \quad (3)$$

Proof. Any family (f_{α}) , where f_{α} is an endomorphism of H_{α} , defines an endomorphism f of M and it is clear that the correspondence $(f_{\alpha}) \mapsto f$ is an injective homomorphism from the right- to the left-hand side in (3). In the other direction consider $f \in \text{End}_R(M)$; since each H_{α} is fully invariant, it is mapped to itself by f . If the restriction of f to H_{α} is denoted by f_{α} then f is just the endomorphism defined by the family (f_{α}) . Thus the homomorphism found is surjective and (3) is established. ■

In any semisimple module M the simple components in a direct decomposition are determined up to isomorphism; on the other hand the decomposition $M = \sum H_{\alpha}$ into α -socles is unique (not merely up to isomorphism).

PROPOSITION 2.8 *Let M be a finitely generated semisimple R -module. Then any two direct decompositions of M into simple modules have the same number of terms, and for suitable numbering corresponding terms are isomorphic.*

Proof. Let $M = P_1 \oplus \cdots \oplus P_n = Q_1 \oplus \cdots \oplus Q_m$ be two direct decompositions into simple submodules, where the number of terms is finite, by Prop. 2.1. By Cor. 2.4, $Q_1 \cong P_j$ for some j , say $j = 1$, and moreover, $P_2 \oplus \cdots \oplus P_n \simeq Q_2 \oplus \cdots \oplus Q_m$. Now the result follows by induction on n . ■

Later, in 4.4 below, we shall see that this result holds quite generally, without assuming M to be finitely generated.

Exercises

- (1) Show in detail how Th. 2.3 may be used to prove the existence of a basis in any vector space over a field.
- (2) Show that the socle of any module is semisimple.

- (3) Let R be any ring. Show that when R is considered as right R -module, the fully invariant submodules are just the two-sided ideals.
- (4) For any homomorphism between R -modules, $f:M \rightarrow N$, define its *graph* as the subset of $M \coprod N$ given by $F = \{(x, xf) | x \in M\}$. Verify that F is a submodule and that $F \cap N = 0$; when is $F \cap M = 0$?
- (5) A module M is said to be *distributive* if the lattice of its submodules, $\text{Lat}(M)$, is distributive. Show that a semisimple module is distributive iff any two distinct simple submodules are non-isomorphic. (*Hint.* Examine $\text{Lat}(P^2)$, where P is simple.)
- (6) (J.-E. Roos) By examining the graph of a homomorphism $M \rightarrow N$ (cf. Ex. (4)), show that if $M \coprod N$ is distributive, then $\text{Hom}_R(M, N) = 0$.
- (7) Show that a module M fails to be distributive iff M has distinct submodules U, V such that $U/(U \cap V) \cong V/(U \cap V)$. (*Hint.* For the necessity take a submodule with two relative complements.)
- (8) Show that if M is distributive and $M = U_1 \oplus \cdots \oplus U_r$, then $\text{End}_R(M) = \prod \text{End}_R(U_i)$.

4.3 Matrix rings

We first encounter matrices in the description of linear mappings between vector spaces; in particular, any matrix ring over a field may be described as an endomorphism ring of a vector space. In order to gain a better understanding of general matrix rings we shall consider direct sums of modules and their endomorphisms. We shall find that an endomorphism of a direct sum can be written as a matrix ring whose (i,j) -entry is a homomorphism from the i th to the j th summand.

Let us consider, for any ring R , a left R -module M which is expressed as a direct sum of certain submodules

$$M = U_1 \oplus \cdots \oplus U_n. \quad (1)$$

Let $\pi_i: M \rightarrow U_i$ be the canonical projections and $\mu_i: U_i \rightarrow M$ the canonical injections ($i = 1, \dots, n$); thus $(x_1, \dots, x_n)\pi_i = x_i$, $x\mu_i = (0, \dots, 0, x, 0, \dots, 0)$, with x in the i th place. It is clear that we have

$$\mu_i \pi_j = \delta_{ij}, \quad (2)$$

$$\sum \pi_i \mu_i = 1. \quad (3)$$

With each endomorphism $f: M \rightarrow M$ we can associate the matrix (f_{ij}) , where $f_{ij}: U_i \rightarrow U_j$ is defined by $f_{ij} = \mu_i f \pi_j$. Similarly, any family (α_{ij}) of homomorphisms $\alpha_{ij}: U_i \rightarrow U_j$ gives rise to an endomorphism $\alpha: M \rightarrow M$ defined by $\alpha = \sum \pi_i \alpha_{ij} \mu_j$. These two processes are mutually inverse: If $\alpha = \sum \pi_i \alpha_{ij} \mu_j$, then $\mu_r \alpha \pi_s = \sum \mu_r \pi_i \alpha_{ij} \mu_j = \alpha_{rs}$ by (2), and if $f_{ij} = \mu_i f \pi_j$, then $\sum \pi_i f_{ij} \mu_j = \sum \pi_i \mu_i f \pi_j \mu_j = f$

by (3). Moreover, the families (f_{ij}) are added and multiplied ‘matrix fashion’: $(f+g)_{ij} = \mu_i(f+g)\pi_j = \mu_i f\pi_j + \mu_i g\pi_j = f_{ij} + g_{ij}$ and $(fg)_{ik} = \mu_i f g \pi_k = \sum \mu_i f \pi_j \mu_j g \pi_k = \sum f_{ij} g_{jk}$. Thus we obtain

THEOREM 3.1 *Let R be any ring. If M is a left R -module, expressed as a direct sum: $M = U_1 \oplus \cdots \oplus U_n$, with projections $\pi_i: M \rightarrow U_i$ and injections $\mu_i: U_i \rightarrow M$, then the elements of $\text{End}_R(M)$ can be written as matrices (f_{ij}) , where $f_{ij}: U_i \rightarrow U_j$, with the usual addition and multiplication of matrices.* ■

In particular, when all the summands are isomorphic, then $M \cong U^n$, and we obtain

COROLLARY 3.2 *Let R be a ring, U a left R -module and $S = \text{End}_R(U)$. Then for any $n \geq 1$ we have*

$$\text{End}_R(U^n) \cong \mathfrak{M}_n(S). \quad \blacksquare \quad (4)$$

Here U^n means the direct sum of n copies of U . It is most convenient to write its elements as rows; then it is clear how the matrix ring $\mathfrak{M}_n(S)$ operates from the right on these rows. Generally, if V is an (R, S) -bimodule, R acting on the left and S on the right, we write “ V ” for the set of all $m \times n$ matrices with entries in V . This is acted on from the left by $\mathfrak{M}_m(R)$ and from the right by $\mathfrak{M}_n(S)$, and it is easily verified to be an $(\mathfrak{M}_m(R), \mathfrak{M}_n(S))$ -bimodule. In the special case $m = 1$ we write ${}^1V^n$ simply as V^n , as remarked earlier. Similarly in the case $n = 1$ we write “ V ” for the set of column vectors. The same notation “ R ” is used in the case of a ring R ; when $m = n$, we have $\mathfrak{M}_n(R)$, which is also denoted by R_n . For example, R is itself an R -bimodule, as we saw, and “ R ” becomes an (R_m, R_n) -bimodule in this way.

Let us return to the case considered in Cor. 3.2. Writing $e_{ij} = \pi_i \mu_j$, we obtain from (2), (3) the equations

$$e_{ij} e_{kl} = \delta_{jk} e_{il}, \quad \sum e_{ii} = 1. \quad (5)$$

It follows that the e_{ij} are just the familiar matrix units in $(\text{End}_R(U))_n$; thus e_{ij} is the matrix with 1 in the (i, j) -position and 0 elsewhere. The isomorphism (4) is given explicitly by

$$\alpha \leftrightarrow (\alpha_{ij}), \quad \text{where } \alpha_{ij} = \mu_i \alpha \pi_j, \quad \alpha = \sum \pi_i \alpha_{ij} \mu_j.$$

It is worth remarking that matrix rings can be defined in terms of their matrix units. To state the result we recall that in any ring R the *centralizer* of a subset X is the set $\{r \in R \mid rx = xr \text{ for all } x \in X\}$. This is easily seen to be a subring of R . In particular the centralizer of R itself is just its *centre*.

PROPOSITION 3.3 *Let R be any ring with n^2 elements e_{ij} satisfying (5). Then $R \cong C_n$, where C is the centralizer in R of the e_{ij} .*

Proof. Given $a \in R$, we define $a_{ij} = \sum_v e_{vi}ae_{jv}$. Then since $e_{jv}e_{kl} = \delta_{jk}e_{jl}$, we have $a_{ij}e_{kl} = \sum_v e_{vi}ae_{jv}e_{kl} = e_{ki}ae_{jl}$ and $e_{kl}a_{ij} = \sum_v e_{kl}e_{vi}ae_{jv} = e_{ki}ae_{jl}$, hence $a_{ij} \in C$. Moreover, $\sum a_{ij}e_{ij} = \sum v e_{vi}ae_{jv}e_{ij} = \sum i e_{ii}ae_{jj} = a$. Conversely, given a matrix (a_{ij}) over C , we define $a = \sum a_{ij}e_{ij}$ and then find that $\sum e_{vi}ae_{jv} = \sum e_{vi}a_{rs}e_{rs}e_{jv} = \sum a_{rs}e_{vi}e_{rs}e_{jv} = a_{ij}\sum e_{vv} = a_{ij}$. Thus we have indeed a bijection $a \leftrightarrow (a_{ij})$ and this is easily seen to be an isomorphism. ■

It is clear that any ring homomorphism $f: K \rightarrow L$ induces a homomorphism of matrix rings $K_n \rightarrow L_n$; we simply map (a_{ij}) to $(a_{ij}f)$. We also have a converse:

COROLLARY 3.4 *Let $f: R \rightarrow S$ be a ring homomorphism. If R is a full matrix ring, say $R = K_n$, then S has the form $S = L_n$ for some ring L , and there is a homomorphism $\varphi: K \rightarrow L$, which induces f .*

Note particularly that f does not have to be surjective, for the conclusion to hold.

Proof. Let e_{ij} be the matrix units in R and write $e'_{ij} = e_{ij}f$. The e'_{ij} again satisfy the equations (5), hence S is a matrix ring; moreover, the centralizer of the e_{ij} , K say, maps to the centralizer of the e'_{ij} in S , L say, and this mapping $\varphi: K \rightarrow L$ induces f ; the details may be left to the reader. ■

Let A, B be rings and M an (A, B) -module. Then we can form a ring which is most easily expressed in matrix form

$$R = \begin{pmatrix} A & M \\ 0 & B \end{pmatrix}.$$

Thus the elements of R are 2×2 matrices $\begin{pmatrix} a & m \\ 0 & b \end{pmatrix}$, $a \in A, b \in B, m \in M$, with the usual matrix operations:

$$\begin{aligned} \begin{pmatrix} a & m \\ 0 & b \end{pmatrix} + \begin{pmatrix} a' & m' \\ 0 & b' \end{pmatrix} &= \begin{pmatrix} a+a' & m+m' \\ 0 & b+b' \end{pmatrix}, \\ \begin{pmatrix} a & m \\ 0 & b \end{pmatrix} \begin{pmatrix} a' & m' \\ 0 & b' \end{pmatrix} &= \begin{pmatrix} aa' & am'+mb' \\ 0 & bb' \end{pmatrix}. \end{aligned}$$

The ring structures on A and B and the module structure on M ensure that R is again a ring. It will be called the *triangular matrix ring* formed from A, B and M . This construction is often useful in forming asymmetric counter-examples. We note the ideal structure of R :

PROPOSITION 3.5 *Let A, B be rings, M an (A, B) -bimodule and let R be the triangular matrix ring formed from A, B, M . Given any left A -submodule N of*

$A \oplus M$ and any left ideal b of B satisfying $Mb \subseteq N$, the set $N \oplus b$ is a left ideal of R and every left ideal of R is of this form.

Proof. Let I be a left ideal of R and consider a typical element $\begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$. Then I contains the products

$$\begin{aligned} \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & y \\ 0 & z \end{pmatrix} &= \begin{pmatrix} ax & ay \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & m \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & y \\ 0 & z \end{pmatrix} = \begin{pmatrix} 0 & mz \\ 0 & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x & y \\ 0 & z \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & bz \end{pmatrix}. \end{aligned} \tag{6}$$

Taking $a = 1$, $b = 1$, we see that I contains each row of $\begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$; thus we have $I = N \oplus b$, where N denotes the set of first rows and b the set of second rows. The equations (6) further show that $AN \subseteq N$, $Bb \subseteq b$ and $Mb \subseteq N$; hence I is of the required form. Conversely, it is clear from (6), that any set of this form is a left ideal. ■

By the symmetry of the construction a similar result holds for right ideals.

We now return to the situation of Cor. 3.2. There we considered U^n as left R -module and found an endomorphism ring of the form S_n . But we can also consider U^n as left R_n -module, and then its endomorphism ring turns out to be S . In this case we shall write " U " and visualize the elements as column vectors; here we shall moreover find a correspondence between submodules. If M is an R -module, we shall write $\text{Lat}_R(M)$ for the lattice of all R -submodules of M .

THEOREM 3.6 *Let M be a left R -module and write $S = \text{End}_R(M)$. Then " M may be regarded as a left R_n -module in a natural way, and*

$$\text{End}_{R_n}(^nM) \cong S. \tag{7}$$

Moreover, there is a lattice-isomorphism

$$\text{Lat}_R(M) \cong \text{Lat}_{R_n}(^nM), \tag{8}$$

and (R, S) -subbimodules correspond to (R_n, S) -subbimodules under this isomorphism.

Proof. In determining the endomorphism ring of $_RM$ we may replace R by its image in $\text{End}(M)$; thus S is the centralizer of R in $\text{End}(M)$. Let e_{ij} be the matrix units in $\text{End}(^nM)$; together with R they generate a ring isomorphic to R_n in $\text{End}(^nM)$, and $\text{End}_{R_n}(^nM)$ is the centralizer of R and the e_{ij} . By Cor. 3.2, the centralizer of R is just S_n , so the required endomorphism ring is the centralizer of the e_{ij} in S_n , i.e. S itself; this proves (7).

Now any R -submodule N of M corresponds to an R_n -submodule $"N$ of $"M$, and the correspondence

$$N \mapsto "N \quad (9)$$

is clearly order-preserving. To establish (8) we need only show that (9) has an inverse which is also order-preserving. Consider the projection on the first factor

$$\pi_1: "M \mapsto M. \quad (10)$$

With any R_n -submodule of $"M$ this associates an R -submodule of M , and we claim that this is the required inverse. If we take an R -submodule N of M and apply first (9) and then (10), we evidently get back to N . Now let P be any R_n -submodule of $"M$; each projection π_i maps P to an R -submodule N_i of M , but all these submodules agree: $N_i = N_1$, because P admits R_n . This becomes clear if we think of the elements of P as column vectors, acted on by R_n from the left. It follows that $P = "N_1$, and since the S -action on M and on $"M$ is unaffected by R , it follows that S -submodules correspond under (8); therefore (R, S) -bimodules correspond to (R_n, S) -bimodules. ■

We can go beyond Th. 3.6 and obtain an equivalence between the categories of left R -modules and left R_n -modules. For any $M \in {}_{R_n}\mathcal{M}$ consider the functors

$$M \mapsto "M = \text{Hom}_R(R^n, M), \quad (11)$$

and for $P \in {}_{R_n}\mathcal{M}$,

$$P \mapsto P^\sharp = \text{Hom}_{R_n}("R, P) = e_{11}P. \quad (12)$$

The functorial property is clear and we have $("M)^\sharp \cong M$ and $"(P^\sharp) \cong P$. This shows that the functors (11) and (12) provide an equivalence between the categories ${}_R\mathcal{M}$ and ${}_{R_n}\mathcal{M}$, for any ring R and any $n \geq 1$.

Two rings R, S with the property that ${}_R\mathcal{M}$ and ${}_S\mathcal{M}$ are equivalent categories are said to be *Morita-equivalent*; it can be shown that this happens iff \mathcal{M}_R is equivalent to \mathcal{M}_S . What we have shown above can be expressed by saying that any ring R is Morita-equivalent to the full matrix ring R_n , for all n . In general R may be Morita-equivalent to rings not of the form R_n (e.g. when R is itself a matrix ring), but there is an important case where this is the only possibility; this will be discussed in 5.1. A general discussion of Morita-equivalence will be reserved for Vol. 3.

Exercises

- (1) Verify that in a full matrix ring R_n the centralizer of the matrix units e_{ij} is precisely R . Show also that $e_{11}R_ne_{11} \cong R$.
- (2) Let R be a ring with two elements u, v such that $u^2 = v^2 = 0$, $uv + vu = 1$. Show that $R = S_2$, where S is the centralizer of u, v .

- (3) Show that there is a natural bijection between the two-sided ideals of R and (R, R_n) -subbimodules of R^n , and likewise between two-sided ideals of R_n and (R, R_n) -subbimodules of R^n . Deduce that there is a natural bijection between the ideals of R and those of R_n .
- (4) Find the centre of the triangular matrix ring $\mathfrak{T}_n(\mathbf{R})$.
- (5) Show that if K is a skew field, then $\mathfrak{M}_n(K)$ for any $n \geq 1$ is a simple ring (cf. 5.3).
- (6) Let K be a skew field and $n \geq 1$. Show that any family of (non-zero) orthogonal idempotents in K_n has at most n members. (Hint. Consider direct sums in K^n .)

- (7) Let $A \subseteq B$ be rings and $R = \begin{pmatrix} A & B \\ 0 & B \end{pmatrix}$, where B is regarded as (A, B) -bimodule. Describe the left and right ideals of R . (Hint. Take first the case where B is a field.)
- (8) (Goodearl) Given rings $A \subseteq B$ and a bimodule $_A U_B$, consider the triangular matrix ring $R = \begin{pmatrix} A & U \\ 0 & B \end{pmatrix}$. For any modules M_A, N_B and a B -module homomorphism $f: M \otimes_A B \rightarrow N$, we can define $M \oplus N$ as a right R -module by the rule

$$(m, n) \begin{pmatrix} a & u \\ 0 & b \end{pmatrix} = (ma, f(m \otimes u) + nb).$$

Verify that this is indeed an R -module and show that every R -module is of this form.

4.4 Free modules

The theory of vector spaces over a field derives its relative simplicity from the fact that every vector space has a basis. As we have seen in 4.2, this holds even for vector spaces over skew fields, whether finite-dimensional or not. But it fails for general rings, and this leads to the notion of a free module, already briefly encountered in 10.4 of Vol. 1.

We recall that in any left R -module M , a family of elements x_1, \dots, x_n is called *linearly independent* if for any $\alpha_i \in R$ the relation $\sum \alpha_i x_i = 0$ holds only when $\alpha_1 = \dots = \alpha_n = 0$. More generally, an infinite family is said to be linearly independent if every finite subfamily is linearly independent; explicitly this means that the family (x_λ) is linearly independent if for any family (α_λ) of elements of R , almost all zero, the relation $\sum \alpha_\lambda x_\lambda = 0$ holds only when $\alpha_\lambda = 0$ for all λ . A linearly independent generating set is called a *basis* and a module is said to be *free* if it has a basis. It is clear that any basis of a free module is a *minimal generating set*, i.e. a generating set such that no proper subset generates the whole module.

For any set I there is a free left R -module F_I with a basis of cardinal $|I|$, unique up to isomorphism. It is obtained by taking the direct sum of $|I|$ copies of R ; explicitly F_I consists of all families of elements of R indexed by I , with almost all components zero, with componentwise addition and multiplication by elements of R . If the family with μ -component $\delta_{\lambda\mu}$ ($\lambda, \mu \in I$) is denoted by u_λ , then $(\alpha_\lambda) \in F_I$

can be uniquely written as $\sum \alpha_\lambda u_\lambda$ and this shows that the u_λ form a basis of F_I . Any other free R -module F' with basis (v_λ) of the same cardinal is isomorphic to F_I via the correspondence $\sum \alpha_\lambda u_\lambda \leftrightarrow \sum \alpha_\lambda v_\lambda$.

Free modules can also be characterized by their universal property:

THEOREM 4.1 *Let R be any ring. Then for any set I there is a left R -module F_I with a map $\varphi: I \rightarrow F_I$ which is universal for mappings into R -modules, i.e. given any mapping $f: I \rightarrow M$ into an R -module, there is a unique homomorphism $f': F_I \rightarrow M$ such that $f = \varphi f'$. The module F_I is in fact free on the image of I under φ .*

Proof. Let us define F_I as above, as the free module on a basis $(u_\lambda) (\lambda \in I)$. Given any map $f: I \rightarrow M$, if there is a homomorphism $f': F_I \rightarrow M$ of the required form, then clearly $u_\lambda f' = \lambda f (\lambda \in I)$; hence we must have

$$\left(\sum \alpha_\lambda u_\lambda \right) f' = \sum \alpha_\lambda (\lambda f). \quad (1)$$

This shows that at most one such homomorphism can exist. Now it is easily checked that the equation (1) indeed defines a homomorphism f'' with the desired properties. ■

As a useful consequence we have

COROLLARY 4.2 *Any short exact sequence*

$$0 \rightarrow M' \xrightarrow{\alpha} M \xrightarrow{\beta} M'' \rightarrow 0, \quad (2)$$

in which M'' is free, is split exact.

Proof. To show that (2) splits it is enough to find a left inverse of β . Let I be a basis of M'' ; since β is surjective, we can for each $\lambda \in I$ find $x_\lambda \in M$ mapped to λ by β . The mapping $\lambda \mapsto x_\lambda$ extends to a homomorphism $\beta': M'' \rightarrow M$, by Th. 4.1, and $\beta' \beta = 1$, by the definition of β' . Hence the sequence (2) splits, as we had to show. ■

The importance of free modules stems from the fact that every module can be expressed as a homomorphic image of a free module.

THEOREM 4.3 *Let R be any ring and M an R -module. Then there is a free R -module F with a submodule G such that $M \cong F/G$.*

Proof. Let M be a left module, say, with generating family $(a_\lambda), (\lambda \in I)$. Take a free left R -module F with basis $u_\lambda (\lambda \in I)$; the mapping $\lambda \mapsto a_\lambda$ extends to a homomorphism $f: \sum \alpha_\lambda u_\lambda \mapsto \sum \alpha_\lambda a_\lambda$, which is surjective, because the a_λ generate M . Hence $G = \ker f$ is a submodule of F such that $M \cong F/G$. ■

It is clear that a module M is finitely generated iff it can be written as F/G , where F is free and finitely generated. If $M \cong F/G$, where F and G are both finitely generated, we say that M is *finitely presented*; if we merely know that G is finitely generated, M is said to be *finitely related*.

We now turn to the question of comparing the cardinals of different bases of a given free module. In the case of vector spaces this cardinal is uniquely defined and is just the dimension of the space. We shall find that for most of the rings encountered here the cardinal is again unique, but this does not hold universally. We therefore consider briefly under what circumstances exceptions can occur.

A free R -module F with a basis of cardinal γ is said to have *rank* γ ; if all its bases have the same rank, F is said to have *unique rank*. In the first place we note that any free module which is not finitely generated always has unique rank. This result actually holds in a somewhat more general form. For any subset X of a left R -module M we shall denote by RX the submodule of M generated by X .

PROPOSITION 4.4 *Let M be a left R -module with a minimal generating set X . If X is infinite, of cardinal α , then any generating set of M has cardinal at least α . In particular, M is not finitely generated and any two minimal generating sets have the same cardinal.*

Proof. Let Y be any generating set of M , of cardinal β . Every $y \in Y$ is a linear combination of a finite number of elements from X , hence there is a finite subset X_y of X such that $y \in RX_y$. We assert that

$$X = \bigcup_{y \in Y} X_y. \quad (3)$$

For clearly $\bigcup X_y \subseteq X$, and $R\left(\bigcup X_y\right)$ is a submodule containing Y and hence equal to M . Thus $\bigcup X_y$ is a generating set, equal to X , by the minimality of the latter. If Y were finite, (3) would express X as a finite union of finite sets, which contradicts the fact that X is infinite. Hence Y must be infinite and from (3) and Prop. 1.2.6,

$$\alpha = |X| \leq \sum |X_y| \leq \aleph_0 \beta = \beta.$$

This shows that Y has cardinal at least α . If Y is also minimal, then by interchanging the roles of X and Y we see that $\beta \leq \alpha$, hence $\beta = \alpha$. ■

With the help of this result we can also establish the general form of Prop. 2.8.

COROLLARY 4.5 *Any two decompositions of a semisimple module into simple summands have the same number of terms and for suitable indexing corresponding terms are isomorphic.*

Proof. Suppose first that M is an isotypic module, say $M = \bigoplus_i S_i$. Take a non-zero element u_i of S_i ; then the family $\{u_i\}$ generates M and it is clearly a minimal generating set. If $M = \bigoplus_j T_j$ is another decomposition, then $|I| = |J|$, because any two minimal generating sets are equipotent, by Prop. 4.4. This proves the result for isotypic modules, since $S_i \cong T_j$ in this case. In the general case we can write M as a direct sum of type components: $M = \bigoplus H_\alpha$; given any expression of M as a direct sum of simple modules, if we group all terms of a given type α together we find the α -socle H_α . Hence in any decomposition the number of terms of type α is uniquely determined as the cardinal of a minimal generating set of H_α . ■

A ring R is said to have the *invariant basis property* or *invariant basis number* (IBN) if every free left R -module has unique rank. By Prop. 4.4 we need only consider finite rank; for a free module F of unique rank r we shall write $\text{rk}(F) = r$. Most rings commonly encountered have IBN, and exceptions are usually reckoned among the pathology of rings (cf. Ex. (2)). A trivial example is given by the trivial ring: here $1 = 0$, so the only module is 0 and $0^m \cong 0^n$ for all m, n . This case will usually be excluded in what follows.

Occasionally a stronger condition than IBN is needed. A ring R is said to be *weakly finite* if every generating set of n elements of R^n is linearly independent. Let R be any ring and suppose that R^n has a generating set of m elements, for some m, n . Then we have a surjective homomorphism $\beta: R^m \rightarrow R^n$, giving rise to an exact sequence

$$0 \rightarrow K \xrightarrow{\alpha} R^m \xrightarrow{\beta} R^n \rightarrow 0, \quad (4)$$

which splits by Cor. 4.2. Thus we have

$$R^n \cong K \oplus R^n. \quad (5)$$

In terms of matrices, let B be the $m \times n$ matrix describing the mapping β in (4), and A the $n \times m$ matrix describing the splitting homomorphism, so that $AB = I_n$. We can now restate the conditions for IBN and weak finiteness:

PROPOSITION 4.6 (i) *For any ring R the following conditions are equivalent:*

- (a) R has invariant basis number;
- (b) $R^m \cong R^n \Rightarrow m = n$;
- (c) if $A \in R^m$, $B \in R^n$ and $AB = I$, $BA = I$, then $m = n$.

(ii) *For any ring R , the following are equivalent:*

- (a) R is weakly finite;
- (b) $R^n \cong K \oplus R^n \Rightarrow K = 0$;
- (c) if $A, B \in R_n$ and $AB = I$, then $BA = I$.

Moreover, any non-trivial weakly finite ring has IBN.

Proof. (i) (b) is practically a restatement of the definition (a). The isomorphism $R^m \cong R^n$ is described by matrices as in (c), so the condition for IBN is that $m = n$.

(ii) To say that an n -element generator set of R^n is linearly independent just amounts to the assertion that $K = 0$, and the condition $K = 0$ holds precisely when $\beta\beta' = 1$, i.e. $BA = I$.

Finally, if IBN fails and $R \neq 0$, then we have $R^m \cong R^n$ for some $m \neq n$, say $m > n$. Then $R^n \cong R^{m-n} \oplus R^n$ and by (ii) (b) it follows that R is not weakly finite. ■

It is clear that the trivial ring is weakly finite, but does not have IBN. For examples of rings with IBN that are not weakly finite, see Cohn (1966). We remark that a ring R with the property $ab = 1 \Rightarrow ba = 1$ is sometimes called ‘directly finite’, ‘v.Neumann-finite’ or ‘inverse symmetric’.

The next result assures us that many of the rings we shall meet are in fact weakly finite (and hence also have IBN).

THEOREM 4.7 (i) *If a ring is weakly finite or has IBN, the same is true of the opposite ring.*

- (ii) *Any commutative ring is weakly finite.*
- (iii) *Any Artinian or Noetherian ring is weakly finite.*
- (iv) *If R is weakly finite, the same holds for any subring.*
- (v) *Given a ring homomorphism $f: R \rightarrow S$, if S has IBN, then so does R .*

Proof. (i) follows by remarking the symmetry of the conditions (c) in Prop. 4.6. To prove (ii) we note that when $AB = I$ over a commutative ring, then $(\det A)(\det B) = 1$, hence A is invertible and $AB = BA = I$.

(iii) Suppose that R fails to be weakly finite; then for some $n \geq 1$, $R^n \cong K \oplus R^n$ with $K \neq 0$, hence R^n is isomorphic to a proper submodule of itself. By repetition this leads to an infinite descending chain of submodules in R^n , which cannot exist if R is left Artinian, by Th. 2.2.8. Similarly, if $R^n \cong K \oplus R^n$, we can represent R^n as a proper homomorphic image of itself and, when we repeat the process, the kernels of these homomorphisms form an infinite ascending chain of submodules of R^n , which cannot exist when R is left Noetherian.

To establish (iv) we use the criterion (ii) (c) of Prop. 4.6, which is clearly inherited by subrings. Finally, to prove (v) suppose that R does not have IBN and let A, B be a pair of rectangular (non-square) matrices over R such that $AB = I$, $BA = I$. Applying f , we obtain a pair of non-square matrices Af, Bf satisfying the same equations, hence IBN also fails for S , and (v) is proved. ■

We conclude this section by noting that any ring other than a field has non-free modules. It is convenient to include several equivalent conditions. In any ring R , if $ab = 1$, we call a a *left inverse* of b and b a *right inverse* of a .

THEOREM 4.8 *Let R be a non-trivial ring. Any element of R is a unit provided it*

has a unique left inverse. Moreover the following conditions on R are equivalent:

- (a) every left R -module is free,
 - (b) every cyclic left R -module is free,
 - (c) R is simple as left R -module,
 - (d) every non-zero element of R has a left inverse,
 - (e) R is a skew field;
- (a°)–(e°) the left-right analogues of (a)–(e).

Proof. If $ab = 1$, then $(a + ba - 1)b = ab + bab - b = 1 + b - b = 1$, so by the uniqueness of left inverses, $ba = 1$, and therefore b is a unit.

(a) \Rightarrow (b) is clear. To prove (b) \Rightarrow (c), let \mathfrak{a} be a maximal left ideal of R . We have an exact sequence

$$0 \rightarrow \mathfrak{a} \rightarrow R \rightarrow F \rightarrow 0,$$

where F is cyclic and hence free, thus $F \cong R/\mathfrak{a}$. Since \mathfrak{a} was maximal, F is simple, and as cyclic free module, $F \cong R$, therefore R is simple as left R -module.

(c) \Rightarrow (d). Take $c \neq 0$ in R . Then $Rc = R$, hence $bc = 1$ for some $b \in R$ and hence c has a left inverse.

(d) \Rightarrow (e). Any non-zero element c has a left inverse b and since $b \neq 0$, it has a left inverse a . Now $ab = bc = 1$, hence $a = a \cdot bc = ab \cdot c = c$, so $c = a$ and $bc = cb = 1$. This shows R to be a skew field. Now (e) \Rightarrow (a) because over a skew field every module is free, as we saw in 4.2, and the equivalence to (a°)–(e°) follows by the symmetry of (e). ■

Exercises

- (1) Verify that the trivial ring is weakly finite but does not have IBN.
- (2) Let V be an infinite-dimensional vector space over a field k and put $R = \text{End}_k(V)$. Show that $R^2 \cong R$ as right R -modules and deduce that R does not have IBN.
- (3) Show that in any non-trivial ring R without IBN there exist positive integers h, k such that $R^m \cong R^n$ iff either $m = n$ or $m, n \geq h$ and $m \equiv n \pmod{k}$. (*Hint.* Take $(h, h+k)$ to be the first pair (m, n) in the lexicographic ordering such that $m \neq n$ and $R^m \cong R^n$.)
- (4) Let R be a ring which does not have IBN. Show that there is an integer h such that every finitely generated R -module can be generated by an h -element set.
- (5) Show that for any ring R the following are equivalent: (a) there is no bound on the number of generators of a finitely generated R -module; (b) for all m, n , $R^n \cong K \oplus R^m$ implies $n \geq m$; (c) for any $A \in {}^m R^n$, $B \in {}^n R^m$, if $AB = I$, then $n \geq m$.
- (6) A ring with the equivalent properties of Ex. (5) is sometimes said to possess UGN (unbounded generating number). Show that any non-trivial weakly finite ring has UGN, and any ring with UGN has IBN. (It can be shown that a ring has UGN iff some non-zero homomorphic image is weakly finite, cf. e.g. Cohn (1985, p. 8).)

4.5 Projective and injective modules

Projective modules form a generalization of free modules which is important in homological algebra. At first sight it looks a little unsatisfactory to replace the (seemingly) well-known notion of ‘free module’ by that of ‘projective module’, but the latter has the advantage of being defined categorically. By contrast, we cannot describe a free module without using a basis, even though the basis is not an invariant of the module.

We begin with a property of functors. Let $F:\mathcal{A} \rightarrow \mathcal{B}$ be a functor between categories of modules. We shall normally assume that F is *additive*, i.e. for any maps α, α' in \mathfrak{U} between the same modules, so that $\alpha + \alpha'$ is defined, we have

$$(\alpha + \alpha')^F = \alpha^F + \alpha'^F.$$

In other words, the mapping $\mathcal{A}(X, Y) \rightarrow \mathcal{B}(X^F, Y^F)$ is a group homomorphism. For example, for any $A \in {}_R\mathcal{M}$, the hom functors $h^A: X \mapsto \text{Hom}_R(A, X)$ and $h_A: X \mapsto \text{Hom}_R(X, A)$ are additive functors from ${}_R\mathcal{M}$ to Ab, the category of abelian groups. On the other hand, the functor $A \mapsto \text{Hom}_R(\text{Hom}_R(A, R), A)$ is not additive.

Consider any sequence

$$A \xrightarrow{\lambda} B \xrightarrow{\mu} C. \quad (1)$$

Any functor takes zero maps to zero maps and so, if $\lambda\mu = 0$, then $\lambda^F \cdot \mu^F = 0$. We shall be particularly interested in functors that preserve exact sequences. A functor F is said to be *exact* if it transforms an exact sequence of the form (1) into an exact sequence

$$A^F \xrightarrow{\lambda^F} B^F \xrightarrow{\mu^F} C^F. \quad (2)$$

Exact functors are rare and we usually have to be satisfied with less. We define a functor F to be *left exact* if the exactness of

$$0 \rightarrow A \xrightarrow{\lambda} B \xrightarrow{\mu} C \rightarrow 0 \quad (3)$$

implies that the sequence

$$0 \rightarrow A^F \xrightarrow{\lambda^F} B^F \xrightarrow{\mu^F} C^F \rightarrow 0 \quad (4)$$

is exact except possibly at C^F . Similarly, if (4) is exact except possibly at A^F , then F is called *right exact*. Clearly F is exact iff it is left and right exact. Similarly a contravariant functor $G: A \rightarrow B$ is *left exact* if the corresponding covariant functor $\text{op. } G: A^\circ \rightarrow B$ is left exact; a *right exact* contravariant functor G is defined similarly by the right exactness of $\text{op. } G$.

To check that F is left exact we need only verify that it preserves kernels. For, given (3), we have $A \cong \ker \mu$ and (4) will be exact at A^F and B^F precisely when

$A^F \cong \ker \mu^F$. Similarly F is right exact iff it preserves cokernels. Further, a contravariant functor is left exact iff it maps cokernels to kernels and right exact iff it maps kernels to cokernels.

THEOREM 5.1 *For any module category $R\mathcal{M}$ the functor $\text{Hom}_R(-, -)$ is left exact in each argument.*

Proof. For any $\mu: B \rightarrow C$ the kernel of the induced mapping $\mu^{h^M}: B^{h^M} \rightarrow C^{h^M}$ is the set of homomorphisms $f: M \rightarrow B$ annihilated by μ , i.e. the maps that factor uniquely through $\ker \mu = A$. Thus $\ker(\mu^{h^M}) = (\ker \mu)^{h^M}$, as claimed. Similarly, for any $\lambda: A \rightarrow B$ we have

$$\ker(\lambda^{h_N}) = (\text{coker } \lambda)^{h_N}. \quad \blacksquare$$

$$\begin{array}{ccccc} & & M & & \\ & \swarrow & f & \searrow & \\ A & \xrightarrow{\lambda} & B & \xrightarrow{\mu} & C \end{array}$$

The lack of exactness expressed in Th. 5.1 lends importance to the following

DEFINITION A module P is called *projective* if the covariant functor $h^P = \text{Hom}_R(P, -)$ is exact; a module I is called *injective* if the contravariant functor $h_I = \text{Hom}_R(-, I)$ is exact.

In checking exactness, the following lemma will be useful.

LEMMA 5.2 *Given a family of modules and homomorphism*

$$A_i \xrightarrow{\alpha_i} B_i \xrightarrow{\beta_i} C_i, \quad (5)$$

the following conditions are equivalent:

- (a) *all the sequences (5) are exact;*
- (b) *the sequence*

$$\prod A_i \rightarrow \prod B_i \rightarrow \prod C_i \quad (6)$$

is exact;

- (c) *the sequence*

$$\coprod A_i \rightarrow \coprod B_i \rightarrow \coprod C_i \quad (7)$$

is exact.

Proof. Take $x \in \prod B_i$, say $x = (x_i)$, and write $\alpha = \prod \alpha_i, \beta = \prod \beta_i$. We have ' $x\beta = 0$ ' \Leftrightarrow ' $x_i\beta_i = 0$ ' for all i ', and ' $x = y\alpha$ ' for some $y \in \prod A_i$ ' \Leftrightarrow ' $x_i = y_i\alpha_i$ for some $y_i \in A_i$ '. Hence if (5) is exact for all i , then (6) is also

exact, and conversely. For (c) the only difference is that $x = (x_i)$ has only finitely many non-zero components, but the argument is the same. ■

PROPOSITION 5.3 *Let (M_i) be a family of modules (over any ring R). Then*

- (i) $\prod M_i$ is injective if and only if each M_i is injective,
- (ii) $\coprod M_i$ is projective if and only if each M_i is projective.

In particular, for each finite family the direct sum $M_1 \oplus \cdots \oplus M_n$ is injective or projective if and only if each M_i is.

Proof. (i) $\prod M_i$ is injective $\Leftrightarrow \text{Hom}(-, \prod M_i)$ is exact $\Leftrightarrow \prod \text{Hom}(-, M_i)$ is exact (by (3) of 4.1) $\Leftrightarrow \text{Hom}(-, M_i)$ is exact for each i (by Lemma 5.2) $\Leftrightarrow M_i$ is injective for each i . This establishes (i); (ii) follows in the same way. ■

This proposition no longer holds when the product and coproduct are interchanged (cf. Ex. (19) and (20) in further exercises).

We can now give a number of alternative descriptions of projective modules:

THEOREM 5.4 *Let R be any ring. For any R -module P the following are equivalent:*

- (a) P is projective,
- (b) every short exact sequence with P in third position

$$0 \rightarrow A \xrightarrow{\lambda} B \xrightarrow{\mu} P \rightarrow 0 \quad (8)$$

splits,

- (c) P is a direct summand of a free module,
- (d) every homomorphism from P to a quotient of a module B can be lifted to B ; thus the diagram can be completed by a map $P \rightarrow B$ to give a commutative triangle.

$$\begin{array}{ccccc} & & P & & \\ & \nearrow f' & \downarrow f & & \\ B & \xrightarrow{\theta} & B'' & \rightarrow & 0 \end{array}$$

Proof. (a) \Rightarrow (b). Given a short exact sequence (8), we find by (a) that the sequence of abelian groups

$$0 \rightarrow \text{Hom}_R(P, A) \rightarrow \text{Hom}_R(P, B) \rightarrow \text{Hom}_R(P, P) \rightarrow 0$$

is exact. Now $1_P \in \text{Hom}_R(P, P)$ and by exactness there exists $\beta \in \text{Hom}_R(P, B)$ such that $\beta\mu = 1_P$, hence (8) splits.

(b) \Rightarrow (c). By Th. 4.3 we can write P as a quotient of a free module:

$$0 \rightarrow \ker \alpha \rightarrow F \xrightarrow{\alpha} P \rightarrow 0.$$

This splits, by (b), and so P is a direct summand of the free module F .

(c) \Rightarrow (d). By (c) we can write $F = P \oplus P'$, where F is free. Given a diagram as shown, we obtain a homomorphism from F to B'' by combining the projection $F \rightarrow P$ with the map $f: P \rightarrow B''$. Let (u_i) be a basis of F ; since θ is surjective, we can choose $a_i \in B$ such that $a_i\theta = u_i g$. Since F is free, there is a homomorphism $g': F \rightarrow B$ which maps u_i to a_i , hence $u_i g' \theta = a_i \theta = u_i g$ and so $g = g'\theta$. It follows that the map $f': P \rightarrow B$ obtained by combining the injection $P \rightarrow F$ with g' satisfies $f = f'\theta$, so it satisfies the conditions of (d).

(d) \Rightarrow (a). Given a short exact sequence, apply $h^P = \text{Hom}_R(P, -)$:

$$0 \rightarrow \text{Hom}_R(P, B') \rightarrow \text{Hom}_R(P, B) \rightarrow \text{Hom}_R(P, B'') \rightarrow 0. \quad (9)$$

By left exactness this can only fail to be exact at $\text{Hom}_R(P, B'')$. But by (c) every map $P \rightarrow B''$ lifts to a map $P \rightarrow B$ and this means that (9) is exact at $\text{Hom}_R(P, B'')$, as claimed. ■

Although a projective module does not in general have a basis, it has a ‘projective coordinate system’ with similar properties. We recall that in a free left R -module F with basis (u_i) every element x can be written as

$$x = \sum a_i u_i \quad (a_i \in R). \quad (10)$$

Here a_i is uniquely determined by x as the value of the i th projection from F to R . If we denote this projection mapping by α_i and write (α_i, x) for its value at x , so that $(\alpha_i, x) = a_i$, then (10) becomes

$$x = \sum (\alpha_i, x) u_i. \quad (11)$$

We shall find that such a representation (11) still holds for projective modules; in fact it characterizes them, although of course the coefficients are no longer unique, as they were for a free module.

PROPOSITION 5.5 (Dual basis lemma) *Let P be a left R -module, over any ring R , and (u_i) a generating system for P . Then P is projective if and only if there exist $\alpha_i \in \text{Hom}_R(P, R)$ such that for any $x \in P$, (α_i, x) vanishes for almost all i and*

$$x = \sum (\alpha_i, x) u_i \quad \text{for all } x \in P. \quad (12)$$

Proof. Let P be projective, take a free module F on a basis (v_i) in bijective correspondence with the u_i and map F to P by the rule $v_i \mapsto u_i$. Since the u_i generate P , we get a surjective mapping α and an exact sequence

$$0 \longrightarrow \ker \alpha \longrightarrow F \xrightarrow{\alpha} P \longrightarrow 0.$$

This sequence splits because P is projective; so there exists $\lambda: P \rightarrow F$ such that $x\lambda\alpha = x$ for all $x \in P$. Let $\pi_i: F \rightarrow R$ be the projection on the i th factor and

put $\alpha_i = \lambda\pi_i: P \rightarrow R$. Then $(\alpha_i, x) = 0$ for almost all i because this is true for π_i . Hence if $x \in P$, by (11) applied to F , we have

$$\begin{aligned} x &= x\lambda\alpha = [\sum (\pi_i, x\lambda)v_i]\alpha \\ &= \sum (\alpha_i, x)u_i. \end{aligned}$$

Conversely, given (12), a homomorphism $f: P \rightarrow B''$ and a surjection $\theta: B \rightarrow B''$ as in Th. 5.4(d), choose $a_i \in B$ to satisfy $a_i\theta = u_i f$, and define $f': P \rightarrow B$ by the rule $xf' = \sum (\alpha_i, x)a_i$. Then f' is an R -homomorphism: $(rx)f' = \sum (\alpha_i, rx)a_i = \sum r(\alpha_i, x)a_i = r(xf')$, and the diagram commutes: $xf'\theta = (\sum (\alpha_i, x)a_i)\theta = \sum (\alpha_i, x)u_i f = (\sum (\alpha_i, x)u_i)f = xf$. ■

The case when the family (u_i) is finite is worth stating separately:

COROLLARY 5.6 *The module P is a direct summand of R^n if and only if there exist $u_1, \dots, u_n \in P$, $\alpha_1, \dots, \alpha_n \in \text{Hom}_R(P, R)$ such that (12) holds.* ■

This means in particular: if P is finitely generated, by n elements say, then it is projective iff it is a direct summand of R^n . For example, a cyclic module is a direct summand of R iff it is projective; but a left ideal of R may well be projective without being a direct summand of R (because it is not principal).

Examples

1. Let k be a field (possibly skew). Then every k -module is free and hence projective; moreover every projective module (in fact every module) is free.
2. $R = \mathbf{Z}$. Every projective module is free (because every subgroup of a free abelian group is free). More generally, over a principal ideal domain R every submodule of a free R -module is free (cf. Vol. 1, Prop. 1 of **10.6**, p. 326, for the case of finite rank; Ex. (18), Further exercises for the general case), hence ‘projective’ and ‘free’ mean the same in this case.
3. For any coprime integers m, n we have $\mathbf{Z}/mn = \mathbf{Z}/m \oplus \mathbf{Z}/n$, by the Chinese remainder theorem (Vol. 1, p. 34). Regarding the summands on the right as (\mathbf{Z}/mn) -modules, we see that they are projective, but they are not free, because the number of elements in a free (\mathbf{Z}/mn) -module is divisible by mn .
4. A commutative integral domain in which all ideals are projective is called a *Dedekind domain* (cf. **9.5** below). Every commutative principal ideal domain is Dedekind, but not conversely; thus the ring of integers in $\mathbf{Q}(\sqrt{-5})$ is a non-principal Dedekind domain.

For injective modules there is a precise analogue of Th. 5.4(a), (b), (d), but not (c); this will not be needed here, and so is reserved for Vol. 3. For the present we note that a module I is injective iff every homomorphism $A' \rightarrow I$, where $A' \subseteq A$, can be extended to a homomorphism $A \rightarrow I$. Thus the diagram shown can be completed by a map $A \rightarrow I$ to make the triangle commutative.

$$\begin{array}{ccccc} 0 & \longrightarrow & A' & \longrightarrow & A \\ & & \downarrow & \nearrow & \\ & & I & & \end{array}$$

This follows because the property stated is equivalent to the exactness of the sequence

$$\text{Hom}(A, I) \rightarrow \text{Hom}(A', I) \rightarrow 0.$$

The condition for I to be injective can still be simplified: we have to extend a homomorphism from a submodule of A into I to a homomorphism $A \rightarrow I$. Since we can ascend to A by adjoining one element at a time, we need only postulate extendability to modules with cyclic quotients. This idea is made precise in

THEOREM 5.7 (Baer's criterion) *Let R be a ring and I a left R -module. Then I is injective if and only if every homomorphism $\mathfrak{a} \rightarrow I$, where \mathfrak{a} is a left ideal of R , can be extended to a homomorphism $R \rightarrow I$.*

Proof. The remark just made shows the condition to be necessary. Suppose that it holds and let an R -module A with a submodule A' be given, with a homomorphism $f: A' \rightarrow I$. We partially order the extensions of f with domain in A by writing $f_1 \leq f_2$ whenever the domain of f_2 contains that of f_1 and f_1, f_2 agree on the latter. These extensions form an inductive family, as is easily checked; hence by Zorn's lemma there is a maximal one $\bar{f}: \bar{A} \rightarrow I$. If $\bar{A} \neq A$, take $a \in \bar{A} \setminus A$ and put $B = \bar{A} + Ra$. The set $\mathfrak{a} = \{r \in R \mid ra \in \bar{A}\}$ is a left ideal of R and the map $f_0: \mathfrak{a} \rightarrow I$ defined by $rf_0 = (ra)\bar{f}$ is a homomorphism from \mathfrak{a} to I . By hypothesis it can be extended to R , i.e. there exists $u \in I$ such that $rf_0 = ru$. Now write

$$(x + ra)f' = x\bar{f} + ru \quad (x \in \bar{A}, \quad r \in R).$$

This is well-defined, for if $x + ra = 0$, then $r \in \mathfrak{a}$ and so $x\bar{f} + ru = x\bar{f} + (ra)\bar{f} = (x + ra)\bar{f} = 0$. Clearly it is a homomorphism of B into I and this contradicts the maximality of \bar{A} . Therefore $\bar{A} = A$ and f has been extended to A . ■

Let R be any integral domain; a left R -module M is said to be *divisible* if the equation in x :

$$u = ax \quad (u \in M, \quad a \in R^\times) \tag{13}$$

always has a solution in M . Every injective module is divisible: the map $ra \mapsto ru$ ($r \in R$) is a homomorphism $Ra \rightarrow M$; if M is injective, this extends to a homomorphism $R \rightarrow M$, and if $1 \mapsto x$ in this homomorphism, then x is a solution of (13). We shall examine the precise relationship between injective and divisible modules in Vol. 3; for the moment we shall show that the converse holds over a principal ideal domain:

PROPOSITION 5.8 *Every injective module over an integral domain is divisible. Over a principal ideal domain the converse holds: every divisible module is injective.*

It only remains to prove the converse; by Th. 5.7 we must show that for a divisible module M , any homomorphism of a left ideal \mathfrak{a} of R into M extends to a homomorphism of R into M . Let $\mathfrak{a} = Ra$ and $a \mapsto u$; then the homomorphism $\mathfrak{a} \rightarrow M$ is given by $f:ra \mapsto ru$. By divisibility the equation (13) has a solution x_1 say. Now the map $f_1:r \mapsto rx_1$ is a homomorphism extending f , for we have $(ra)f_1 = rax_1 = ru$. ■

In particular this tells us what the injective \mathbf{Z} -modules are. Examples of divisible abelian groups are \mathbf{Q} and $\mathbf{Z}(p^\infty)$, the group of all p^n th roots of unity, $n = 1, 2, \dots$. It can be shown that every divisible abelian group is a direct sum of groups of the above types. We shall not stop to show this as it will not be needed in the sequel (cf. Fuchs (1970) and Ex. (5) below).

Exercises

- (1) Show that any finitely generated projective module is finitely presented. Show that a finitely related projective module is a direct sum of a finitely presented module and a free module.
- (2) Show that if M is a finitely generated module over a Noetherian ring R , then $M^* = \text{Hom}_R(M, R)$ is again finitely generated.
- (3) Show that any quotient of a divisible module is again divisible.
- (4) Show that a torsion-free divisible module over a commutative integral domain is injective. (M is *torsion-free* if $rx = 0$, $0 \neq x \in M \Rightarrow r = 0$.)
- (5) Show that a divisible abelian p -group is a direct sum of copies of $\mathbf{Z}(p^\infty)$; show that a torsion-free divisible abelian group is a direct sum of copies of \mathbf{Q} . Hence determine the structure of a general divisible abelian group.
- (6) Let I be an injective left R -module and \mathfrak{a} an ideal in R . Show that the submodule of I annihilated by \mathfrak{a} is injective as (R/\mathfrak{a}) -module.

4.6 Duality of finite abelian groups

Let G be an abelian group, written multiplicatively. By a *character* on G one understands a homomorphism of G into the multiplicative group of non-zero complex numbers, $\chi: G \rightarrow \mathbf{C}^\times$. The characters of G can again be multiplied and form a group

$$\hat{G} = \text{Hom}(G, \mathbf{C}^\times),$$

called the *character group* or *dual* of G . For example, if $G = \mathbf{C}_n$ is cyclic of order n , with generator s , and $\chi: \mathbf{C}_n \rightarrow \mathbf{C}^\times$ is a character, then $\omega = \chi(s)$ satisfies the equation $\omega^n = 1$, and conversely, every function $\chi(s^i) = \omega^i$, where ω is an n th root of 1, is a character of \mathbf{C}_n . Further, the mapping $\chi \mapsto \chi(s)$ is a homomorphism from \hat{G} to \mathbf{C}^\times . In particular, choosing for $\chi(s)$ a primitive n th root of 1, we obtain a generator for $\hat{\mathbf{C}}_n$, clearly of order n ; thus we see that $\hat{\mathbf{C}}_n \cong \mathbf{U}_n$, where \mathbf{U}_n is the group of n th roots of 1 in \mathbf{C}^\times . This shows $\hat{\mathbf{C}}_n$ to be again cyclic of order n .

Let us now write our abelian group additively and put $\mathbf{T} = \mathbf{R}/\mathbf{Z}$. We observe that \mathbf{T} , the additive group of real numbers (mod 1), is isomorphic to the multiplicative group of complex numbers of modulus 1, via the mapping $x \mapsto \exp(2\pi i x)$. Since \mathbf{T} is divisible, it is injective as \mathbf{Z} -module and it follows that the functor $A \mapsto \hat{A} = \text{Hom}_{\mathbf{Z}}(A, \mathbf{T})$ is exact, so for any subgroup B of A we have the exact sequence

$$0 \rightarrow \widehat{A/B} \rightarrow \hat{A} \rightarrow \hat{B} \rightarrow 0. \quad (1)$$

Using these facts, we easily obtain

THEOREM 6.1 *Every finite abelian group is isomorphic to its dual:*

$$\hat{A} \cong A. \quad (2)$$

For the cyclic case has been checked above. In general we can write $A = A_1 \times \dots \times A_r$, where each A_i is cyclic, by the basis theorem for abelian groups. Now it is easily seen that

$$\text{Hom}(B \times C, \mathbf{T}) \cong \text{Hom}(B, \mathbf{T}) \times \text{Hom}(C, \mathbf{T}),$$

i.e. $B \times C \cong \hat{B} \times \hat{C}$. By induction it follows that $\hat{A} \cong \hat{A}_1 \times \dots \times \hat{A}_r$ and we have seen that $\hat{A}_i \cong A_i$, therefore $\hat{A} \cong A$, as claimed. ■

We note that the isomorphism depends essentially on the choice of basis in A ; there is no natural transformation from A to \hat{A} . However, we observe that there is a natural transformation

$$\alpha: A \rightarrow \hat{\hat{A}}, \quad (3)$$

of A into its *bidual* (i.e. second dual), defined as follows. Given $x \in A$, we define $\alpha_x \in \hat{\hat{A}}$ by the rule

$$\alpha_x: \chi \mapsto \chi(x). \quad (4)$$

This is a homomorphism from \hat{A} to \mathbf{T} , as is easily checked. We thus have an element α_x of $\hat{\hat{A}}$, for each $x \in A$, and it is clear that the mapping $x \mapsto \alpha_x$ is a homomorphism from A to $\hat{\hat{A}}$, indicated by (3). To show that α is a natural transformation, take a homomorphism of abelian groups $f: A \rightarrow B$. It induces a homomorphism $\hat{f}: \hat{B} \rightarrow \hat{A}$ and hence $\hat{\hat{f}}: \hat{\hat{A}} \rightarrow \hat{\hat{B}}$, and we have a diagram as shown below.

$$\begin{array}{ccc}
 A & \xrightarrow{f} & B \\
 \downarrow \alpha & & \downarrow \alpha \\
 \hat{A} & \xrightarrow{\hat{f}} & \hat{B}
 \end{array}$$

Here \hat{f} is defined by the equation

$$\alpha_x \hat{f} = \alpha_{xf}; \quad (5)$$

to say that α is natural means that this diagram commutes, which is just the condition (5).

To find the kernel of α we note that $\alpha_x = 0$ means: $\chi(x) = 0$ for all $\chi \in \hat{A}$. Now if $x \neq 0$, then there is a character χ_1 on the subgroup generated by x which does not vanish on x , and by (1), χ_1 can be extended to a character on A . Therefore we can find $\chi \in \hat{A}$ such that $\chi(x) \neq 0$, and this shows α_x in (4) to be non-zero when $x \neq 0$. Hence the mapping (3) is injective, for any abelian group A .

Suppose now that A is finite; then A and \hat{A} have the same order, by Th. 6.1, and hence α is then an isomorphism. If we only know that \hat{A} is finite, then so is \hat{A} , and since (3) is injective in any case, we again find that A is finite. Thus we have

THEOREM 6.2 *Let A be an abelian group. If either A or \hat{A} is finite, then so is the other and (3) is an isomorphism.* ■

We observe that for a finite abelian group A , of exponent m say, every character takes values in \mathbf{U}_m , the group of m th roots of 1; hence \hat{A} may then also be defined as $\text{Hom}(A, \mathbf{U}_m)$. For this reason all that has been said still applies if we take \mathbf{Q}/\mathbf{Z} instead of \mathbf{R}/\mathbf{Z} ; we remark that \mathbf{Q}/\mathbf{Z} is just the torsion subgroup of \mathbf{R}/\mathbf{Z} .

We conclude by noting the orthogonality relations of characters. Here G is taken to be multiplicative and the values are taken in \mathbf{C}^\times .

THEOREM 6.3 *Let G be a finite abelian group of order n . Then*

(i) *for any $\chi, \psi \in \hat{G}$,*

$$(1/n) \sum_{x \in G} \chi(x)\psi(x^{-1}) = \begin{cases} 1 & \text{if } \chi = \psi \\ 0 & \text{if } \chi \neq \psi; \end{cases}$$

(ii) *for any $x, y \in G$,*

$$(1/n) \sum_{x \in G} \chi(x)\chi(y^{-1}) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$$

Proof. (i) For $\chi = \psi$ the result is clear because $\chi(x)\chi(x^{-1}) = \chi(xx^{-1}) = 1$. Otherwise take $a \in G$ such that $\chi(a) \neq \psi(a)$ and replace x by ax . Since ax runs over

G as x does, we have

$$\sum \chi(x)\psi(x^{-1}) = \sum \chi(ax)\psi(x^{-1}a^{-1}) = \chi(a)\psi(a^{-1})\sum \chi(x)\psi(x^{-1}).$$

Hence $[1 - \chi(a)\psi(a^{-1})]\sum \chi(x)\psi(x^{-1}) = 0$, and here the first factor is not zero, by the choice of a , hence $\sum \chi(x)\psi(x^{-1}) = 0$. This proves (i). Now (ii) follows by applying the result to \hat{G} and using Th. 6.2. ■

Exercises

- (1) Let G be a finite abelian group and H a subgroup. Show that the annihilator of H in \hat{G} is a subgroup of \hat{G} of order $(G:H)$.
- (2) Show that for any field K (even skew) the correspondence $V \mapsto V^* = \text{Hom}_K(V, K)$ is an exact contravariant functor from left to right vector spaces.
- (3) Let R be a ring and T an R -bimodule. For any left R -module M define $M^* = \text{Hom}_R(M, T)$ as right R -module; further, for $N \subseteq M$ define N^\perp as the submodule of M^* annihilating N and for $P \subseteq M^*$ define P^\perp as the submodule of M annihilating P . Show that the correspondence $N \mapsto N^\perp$, $P \mapsto P^\perp$ is a Galois connexion.
- (4) Let R be a commutative ring and T an R -module which is injective and such that $M^* \neq 0$ whenever $M \neq 0$ (T is an *injective cogenerator*). Show that any submodule N of M satisfies $N^{\perp\perp} = N$.
- (5) Let φ be a non-trivial character on the additive group of a finite field F . Show that every character α of F has the form $\alpha = \alpha_b$, where $\alpha_b(x) = \varphi(bx)$.
- (6) Let α be a character of the additive group and β a character of the multiplicative group of \mathbf{F}_q . Show that $\overline{\alpha(x)} = \alpha(-x)$, $\overline{\beta(x)} = \beta(x^{-1})$. Show further that

$$\left| \sum_{x \in \mathbf{F}_q} \alpha(x)\beta(x) \right| = q^{1/2}$$

unless α or β is trivial; the sum is $q - 1$ if both are trivial, 0 if α is trivial and -1 if β is trivial.
(Hint. Evaluate the square.)

4.7 The tensor product of modules

The tensor product may be defined for any pair of bimodules, but before doing this we shall examine the simpler case of modules over a commutative ring. This will help us to understand the general case; it is also enough for many applications.

Let K be a commutative ring and U, V, W any K -modules. We shall write U, V, W as left modules, although, as we have seen, it is only a matter of notation whether a module over a commutative ring is regarded as a left or right module.

We want to consider bilinear mappings from U, V to W , i.e. mappings

$$f: U \times V \rightarrow W, \quad (1)$$

such that f is K -linear in each argument. Our object will be to construct a K -module T and a bilinear mapping $\lambda: U \times V \rightarrow T$ which is universal for all bilinear mappings (1), in the sense that to any bilinear mapping f as in (1) there corresponds a unique homomorphism $\bar{f}: T \rightarrow W$ such that the accompanying triangle commutes.

$$\begin{array}{ccc} U \times V & \xrightarrow{\lambda} & T \\ & \searrow f & \downarrow \bar{f} \\ & & W \end{array}$$

A module T with these properties is called a *tensor product* of U and V and is denoted by $U \otimes_K V$ or simply $U \otimes V$. If it exists it is clearly unique up to isomorphism (as a universal object), and we shall speak of the tensor product.

To prove the existence of T we form the free K -module A on the set* $U \times V$ and in A consider the submodule B generated by all the elements

$$\begin{aligned} (u + u', v) - (u, v) - (u', v) & \quad u, u' \in U, \\ (u, v + v') - (u, v) - (u, v') & \quad v, v' \in V, \\ (\alpha u, v) - \alpha(u, v), \quad (u, \alpha v) - \alpha(u, v), & \quad \alpha \in K. \end{aligned} \tag{2}$$

There is a mapping $\lambda: U \times V \rightarrow A/B$, obtained by taking the inclusion mapping $i: U \times V \rightarrow A$, followed by the natural homomorphism $v: A \rightarrow A/B$. This mapping is bilinear, for the elements (2) generating B were just chosen to ensure this. We set $T = A/B$ and claim that T , with the mapping λ , is the required tensor product. Let $f: U \times V \rightarrow W$ be any bilinear mapping; regarded as a set mapping, i.e. ignoring bilinearity, it may be extended to a unique homomorphism $f_1: A \rightarrow W$, because A is free on the elements (u, v) . We claim that $\ker f_1 \supseteq B$; for we have

$$\begin{aligned} [(u + u', v) - (u, v) - (u', v)] f_1 &= (u + u', v)f - (u, v)f - (u', v)f = 0, \\ [(\alpha u, v) - \alpha(u, v)] f_1 &= (\alpha u, v)f - \alpha[(u, v)f] = 0, \end{aligned}$$

by the bilinearity of f , and similarly for the other relations. Hence f_1 may be taken via T , by the factor theorem, and this provides the required mapping $g: T \rightarrow W$.

$$\begin{array}{ccccc} & U \times V & & & \\ & \swarrow i & \downarrow \lambda & \searrow f & \\ A & \xrightarrow{\nu} & T & \xrightarrow{f_1} & W \\ & \searrow g & \downarrow & \nearrow & \\ & & & & \end{array}$$

The mapping g is unique since its values are determined on the images of (u, v) in T and these form a generating set. Our conclusions may be summed up as follows:

*This is a set (of pairs), not to be confused with the K -module $U \oplus V$.

THEOREM 7.1 Let U, V be modules over a commutative ring K . Then there exists a K -module $U \otimes V$ together with a bilinear mapping $\lambda: U \times V \rightarrow U \otimes V$ which is universal for bilinear mappings from $U \times V$ to K -modules. ■

The image of (u, v) in $U \otimes V$ is denoted by $u \otimes v$. Thus $U \otimes V$ is a K -module with generating set $\{u \otimes v | u \in U, v \in V\}$ and defining relations

$$\begin{aligned} (u + u') \otimes v &= u \otimes v + u' \otimes v, & u, u' \in U, \\ u \otimes (v + v') &= u \otimes v + u \otimes v', & v, v' \in V, \\ (\alpha u) \otimes v &= u \otimes (\alpha v) = \alpha(u \otimes v) & \alpha \in K. \end{aligned}$$

There is another way of expressing Th. 7.1 which is often useful. Th. 7.1 states in effect that for any K -modules U, V, W there is a natural bijection between the bilinear mappings $U \times V \rightarrow W$ and the homomorphisms $U \otimes V \rightarrow W$. Now a mapping $f: U \times V \rightarrow W$ is linear in the second variable iff for each $u_0 \in U$, the mapping $V \rightarrow U \times V \rightarrow W$ given by $v \mapsto (u_0, v) \mapsto (u_0, v)f$ is linear. Further, f is bilinear iff in addition the mapping $U \rightarrow \text{Hom}_K(V, W)$ given by $u \mapsto (u, -)f$ is linear, i.e. $f \in \text{Hom}(U, \text{Hom}(V, W))$. Hence there is a natural bijection

$$\text{Hom}_K(U, \text{Hom}_K(V, W)) \cong \text{Hom}_K(U \otimes_K V, W). \quad (3)$$

This is easily verified to be an isomorphism of K -modules. The property (3) is known as *adjoint associativity*; later we shall see its general form for bimodules.

From the definition it is easy to check that tensor products satisfy the associative and commutative laws:

PROPOSITION 7.2 Let U, V, W be any K -modules, where K is any commutative ring. Then

$$U \otimes (V \otimes W) \cong (U \otimes V) \otimes W, \quad (4)$$

$$U \otimes V \cong V \otimes U. \quad (5)$$

Proof. We begin with (5). The rule $(u, v) \mapsto v \otimes u$ is a bilinear mapping $U \times V \rightarrow V \otimes U$, and hence gives rise to a homomorphism $\alpha: U \otimes V \rightarrow V \otimes U$, in which $u \otimes v \mapsto v \otimes u$. The general element of $U \otimes V$ has the form $\sum u_i \otimes v_i$ and it follows that $\alpha: \sum u_i \otimes v_i \mapsto \sum v_i \otimes u_i$. The same argument shows that $\beta: \sum v_i \otimes u_i \mapsto \sum u_i \otimes v_i$ is a homomorphism; clearly it is inverse to α , hence the latter is an isomorphism and (5) follows.

The proof of (4) is quite similar. We consider the mapping $\alpha: U \times V \times W \rightarrow U \otimes (V \otimes W)$ given by $(u, v, w) \mapsto u \otimes (v \otimes w)$. For fixed w this is bilinear in u, v and hence gives rise to a unique mapping $\alpha': (U \otimes V) \times W \rightarrow U \otimes (V \otimes W)$. It is easily checked that α' is bilinear and so gives rise to a mapping $\alpha'': (U \otimes V) \otimes W \rightarrow U \otimes (V \otimes W)$, in which $(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$. The inverse mapping is constructed in the same way and this shows α'' to be an isomorphism, which proves (4). ■

We observe that it is possible to define $U \otimes V \otimes W$ directly by the universal property for trilinear mappings, and a similar proof will show that it is isomorphic to either of the modules in (4). The same holds for more than three factors; this is just the generalized associative law (Vol. 1, p. 43). We shall therefore omit brackets in repeated tensor products.

Next we prove a ‘distributive’ law:

PROPOSITION 7.3 *For any K -modules U, V', V'' we have*

$$U \otimes (V' \oplus V'') \cong U \otimes V' \oplus U \otimes V''. \quad (6)$$

Proof. We show that the module on the right of (6) satisfies the universal property of the tensor product. A bilinear mapping from $U \times (V' \oplus V'')$ is given by $(u, v', v'') \mapsto (u \otimes v', u \otimes v'')$. If $f: U \times (V' \oplus V'') \rightarrow W$ is any bilinear mapping, then

$$(u, v' \oplus v'')f = (u, v')f + (u, v'')f,$$

and the expression on the right can be regarded as a mapping from $U \otimes V' \oplus U \otimes V''$; thus f is uniquely factored by the standard bilinear mapping, and the result follows. ■

The definition of the tensor product by a universal property is useful for proving the existence of mappings from $U \otimes V$ to a K -module, for we need only find the appropriate bilinear mapping from $U \times V$. It also has the merit of generality; but the definition is not such that it allows the structure of $U \otimes V$ to be read off. For example, if r, s are coprime integers, then $\mathbf{Z}/r \otimes \mathbf{Z}/s = 0$. To prove this we proceed as follows. Since r, s are coprime, there exist m, n such that $mr + ns = 1$. Now for any $a \in \mathbf{Z}/r, b \in \mathbf{Z}/s$ we have

$$a \otimes b = mr(a \otimes b) + ns(a \otimes b) = m(ra \otimes b) + n(a \otimes sb) = 0.$$

It follows that $\mathbf{Z}/r \otimes \mathbf{Z}/s = 0$, because the tensor product is generated by elements of the form $a \otimes b$.

It is important to bear in mind that the general element of $U \otimes V$ is not of the form $u \otimes v$, but is a sum of such terms: $\sum u_i \otimes v_i$. For example, if V is a free K -module, with basis e_1, \dots, e_n , then every element of $U \otimes V$ can be written uniquely in the form $\sum u_i \otimes e_i$ ($u_i \in U$), i.e. $U \otimes K^n \cong U^n$. To prove this, let us first take the case $n = 1$:

$$U \otimes K \cong U. \quad (7)$$

We have a bilinear mapping $\theta: (u, \lambda) \mapsto u\lambda$ from $U \times K$ to U , and if $f: U \times K \rightarrow W$ is any bilinear mapping, then $(u, \lambda)f = (u\lambda, 1)f$, hence $f = \theta f'$, where $f': u \mapsto (u, 1)f$, and clearly f' is the only mapping with this property. Thus U satisfies the universal property of Th. 7.1 and (7) follows. Now $U \otimes K^n \cong U^n$ follows by induction on n , using the distributive law (Prop. 7.3). Thus we obtain

PROPOSITION 7.4 *For any K -module U , the tensor product with a free K -module of rank n is a direct sum of n copies of U :*

$$U \otimes K^n \cong U^n. \quad \blacksquare \quad (8)$$

By symmetry a corresponding result holds for the first factor, and combining the two, we obtain

COROLLARY 7.5 *If U and V are free K -modules of finite rank, say $U \cong K^m$, $V \cong K^n$, then $U \otimes V \cong K^{mn}$. In particular, this applies for finite-dimensional vector spaces over a field, and we then have*

$$\dim(U \otimes V) = \dim U \cdot \dim V. \quad \blacksquare$$

Explicitly, if e_1, \dots, e_m is a basis for U and f_1, \dots, f_n a basis for V , then the elements $e_i \otimes f_j (i = 1, \dots, m, j = 1, \dots, n)$ form a basis for $U \otimes V$.

We record the property noted before (7), namely the independence property of the tensor product:

Let U be any K -module and V a free K -module with basis e_1, \dots, e_n . Then every element of $U \otimes V$ can be uniquely expressed in the form

$$\sum u_i \otimes e_i, \quad \text{where } u_i \in U. \quad (9)$$

Caution is needed in applying this result. Thus if $\sum u_i \otimes v_i = 0$ in $U \otimes V$, and the v_i are linearly independent over K , it does not follow that the u_i must vanish. If the submodule generated by the v_i is denoted by V' (so that the v_i form a basis for V'), then all we can conclude is that the u_i all vanish if $\sum u_i \otimes v_i = 0$ in $U \otimes V'$. Now the inclusion $V' \rightarrow V$ induces a homomorphism

$$U \otimes V' \rightarrow U \otimes V, \quad (10)$$

which however may not be injective. For example, the inclusion map $2\mathbf{Z} \rightarrow \mathbf{Z}$ is injective, but it does not remain so on tensoring with $\mathbf{Z}/2$. If $\mathbf{Z}/2$, \mathbf{Z} , $2\mathbf{Z}$ are generated by e, f, f' respectively, then $(\mathbf{Z}/2) \otimes \mathbf{Z}, (\mathbf{Z}/2) \otimes 2\mathbf{Z}$ are both isomorphic to $\mathbf{Z}/2$, by (7), with generators $e \otimes f, e \otimes f'$ respectively. But f' maps to $2f$ and $e \otimes f' \mapsto e \otimes 2f = 2e \otimes f = 0$. Thus (10) is the zero mapping in this case. A more precise analysis of this phenomenon will be undertaken in Vol. 3. For the moment we note that (10) is certainly injective if V' is a direct summand in V , by Prop. 7.3; so in that case we can identify $U \otimes V'$ with its image in $U \otimes V$. We note that this always holds when K is a field.

Let us next consider the effect of the tensor product on homomorphisms. Given any K -linear maps $\alpha: U \rightarrow U'$, $\beta: V \rightarrow V'$, there is a unique K -linear map $\alpha \otimes \beta: U \otimes V \rightarrow U' \otimes V'$ such that the left-hand square of the diagram below commutes:

$$\begin{array}{ccccc}
 U \times V & \xrightarrow{\alpha \times \beta} & U' \times V' & \xrightarrow{\alpha' \times \beta'} & U'' \times V'' \\
 \downarrow \lambda & & \downarrow \lambda' & & \downarrow \lambda'' \\
 U \otimes V & \xrightarrow{\alpha \otimes \beta} & U' \otimes V' & \xrightarrow{\alpha' \otimes \beta'} & U'' \otimes V''
 \end{array} \quad (11)$$

For the mapping $(u, v) \mapsto u\alpha \otimes v\beta$ from $U \times V$ to $U' \otimes V'$ is bilinear, and hence can be taken via $U \otimes V$, by the universal property of $U \otimes V$.

If $\alpha': U' \rightarrow U''$, $\beta': V' \rightarrow V''$ is another pair of homomorphisms, we obtain a commutative diagram (11). Since $(u, v)(\alpha \times \beta)(\alpha' \times \beta') = (u\alpha\alpha', v\beta\beta')$ for any $u \in U$, $v \in V$, we have

$$(\alpha \times \beta)(\alpha' \times \beta') = \alpha\alpha' \times \beta\beta',$$

and it follows from the diagram (11) that

$$\alpha\alpha' \otimes \beta\beta' = (\alpha \otimes \beta)(\alpha' \otimes \beta'). \quad (12)$$

Let us consider the special case $V'' = V' = V$, $\beta' = \beta = 1$. Then (12) reduces to

$$\alpha\alpha' \otimes 1 = (\alpha \otimes 1)(\alpha' \otimes 1), \quad (13)$$

and together with the obvious equation $1 \otimes 1 = 1$ this shows that the assignment $U \mapsto U \otimes V$ is a functor from K -modules to K -modules, for any given V . By symmetry the assignment $V \mapsto U \otimes V$ is also a functor for fixed U . Thus the tensor product is a bifunctor.

The above diagram shows that there is a correspondence between pairs of maps $(\alpha, \beta) \in \text{Hom}(U, U') \times \text{Hom}(V, V')$ and maps $\alpha \otimes \beta \in \text{Hom}(U \otimes V, U' \otimes V')$. Thus we have a mapping $(\alpha, \beta) \mapsto \alpha \otimes \beta$ which is clearly bilinear; by the universal property of the tensor product it induces a linear mapping

$$\text{Hom}_K(U, U') \otimes \text{Hom}_K(V, V') \rightarrow \text{Hom}_K(U \otimes V, U' \otimes V'). \quad (14)$$

We remark that for a pair of mappings $\alpha: U \rightarrow U'$, $\beta: V \rightarrow V'$ the expression $\alpha \otimes \beta$ is ambiguous: it may mean the element of the left of (14) or the induced homomorphism from $U \otimes V$ to $U' \otimes V'$, and one of these is mapped to the other in (14). It will usually be clear from the context which interpretation is intended; in some important cases the mapping (14) is an isomorphism and the ambiguity disappears. For example, when U and V are free of finite rank, say $U \cong K^n$, $V \cong K^m$, then (14) reduces to $U'^m \otimes V'^n \cong (U' \otimes V')^{mn}$, by a double application of Prop. 7.3, together with the relation

$$\text{Hom}(K^n, U) \cong U^n,$$

which follows by associating with $(u_1, \dots, u_n) \in U^n$ the map $\varphi: e_i \mapsto u_i$, where e_1, \dots, e_n is the standard basis of K^n . In particular, when $U' = U$, $V' = V$, we obtain

PROPOSITION 7.6 *If U, V are free K -modules of finite rank, then the mapping*

(14) induces the isomorphism

$$\text{End}_K(U) \otimes \text{End}_K(V) \cong \text{End}_K(U \otimes V). \quad \blacksquare$$

When we come to consider tensor products over a non-commutative ring, the corresponding construction leads in the first instance to abelian groups rather than modules. Thus let R be any ring, U a right and V a left R -module, and for any abelian group W consider mappings

$$f: U \otimes V \rightarrow W$$

which are *biadditive*, i.e. additive in each argument, and *R -balanced*, i.e.

$$f(ur, v) = f(u, rv) \quad \text{for all } u \in U, v \in V, r \in R.$$

We can again construct $U \otimes_R V$, now merely an abelian group, universal for R -balanced biadditive maps from $U \times V$ to abelian groups. The existence is proved as before, $U \otimes V = A/B$, where A is the free abelian group on $U \times V$ and B is the subgroup generated by

$$\begin{aligned} (u + u', v) - (u, v) - (u', v), & \quad u, u' \in U, \\ (u, v + v') - (u, v) - (u, v'), & \quad v, v' \in V, \\ (ur, v) - (u, rv), & \quad r \in R. \end{aligned}$$

Suppose now that U is an (S, R) -bimodule and V is an (R, T) -bimodule. Then the tensor product $U \otimes V$ just defined may be regarded as an (S, T) -bimodule in the following way. Take $s \in S$ and consider the mapping $\lambda_s: U \times V \rightarrow U \otimes V$ defined by

$$\lambda_s: (u, v) \mapsto su \otimes v.$$

Clearly this is biadditive and balanced; e.g. to prove the latter, we have $s(ur) \otimes v = (su)r \otimes v = su \otimes rv$, by the bimodule property of U . It follows that λ_s induces a homomorphism $U \otimes V \rightarrow U \otimes V$ which is simply denoted by s ; thus we have

$$s(\sum u_i \otimes v_i) = \sum su_i \otimes v_i. \quad (15)$$

If we do this for each $s \in S$, we obtain a left S -module structure on $U \otimes V$, for we have, for any $s, s' \in S$,

$$(ss')(u \otimes v) = (ss')u \otimes v = s(s'u) \otimes v = s[s'u \otimes v] = s(s'(u \otimes v)),$$

and of course $1(u \otimes v) = u \otimes v$. Similarly we can define a right T -module structure on $U \otimes V$ such that $(u \otimes v)t = u \otimes vt$ for $t \in T$, and $U \otimes V$ is an (S, T) -bimodule, because

$$s[(u \otimes v)t] = s[u \otimes vt] = su \otimes vt = (su \otimes v)t = [s(u \otimes v)]t.$$

Given any (S, T) -bimodule W , we can as before regard any homomorphism $f: U \times V \rightarrow W$ which is S -linear in the first, T -linear in the second argument and

R -balanced, as defining for each $u \in U$ a T -linear map $f_u: v \mapsto (u, v)f$. The set of all these T -linear maps has a natural (S, R) -bimodule structure induced from $\text{Hom}_T(V, W)$ and the map $u \mapsto f_u$ is a homomorphism of (S, R) -bimodules: $ur \mapsto f_{ur}$ and $(ur, v)f = (u, rv)f$ because f is R -balanced. Thus the natural homomorphism (3) leads to an isomorphism of S -bimodules, again called *adjoint associativity*:

$$\text{Hom}_T(U \otimes_R V, W) \cong \text{Hom}_R(U, \text{Hom}_T(V, W)) \quad ({}_S U_R, {}_R V_T, {}_S W_T). \quad (16)$$

By symmetry we likewise have an isomorphism of T -bimodules

$$\text{Hom}_S(U \otimes_R V, W) \cong \text{Hom}_R(V, \text{Hom}_S(U, W)). \quad (17)$$

Exercises

- (1) Define $U \otimes V \otimes W$ directly by the universal property for trilinear mappings and show that it is isomorphic to $U \otimes (V \otimes W)$. Do the same for products of more than three factors.
- (2) Give a direct proof that $\text{Hom}(U, \text{Hom}(V, W)) \cong \text{Hom}(V, \text{Hom}(U, W))$, for K -modules U, V, W .
- (3) Let A be a commutative integral domain with field of fractions K . Show that for any A -module U the kernel of the natural homomorphism $U \rightarrow K \otimes U$ is the torsion submodule of U , i.e. $\{u \in U \mid \alpha u = 0 \text{ for some } \alpha \in A^\times\}$.
- (4) Show that tensor products preserve infinite direct sums, but not necessarily direct products. Thus prove that $(\coprod_i U_i) \otimes V \cong \coprod_i (U_i \otimes V)$, but the natural mapping $(\prod_i U_i) \otimes V \rightarrow \prod_i (U_i \otimes V)$ need not be an isomorphism. (*Hint.* Take $U_i = \mathbf{Z}/2^i$, $i \in \mathbf{N}$, $V = \mathbf{Q}$.)
- (5) Let $U = \mathbf{Z}/2$ with basis u and $V = \mathbf{Z}^2$ with basis e_1, e_2 . Verify that $v_1 = e_1 + e_2, v_2 = e_1 - e_2$ are linearly independent and $u \otimes v_1 - u \otimes v_2 = 0$, but $u \neq 0$.
- (6) Let U, V be K -modules and $U^* = \text{Hom}(U, K)$, $V^* = \text{Hom}(V, K)$ their duals. Find a natural homomorphism $U^* \otimes V^* \rightarrow (U \otimes V)^*$ and give conditions on U and V under which this is an isomorphism.
- (7) Show that there is a natural homomorphism $U^* \otimes V \rightarrow \text{Hom}(U, V)$ under which $\alpha \times v$ corresponds to the map $u \mapsto \langle \alpha, u \rangle v$. Verify that this is an isomorphism if U is free of finite rank.
- (8) Show that for any K -module U , the mapping $\alpha \times u \mapsto \langle \alpha, u \rangle$ defines a homomorphism $U^* \otimes U \rightarrow K$. Verify that if U is free of finite rank, this defines a linear map $\text{End}(U) \rightarrow K$, which is just the trace.
- (9) Write down the matrix multiplication implicit in (12) and deduce that $\det(A \otimes B) = (\det A)^n (\det B)^m$, $\text{tr}(A \otimes B) = \text{tr } A \cdot \text{tr } B$.
- (10) Show that for infinite-dimensional spaces (14) need not be surjective.

Further exercises for Chapter 4

- (1) Let R be a ring such that in any left R -module any maximal linearly independent set is a basis. Give a direct proof that R is a skew field.
- (2) Let R be a ring such that in any torsion-free left R -module, given any linearly dependent set, one (suitably chosen) element of the set is a linear combination of the rest. Show that R is a local ring (i.e. the set of all non-units is an ideal).
- (3) Let M be a left R -module with annihilator $\mathfrak{a} = \{x \in R \mid xM = 0\}$. Show that if M is Noetherian, then R/\mathfrak{a} is left Noetherian.
- (4) Show that any module of finite composition length is finitely generated.
- (5) Let $0 \rightarrow M_1 \rightarrow \dots \rightarrow M_r \rightarrow 0$ be an exact sequence of modules. Show that if each module M_i has finite length t_i then $\sum (-1)^i t_i = 0$. Show that this still holds if the length is replaced by the multiplicity of a given type of simple module in a composition series.
- (6) Show that if f is an idempotent endomorphism of a module M , then $M = \text{im } f \oplus \ker f$.
- (7) For which integers n is \mathbb{Z}/n semisimple as \mathbb{Z} -module?
- (8) A module M is said to be an *essential extension* of a submodule N if N meets every non-zero submodule non-trivially. Show that a semisimple module cannot be an essential extension of a proper submodule.
- (9) Show that a module is an essential extension of its socle iff every non-zero submodule contains a simple submodule. Deduce that over an Artinian ring every module is an essential extension of its socle. Give an example of a Noetherian module which is not an essential extension of its socle.
- (10) Let $M = \bigoplus_I P_i$ be a direct sum of modules (not necessarily simple) and for $A \subset I$ write \sum_A for the sum of all P_i with $i \in A$. Show that $\sum_A \cap \sum_B = \sum_{A \cap B}$ for any subsets A, B of I .
- (11) Show that the submodules of a module M have unique complements iff M is semisimple and each α -socle is simple.
- (12) A module is called *semi-Artinian* if every non-zero quotient contains a simple submodule. Show that a module which is Noetherian and semi-Artinian has finite composition length.
- (13) Show that a left R -module is simple iff it is isomorphic to R/\mathfrak{a} , for some maximal left ideal \mathfrak{a} of R .
- (14) Let R be a commutative ring in which 1 has infinite additive order (e.g. a field of characteristic 0). By comparing traces of matrices show that R has IBN.

- (15) Show that $\text{Hom}_R(R, M) \cong M$, for any R -module M .
- (16) Show that $\text{Hom}(\mathbf{Z}/r, \mathbf{Z}/s) \cong \mathbf{Z}/d$, where $d = (r, s)$.
- (17) Show that $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Q} \cong \mathbf{Q}$, $A \otimes_{\mathbf{Z}} \mathbf{Q} = 0$ for any abelian torsion group A .
- (18) Let R be a principal ideal domain. Show that a submodule of a free R -module is free. (*Hint.* Use the projections on R and apply transfinite induction.)
- (19) (R. Baer) Fix a prime p and in $\mathbf{Z}^{\mathbb{N}}$ denote by B the subgroup of elements (a_i) such that for any $m \geq 0$ almost all the a_i are divisible by p^m . Assuming that $\mathbf{Z}^{\mathbb{N}}$ is free, show that B is free of uncountable rank; show that B/pB is elementary abelian of countable order and obtain a contradiction. Deduce that $\mathbf{Z}^{\mathbb{N}}$ is not free.
- (20) Show that over a Noetherian ring any direct sum of injective modules is injective (the Noetherian condition is necessary as well as sufficient, cf. Vol. 3).
- (21) Show that for any finitely generated projective R -module P , $\text{Hom}_R(P, M) \cong P^* \otimes_R M$, where $P^* = \text{Hom}_R(P, R)$. (*Hint.* Establish a natural transformation and take $P = R^n$.)

5

Rings and algebras

Historically the first rings to be studied (in the second half of the 19th century) were the rings of integers in algebraic number fields. At about the same time the theory of algebras began to develop; its most important landmarks were the Wedderburn structure theorems for semisimple algebras, and the study of the radical. The theories merged when it was realized that the Wedderburn theorems could be stated more generally for Artinian rings. This is the form in which the results will be presented here; the formulation for general rings (Jacobson radical and density theorem) will be reserved for Vol. 3.

5.1 Algebras: definition and examples

Most of our rings have a coefficient ring; this is often given as a field, so that the ring under consideration is a vector space. But it is convenient to frame the definitions more generally; all we need to impose on the coefficient ring is commutativity. At a later stage we shall see how even this restriction can be lifted.

Thus let K be a commutative ring; by a K -algebra we understand a ring A which is also a K -module, such that the multiplication in A is bilinear. Explicitly we have, in addition to the ring laws,

$$\alpha x \cdot y = x \cdot \alpha y = \alpha(xy), \quad \text{for any } x, y \in A, \quad \alpha \in K. \quad (1)$$

For example, K itself is always a K -algebra in a natural way. Let us also note that any ring R may be regarded as a \mathbf{Z} -algebra by putting

$$na = a + a + \cdots + a \quad (n \text{ terms}), \quad (-n)a = -na, \quad a \in R, n \in \mathbf{N}. \quad (2)$$

It is easily verified that R becomes a \mathbf{Z} -algebra in this way.

A *homomorphism* of K -algebras is a ring homomorphism which is K -linear, and a *subalgebra* of a K -algebra A is a subring admitting multiplication by all the elements of K . For example, the *centre* of A , defined as $C = \{z \in A \mid zx = xz \text{ for all } x \in A\}$ is a subalgebra of A , as is easily checked.

From (1) we find by taking $y = 1$ that $x \cdot \alpha 1 = \alpha x$, while $x = 1$ gives $\alpha 1 \cdot y = \alpha y$. Therefore

$$\alpha 1 \cdot x = x \cdot \alpha 1 = \alpha x, \quad x \in A, \quad \alpha \in K.$$

Taking $x = \beta 1$, we further obtain $\alpha 1 \cdot \beta 1 = \alpha(\beta 1) = \alpha\beta \cdot 1$. These rules, together with the equations $(\alpha + \beta) \cdot 1 = \alpha 1 + \beta 1$ and $1 \cdot 1 = 1$, which hold in any K -module, show that the map $\alpha \mapsto \alpha 1$ is a homomorphism of K into the centre of A . Conversely, given any ring R and a homomorphism f from K to the centre of R , we can define R as a K -algebra by putting

$$\alpha x = (\alpha f)x, \quad x \in R, \quad \alpha \in K.$$

This shows a K -algebra A to be nothing more than a ring with a homomorphism from K to the centre of A . In particular, any ring whose centre contains K as a subring may be regarded as a K -algebra.

Sometimes a K -algebra is defined more generally as a K -module A with a bilinear multiplication. What we described above is the special case when A is associative and has a 1, and so is a ring. Occasionally we shall want to consider the term in this more general sense, but when nothing is said to the contrary, K -algebras are understood to be associative with 1.

We remark that for a K -algebra A lacking a 1, we can always adjoin a 1 by forming the direct sum $K \oplus A$ with the multiplication

$$(\alpha, a)(\beta, b) = (\alpha\beta, \alpha b + \beta a + ab), \quad a, b \in A, \quad \alpha, \beta \in K. \quad (3)$$

The resulting K -algebra is denoted by A^1 ; it has the one $(1, 0)$ and is associative if A is.

A K -algebra A is said to be *augmented* if there is an algebra homomorphism $\varepsilon: A \rightarrow K$. The kernel of ε is an ideal of A , called the *augmentation ideal*. Augmented algebras may be described as follows.

THEOREM 1.1 *A K -algebra A is augmented if and only if the homomorphism $\alpha \mapsto \alpha 1$ ($\alpha \in K$) is an embedding and A contains an ideal \mathfrak{a} such that*

$$A = \mathfrak{a} \oplus K \cdot 1. \quad (4)$$

Proof. Assume that A is an augmented K -algebra, and write $\mathfrak{a} = \ker \varepsilon$ for the augmentation ideal. Since ε is a homomorphism, we have $(\alpha 1)\varepsilon = \alpha$, so $K \cdot 1$, the image of K in A , is isomorphic to K . Clearly $\mathfrak{a} \cap K \cdot 1 = 0$ and for any $x \in A$, $(x - x\varepsilon \cdot 1)\varepsilon = x\varepsilon = 0$, hence we have the direct sum (4).

Conversely, when A contains an ideal \mathfrak{a} and a copy $K \cdot 1$ of K such that (4) holds, then the projection on $K \cdot 1$ is an augmentation, as is easily verified. ■

We observe that for any algebra A without a 1 the algebra A^1 , formed as in (3) is augmented. Conversely, given an augmented algebra as in (4), we have $A = \mathfrak{a}^1$; the verification is straightforward and may be left to the reader.

If our coefficient ring is a field k , any k -algebra A is a vector space over k and so has a basis $\{u_i\}$ say. Every element of A can then be uniquely expressed in the form $\sum \alpha_i u_i$ ($\alpha_i \in k$, almost all 0). In particular, the multiplication in A is described

completely by the products $u_i u_j$. These take the form

$$u_i u_j = \sum \gamma_{ijr} u_r \quad \gamma_{ijr} \in k.$$

The elements γ_{ijr} are called the *structure constants* of A . They determine the algebra structure completely, by bilinearity, but it must be remembered that (i) the γ_{ijr} depend on the choice of the basis in A , and (ii) the γ_{ijr} cannot be assigned arbitrarily, for they will need to satisfy certain equations to ensure that A is associative and has a 1.

We recall that a ring is called a *division ring* or a *skew field* if every non-zero element is a unit, i.e. it has an inverse. A finite-dimensional k -algebra which is a skew field is also called a *division algebra*.

Let us give some examples of K -algebras (where K is any commutative ring):

1. Let V be any K -module and put $A = \text{End}_K(V)$. We can regard A as a K -module by defining for $f \in A$, $\alpha f: x \mapsto \alpha(xf) = (\alpha x)f$. With this definition it is easily checked that A is indeed a K -algebra. In particular, when V is the free K -module of rank n , $V = K^n$, then $A \cong \mathfrak{M}_n(K)$, the full $n \times n$ matrix ring over K (this is easily verified directly but also follows from Th. 1.3 below and Cor. 4.3.2).

2. $\mathfrak{T}_n(K)$, the set of all upper triangular $n \times n$ matrices over K , i.e. matrices $C = (c_{ij})$, where $c_{ij} = 0$ for $i > j$.

3. Let $M = \{u_i | i \in I\}$ be a monoid and take A to be the free K -module on M as basis, with multiplication $u_i u_j = u_r$ as in M . By linearity this defines a multiplication on A and it is not hard to see that A becomes a K -algebra in this way. The associativity follows from the associativity in M , and the unit element in M is also the 1 for A . This algebra A is usually denoted by KM and is called the *monoid algebra* on M ; in particular, for a group G we obtain the *group algebra* KG . The mapping $\varepsilon: \sum \alpha_i u_i \mapsto \sum \alpha_i$ is an augmentation for KM . The following are particular cases of this construction.

3.i Let $M = \{1, x, x^2, \dots\}$ consist of all powers of a single element x . Then M is the infinite cyclic monoid and KM consists of all polynomials in x over K , with the usual multiplication, thus $KM \cong K[x]$.

3.ii Let G be the infinite cyclic group, with generator x say. The group algebra is $K[x, x^{-1}]$, the ring of all *Laurent polynomials* in the indeterminate x .

3.iii Let $G = \mathbf{C}_n$ be the cyclic group of order n , with generator u . Then $K\mathbf{C}_n$ consists of all ‘polynomials’ $\alpha_0 + \alpha_1 u + \dots + \alpha_{n-1} u^{n-1}$ with the usual multiplication, but subject to $u^n = 1$.

4. Any subring of \mathbf{C} , the complex numbers, contains \mathbf{Z} and so is a \mathbf{Z} -algebra. Given $a \in \mathbf{C}$, consider $\mathbf{Z}[a]$, the subring generated by a . As abelian group this is not usually finitely generated; it is so precisely when a satisfies an equation

$$x^n + \alpha_1 x^{n-1} + \dots + \alpha_n = 0, \quad \text{where } \alpha_i \in \mathbf{Z}. \quad (5)$$

For if a satisfies (5), then we can express all powers of a as linear combinations of

$1, a, a^2, \dots, a^{n-1}$, so $\mathbf{Z}[a]$ is then finitely generated as abelian group. The converse also holds, as we shall see in 8.4. A complex number satisfying a monic equation (5) with integer coefficients is called an *algebraic integer*. Since the conjugates satisfy the same equation, they are again algebraic integers (though they need not lie in $\mathbf{Z}[a]$). By Gauss's lemma (Lemma 3.7.1) the irreducible polynomial with integer coefficients satisfied by a is again monic, hence it follows that the trace and norm of an algebraic integer are rational integers, i.e. they lie in \mathbf{Z} .

5. Finite-dimensional algebras over a field. Let k be a field and A an n -dimensional k -algebra with basis u_1, \dots, u_n . As we have seen, A is completely determined by the n^3 structure constants γ_{ijr} . If we choose the basis so that $u_1 = 1$, then

$$\gamma_{1ir} = \gamma_{i1r} = \delta_{ir},$$

while the associativity is expressed by the equations

$$\sum_v \gamma_{ijv} \gamma_{vrs} = \sum_v \gamma_{ivs} \gamma_{jrv}.$$

6. Any Boolean ring may be regarded as an \mathbf{F}_2 -algebra, as we saw in 2.3.

Let A be an n -dimensional algebra over a field k and consider the right multiplication in A ,

$$\rho_a : x \mapsto xa \quad (x \in A). \quad (6)$$

This is a k -linear mapping and hence can be represented by an $n \times n$ matrix over k . By a *matrix representation*, or simply a *representation*, one understands a k -algebra homomorphism into a full matrix ring over k . Now the right multiplication ρ_a is a representation of A , as is easily checked. It is called the *regular representation* of A ; more precisely, it is the *right regular representation*, the *left regular representation* being given by $\lambda_a : x \mapsto ax$. Strictly speaking, λ is an *antirepresentation*, because we have

$$\lambda_{ab} = \lambda_b \lambda_a. \quad (7)$$

To describe ρ_a explicitly, if $a = \sum \alpha_i u_i$, then

$$\rho_a : u_i \mapsto \sum u_i \alpha_j u_j = \sum_{jv} \alpha_j \gamma_{ijv} u_v.$$

Hence ρ_a is represented by the matrix

$$(\rho_a)_{ij} = \sum_v \alpha_v \gamma_{ivj}$$

relative to the basis $\{u_i\}$. We further note that ρ_A is *faithful*, i.e. its kernel as a homomorphism is 0. For if $\rho_a = 0$, then $a = 1 \cdot a = 1\rho_a = 0$. We thus obtain the analogue of Cayley's theorem for groups:

THEOREM 1.2 *Let A be an n -dimensional algebra over a field k . Then A is isomorphic to a subalgebra of the matrix ring $\mathfrak{M}_n(k)$.* ■

The regular representation can of course be defined for any K -algebra as a subring of $\text{End}_K(A)$, or indeed for any ring. It turns out that the rings of left and right multiplications are each others' centralizers in $\text{End}(R)$. This is an important result, best stated for general rings:

THEOREM 1.3 *Let R be any ring and regard R as left R -module over itself, by left multiplication. Then*

$$\text{End}_R(R) \cong R. \quad (8)$$

Similarly, $\text{End}_R(R_R) \cong R^\circ$. Moreover, if R is a K -algebra, then (8) is a K -algebra isomorphism.

Proof. Consider the mapping $\rho: a \mapsto \rho_a$ from R to $\text{End}_R(R)$, where ρ_a is given by (6). The image is in $\text{End}_R(R)$ because $(bx)a = b(xa)$ for any $b \in R$, and it is easily verified that $\rho_{a+b} = \rho_a + \rho_b$, $\rho_{ab} = \rho_a \rho_b$, $\rho_1 = 1$, and if R is a K -algebra, then $\rho_{\alpha a} = \alpha \rho_a$ for $\alpha \in K$. We have already seen that ρ is injective, and it only remains to prove surjectivity. Let $\theta \in \text{End}_R(R)$, say $1\theta = c$. Then for all $x \in R$, $x\theta = (x.1)\theta = x.1\theta = xc = x\rho_c$, hence $\theta = \rho_c$ and this shows ρ to be surjective; therefore it is an isomorphism. The assertion for λ follows similarly, but we have an anti-isomorphism this time, because of (7). ■

The result may also be expressed by saying that for any ring R , the multiplication rings ρ_R, λ_R are subrings of $\text{End}(R)$ and are centralizers of each other.

We can now give the description, promised in 4.3, of a class of rings whose only Morita-equivalents are the full matrix rings.

THEOREM 1.4 *Let R be a ring such that every finitely generated projective R -module is free. Then the rings Morita-equivalent to R are precisely the full matrix rings $\mathfrak{M}_n(R)$, $n = 1, 2, \dots$*

Proof. We saw in 4.3 that R_n is always Morita-equivalent to R . Conversely, suppose that S is Morita-equivalent to R ; then the categories ${}_S\mathcal{M}$ and ${}_R\mathcal{M}$ are equivalent, and finitely generated projective modules in these categories correspond to each other, for being projective is a categorical property and so is being finitely generated: a module is finitely generated iff it cannot be written as the union of a chain of proper submodules. By hypothesis every finitely generated left R -module has the form R^n , for some n . If S corresponds to R^n under the category equivalence, then these modules have isomorphic endomorphism rings. Now $\text{End}_R(R) \cong R$, by Th. 1.3, hence by Cor. 4.3.2, $S \cong \text{End}_S(S) \cong \text{End}_R(R^n) \cong \mathfrak{M}_n(R)$. ■

A ring with IBN having the property of Th. 1.4 is sometimes called *projective-free*.

Exercises

- (1) Show that every K -algebra (for any commutative ring K) on a single generator is commutative, and is a homomorphic image of the polynomial ring $K[x]$.
- (2) Show that every two-dimensional \mathbf{R} -algebra has a basis $1, u$, where u^2 is either 0 or 1 or -1 , and verify that no two of these are isomorphic. Classify all two-dimensional k -algebras (i) with 1, (ii) without 1.
- (3) Let A be a k -algebra without 1, where k is a field of characteristic not 2, and suppose that $x^2 = 0$ for all $x \in A$. Show that $xyz = 0$ for all $x, y, z \in A$.
- (4) Show that the group algebra of every finite non-trivial group has zero-divisors.
- (5) Let S be a semigroup and kS its semigroup algebra over a field k (defined as for monoids). Show that if kS has a 1, then so does S .
- (6) Let R be any ring and K a commutative ring. Show that $R \otimes_{\mathbf{Z}} K$ is a K -algebra and that the map $\lambda: R \rightarrow R \otimes_{\mathbf{Z}} K$ given by $r \mapsto r \otimes 1$ is a ring homomorphism such that for any ring homomorphism $f: R \rightarrow A$, where A is a K -algebra, there is a unique K -algebra homomorphism $f': R \otimes K \rightarrow A$ such that $f = \lambda f'$.
- (7) Verify that the equations (2) define a \mathbf{Z} -algebra structure on any ring.
- (8) Show that in any non-trivial ring R , if $axa = a$ has a unique solution for each $a \in R$, then R is a skew field.
- (9) Find a non-trivial ring R such that $\mathfrak{M}_2(R) \cong R$. (Hint. Use Ex. (2) of 4.4 and Th. 1.4.)
- (10) Let k be a field. Show that (i) a non-zero k -algebra A contains an isomorphic copy of k in its centre and (ii) any ring R whose centre contains a subfield isomorphic to k can be defined as a k -algebra. Examine what goes wrong if R contains a subfield isomorphic to k , but not contained in the centre of R .

5.2 Direct products of rings

Let (R_i) , $i \in I$, be any family of rings. Its *direct product* is the ring obtained by forming the Cartesian product

$$R = \prod R_i \tag{1}$$

and defining all operations componentwise, thus $(x_i) + (y_i) = (x_i + y_i)$, $(x_i)(y_i) = (x_i y_i)$, $1 = (1_i)$, where 1_i is the one of R_i . It is easily seen that R is again a ring; the natural projections $\pi_i: R \rightarrow R_i$ are homomorphisms, while the injections $\mu_i: R_i \rightarrow R$ preserve addition and multiplication, but not the 1, and so are not homomorphisms. For the same reason the images of the R_i in R are not subrings, although they are ideals in R .

We shall mainly be interested in finite direct products; they may be described as follows:

THEOREM 2.1 *Let R_1, \dots, R_t be any rings and denote their direct product by R , with natural projections $\pi_i: R \rightarrow R_i$ and injections $\mu_i: R_i \rightarrow R$. Then $\mathfrak{a}_i = \text{im } \mu_i$ is an ideal in R and*

$$R = \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_t. \quad (2)$$

Moreover,

$$\mathfrak{a}_i \mathfrak{a}_j = 0 \quad \text{if } i \neq j, \quad \mathfrak{a}_i \mathfrak{a}_i \subseteq \mathfrak{a}_i. \quad (3)$$

Conversely, any ring R of the form (2), where each \mathfrak{a}_i is an ideal in R , may be expressed as a direct product of rings R_1, \dots, R_t , where R_i is isomorphic to \mathfrak{a}_i , qua ring*.

Proof. We have seen that a finite direct product of abelian groups is isomorphic to their direct sum. Hence in the finite case, (2) follows from (1). Now whenever (2) holds, then $\mathfrak{a}_i \mathfrak{a}_j \subseteq \mathfrak{a}_i \cap \mathfrak{a}_j$ because the \mathfrak{a}_i are ideals in R , and since $\mathfrak{a}_i \cap \mathfrak{a}_j = 0$ for $i \neq j$, (3) follows. Conversely, any ring of the form (2) is a direct product of the \mathfrak{a}_i as abelian groups, and (3) ensures that each \mathfrak{a}_i is an ideal in R and the multiplication is performed componentwise. Moreover, the natural homomorphism from R with kernel $\sum_{j \neq i} \mathfrak{a}_j$ has image \mathfrak{a}_i which is therefore a ring R_i say, and so $R \cong \prod R_i$. ■

Let R be a ring with ideals $\mathfrak{c}_1, \dots, \mathfrak{c}_t$ and write $R_i = R/\mathfrak{c}_i$. Then we have the natural homomorphisms $f_i: R \rightarrow R_i$ which can be combined to give a homomorphism

$$f: R \rightarrow \prod R_i. \quad (4)$$

Explicitly we have $f: x \mapsto (xf_1, \dots, xf_t)$. We ask: When is (4) an isomorphism? This is answered by the next result, which generalizes the Chinese remainder theorem (Th. 5, 2.3, p. 34 of Vol. 1). In a ring R , two submodules U, V (e.g. ideals) will be called *comaximal* if $U + V = R$; for left (or right) ideals this holds precisely when $1 \in U + V$.

THEOREM 2.2 *Let R be a ring with ideals $\mathfrak{c}_1, \dots, \mathfrak{c}_t$. Put $R_i = R/\mathfrak{c}_i$ and let $f: R \rightarrow \prod R_i$ be the homomorphism (4). Then*

- (i) *f is injective if and only if $\bigcap \mathfrak{c}_i = 0$,*
- (ii) *f is surjective if and only if the \mathfrak{c}_i are pairwise comaximal.*

*On account of (2), R used to be called the *direct sum* of the R_i by some authors, but this is rather misleading when rings with 1 are considered, as is the case here.

Proof. (i) Clearly $\ker f = \cap \ker f_i = \cap c_i$. (ii) If f is surjective, take $x \in R$ mapping to $(1, 0, \dots, 0)$; then $x \equiv 1 \pmod{c_1}$, $x \equiv 0 \pmod{c_2}$, so $1 = 1 - x + x \in c_1 + c_2$, hence c_1 and c_2 are comaximal; similarly for other pairs c_i, c_j . Conversely, assume that the c_i are pairwise comaximal and put

$$a_i = \bigcap_{j \neq i} c_j;$$

since $c_1 + c_i = R$ for $i > 1$, there exist $x_i \in c_i, y_i \in c_i$ such that $x_i + y_i = 1$, and therefore $y_2 \dots y_t = (1 - x_2) \dots (1 - x_t) \equiv 1 \pmod{c_1}$ and $y_2 \dots y_t \in a_1$, so $c_1 + a_1 = R$. It follows that there exists $e_1 \in R$ which maps to $(1, 0, \dots, 0)$ under f . By symmetry there exists $e_i \in R$ mapping to an element with the i th component 1 and the rest 0. Now for any $a_1, \dots, a_t \in R$, $\sum a_i e_i \mapsto (a_1 f_1, \dots, a_t f_t)$, hence f is surjective, as claimed. ■

In the decomposition (2) the unit element e_i of a_i is *idempotent*: $e_i^2 = e_i$; *central*: $e_i x = x e_i$ for all $x \in R$; and distinct e 's are *orthogonal*: $e_i e_j = 0$ for $i \neq j$. Moreover, their sum is 1:

$$e_1 + \dots + e_t = 1. \quad (5)$$

Thus a direct decomposition (2) of R yields a decomposition of 1 as a sum of pairwise orthogonal central idempotents. Conversely, let R be a ring in which 1 admits a decomposition (5) as a sum of pairwise orthogonal central idempotents. Then $a_i = e_i R$ is an ideal in R for which (2) and (3) hold, so R is then a direct product of the a_i , by Th. 2.1. This proves

PROPOSITION 2.3 *For any ring R the representations (2) of R as a finite direct product correspond to the decompositions of 1 as a sum of pairwise orthogonal central idempotents.* ■

A central idempotent e is said to be *centrally primitive* if $e \neq 0$ and e cannot be written in the form $e = u + v$, where u, v are central idempotents and $uv = 0$, $u, v \neq 0$. Clearly a non-zero ring R is indecomposable (into a direct product of several factors) iff 1 is centrally primitive in R .

Any decomposition (5) with a maximal number of terms has centrally primitive summands. Such maximal decompositions exist in any Noetherian or Artinian ring (but need not exist in general rings). For let R be Noetherian say, and choose a direct decomposition $R = a_1 \oplus a'_1$ in which a'_1 is a (proper) maximal ideal. Then a_1 will be directly indecomposable. Next write $a'_1 = a_2 \oplus a'_2$, where a'_2 is maximal among ideals strictly contained in a'_1 . Continuing in this way, we obtain a sequence of decompositions

$$R = a_1 \oplus \dots \oplus a_t \oplus a'_t;$$

since R is Noetherian, the ascending chain of ideals $a_1 \oplus \dots \oplus a_t \oplus a'_t$ ($t = 1, 2, \dots$) must break off and we conclude that $a'_t = 0$ at some stage. Clearly each a_i is directly

indecomposable (by the maximality property of a'_i), so we have achieved the desired decomposition of R . In an Artinian ring we proceed similarly, choosing a_i minimal $\neq 0$ at each stage. Then the a'_i form a descending chain which must again break off. So in either case we obtain a decomposition of R as a direct product of indecomposable rings. Such a decomposition must be unique, for if

$$R = a_1 \oplus \cdots \oplus a_t = b_1 \oplus \cdots \oplus b_s,$$

where all the a_i, b_j are indecomposable as rings, then $a_1 = a_1 b_1 \oplus \cdots \oplus a_1 b_s$ and by indecomposability $a_1 b_j \neq 0$ for just one suffix j , say $j = 1$. Hence $a_1 = a_1 b_1$ and similarly $a_1 b_1 = b_1$, so $a_1 = b_1$. Now $a_2 \oplus \cdots \oplus a_t = b_2 \oplus \cdots \oplus b_s$, for each side is the annihilator of a_1 in R . By induction on t we have $s = t$ and for suitable numbering, $a_i = b_i$. Thus we have

PROPOSITION 2.4 *Any Artinian or Noetherian ring can be written in just one way as a direct product of a finite number of indecomposable rings; the factors are uniquely determined as the ideals generated by the centrally primitive central idempotents.* ■

In rings without finiteness conditions the direct product decomposition (1) may be replaced by a representation of R as a sheaf of simpler rings over a certain topological space, the *decomposition space* of R . This is defined as the Boolean space associated with the Boolean algebra of central idempotents in R (cf. Pierce 1967).

Exercises

- (1) In any ring R define the *additive exponent* as the additive order of 1 (if finite). If R has additive exponent n , show that $nR = 0$, and R can be written as a direct product of rings whose additive exponents are prime powers.
- (2) Show that if e is a non-zero idempotent in R , then for any $x \in R$, $e + ex(1 - e)$ is another non-zero idempotent. Deduce that e is central iff it commutes with every idempotent in R .
- (3) Show that any two decompositions of 1 into sums of pairwise orthogonal central idempotents have a common refinement (obtained by decomposing the terms further, resp. taking the products of the idempotents from the two decompositions). Deduce that a decomposition with a maximal number of terms is unique.
- (4) Show that if R is a *reduced* ring (i.e. a ring with no non-zero nilpotent elements), then every idempotent in R is central.
- (5) A ring R is said to be *strongly regular* if for each $a \in R$, the equation $xa^2 = a$ has a solution. Show that R is strongly regular iff every principal (left or) right ideal is generated by a central idempotent. (*Hint.* Show that R is reduced and compute $(axa - a)^2$.)

(6) A ring is called *completely primary* if every element of it is either invertible or nilpotent. Show that every Artinian commutative ring can be written as a finite direct product of completely primary rings. (*Hint.* If $c \in R$ is neither invertible nor nilpotent, find $a \in R$ and $n \in \mathbb{N}$ such that ac^n is an idempotent $\neq 0, 1$.)

(7) Let k be a field and $\alpha_1, \dots, \alpha_n$ distinct elements of k . Show that the ideals $(x - \alpha_i)$ are pairwise comaximal in $k[x]$; hence find, for any $\beta_1, \dots, \beta_n \in k$, a polynomial f in $k[x]$ such that $f \equiv \beta_i \pmod{(x - \alpha_i)}$. Deduce the Lagrange interpolation formula in the form $f = \sum \beta_i e_i$, where e_i is the unique polynomial of degree at most $n - 1$ determined by $e_i(x)(x - \alpha_i) = c_i f(x)$, the constant c_i being chosen so that $e_i(\alpha_i) = 1$. Deduce that the e_i are pairwise orthogonal idempotents whose sum is 1.

5.3 The Wedderburn structure theorems

We now come to one of the central results of ring theory, giving an explicit description of simple and semisimple Artinian rings. A ring R is called *simple* if $R \neq 0$ and R has no ideals other than 0 or R . R is called *left semisimple* if it is semisimple as left R -module over itself. Below we shall see that a left semisimple ring is also right semisimple and (left and right) Artinian, so we just speak of semisimple rings. By contrast there are simple rings that are not Artinian.

We begin with a lemma which is fundamental in much that follows.

LEMMA 3.1 (Schur's lemma) *Let R be any ring. If M is a simple R -module, then $\text{End}_R(M)$ is a skew field.*

Proof. We have to show that every non-zero endomorphism of M is an automorphism. Let $\alpha: M \rightarrow M$ be a non-zero endomorphism. Then $\ker \alpha$ is a submodule of M ; since $\alpha \neq 0$, it follows that $\ker \alpha \neq M$, and so by simplicity, $\ker \alpha = 0$. Similarly, $\text{im } \alpha$ is a submodule, non-zero and hence equal to M . Therefore α is bijective and its inverse is easily seen to be an endomorphism. Thus every non-zero endomorphism of M has an inverse, so $\text{End}_R(M)$ is a skew field. ■

We begin by describing simple Artinian rings:

THEOREM 3.2 (Wedderburn's first structure theorem) *For any ring R the following conditions are equivalent:*

- (a) R is simple and left Artinian,
- (b) R is left semisimple and all simple left R -modules are isomorphic,
- (c) $R \cong \mathfrak{M}_n(K)$, where K is a skew field and $n \geq 1$,
- (a°)–(c°) the right-hand analogues of (a)–(c).

Moreover, the integer n in (c) is unique and K is unique up to isomorphism.

Proof. (a) \Rightarrow (b). Let Rx be a minimal left ideal of R . By the simplicity of R we

have $R = RxR = \sum Rx a_i$, where a_i ranges over R . The left ideal $Rx a_i$ is a homomorphic image of Rx , by the map $rx \mapsto rxa_i$; hence by the minimality of Rx , either $Rx a_i = 0$ or $Rx a_i \cong Rx$. Hence R is a sum of left ideals isomorphic to Rx and by Th. 4.2.3, a direct sum, so (b) holds. Further, any simple R -module is a quotient of R by a left ideal, hence isomorphic to a minimal left ideal.

(b) \Rightarrow (c). Since R is finitely generated (by 1) as left R -module, and semisimple by hypothesis, it is a direct sum of finitely many left ideals, all isomorphic among themselves. Let U be a minimal left ideal and suppose that $R \cong U^n$. By Schur's lemma, $K = \text{End}_R(U)$ is a skew field; by Cor. 4.3.2, $\text{End}_R(U^n) \cong \mathfrak{M}_n(K)$; and by Th. 1.3, $\text{End}_R(R) \cong R$; hence $R \cong \mathfrak{M}_n(K)$, as claimed. Here n is uniquely determined as the composition length of $_R R$, while K is unique up to isomorphism as the endomorphism ring of the unique simple left R -module type.

(c) \Rightarrow (a). $K_n = \mathfrak{M}_n(K)$ has finite dimension as left K -space; every left ideal is a subspace, so the descending chain condition holds and K_n is left Artinian. To show that K_n is simple, take any $a = (a_{ij}) \neq 0$, say $a_{rs} \neq 0$. Then $e_{ir}ae_{sj}a_{rs}^{-1} = e_{ij}$, hence the ideal generated by a contains all the e_{ij} and so coincides with K_n . This shows K_n to be simple.

Finally, since condition (c) is left-right symmetric, (a^o) and (b^o) also hold, and again imply (c). ■

COROLLARY 3.3 *Any simple Artinian ring is an algebra over a field.*

For $\mathfrak{M}_n(K)$ is an algebra over the centre of K . ■

This result actually holds without the Artinian hypothesis (Ex. (9)).

Next we turn to semisimple rings.

THEOREM 3.4 (Wedderburn's second theorem) *Every left semisimple ring is a finite direct product of full matrix rings over skew fields:*

$$R \cong \mathfrak{M}_{n_1}(K_1) \times \cdots \times \mathfrak{M}_{n_r}(K_r), \quad (1)$$

where the n_i and the isomorphism types of the K_i are determined by R . Conversely, every ring of the form (1) is semisimple; in particular, every left semisimple ring is right semisimple and (left and right) Artinian. Moreover, two minimal left ideals of R are isomorphic if and only if they lie in the same factor on the right of (1).

Proof. Since R is left semisimple and finitely generated as left ideal, we have $R = H_1 \oplus \cdots \oplus H_r$, where $H_r \cong I_i^{n_i}$, I_i being a minimal left ideal and different I 's being non-isomorphic. By Schur's lemma, $\text{End}_R(I_i) = K_i$ is a skew field, $\text{End}_R(H_i) \cong \mathfrak{M}_{n_i}(K_i)$ and using Th. 1.3 and Cor. 4.2.7, we have

$$R \cong \text{End}_R(R) \cong \prod \mathfrak{M}_{n_i}(K_i).$$

Here n_i and the isomorphism type of K_i is determined by the type component H_i of R , which itself is unique, as we have seen in 4.2.

Conversely, for any skew field K and any $n \geq 1$, we have $\mathfrak{M}_n(K) \cong I^n$, where I is a minimal left K_n -module, represented for example by a single column of the matrix ring $\mathfrak{M}_n(K)$. Hence $\prod \mathfrak{M}_n(K_i) \cong \bigoplus I_i^{n_i}$ is left semisimple. It has finite composition length and so is left Artinian. By the evident symmetry of $\mathfrak{M}_n(K)$ it is also right semisimple and right Artinian. ■

For finite-dimensional algebras there is a stronger conclusion, for which we shall need a sharper form of Schur's lemma:

LEMMA 3.5 *Let k be an algebraically closed field, R a k -algebra and U a simple R -module which is a finite-dimensional k -space. Then $\text{End}_R(U) \cong k$.*

Proof. As R -module, U is also a k -module and by hypothesis it is finite-dimensional, so each R -endomorphism α of U represented by a matrix $\rho(\alpha)$. Since k is algebraically closed, $\rho(\alpha)$ has an eigenvalue λ in k . Thus $\rho(\alpha) - \lambda \cdot 1$ is singular, and it defines an endomorphism of U , which by Schur's lemma can only be the zero map. Hence $\rho(\alpha) = \lambda \cdot 1$ and so $\text{End}_R(U) \cong k$. ■

PROPOSITION 3.6 *Let k be an algebraically closed field. Then any finite-dimensional k -algebra which is semisimple is a direct product of full matrix rings over k .*

Proof. If we go through the proof of Th. 3.4, we now find that each I_i is a finite-dimensional k -space, hence by Lemma 3.5, $\text{End}_R(I_i) \cong k$ and the conclusion follows from Th. 3.4. ■

Ths. 3.2 and 3.4 were proved by J. H. M. Wedderburn in 1908 for finite-dimensional k -algebras. In 1928 E. Artin made the observation that these theorems were valid more generally for any rings satisfying the ascending and descending chain conditions on right ideals. E. Noether remarked in 1929 that Th. 3.2 needed only the descending chain condition. Then in 1939 C. Hopkins proved that in any ring the ascending chain condition is a consequence of the descending chain condition (cf. 5.4 below). Thus any Artinian ring is necessarily Noetherian (here the presence of a unit element is essential). For this reason the Wedderburn theorems are now usually stated for Artinian rings. I. Schur first proved the lemma that bears his name in his dissertation in 1901, where he gave a simplified treatment of group representations, a topic we shall take up in Ch. 7.

From the description of semisimple rings in Th. 3.4 we easily derive other conditions which are sometimes useful.

THEOREM 3.7 *For any ring R the following conditions are equivalent:*

- (a) *R is semisimple,*
- (b) *every left R -module is semisimple,*

- (c) every left R -module is projective,
- (d) every left R -module is injective,
- (a°)–(d°) the right-hand analogues of (a)–(d).

Proof. (a) \Leftrightarrow (b). (a) is a special case of (b); when (a) holds, then every direct sum of copies of R is semisimple, hence so is every homomorphic image, but this includes every R -module, by Th. 4.4.3.

(b) \Leftrightarrow (c) \Leftrightarrow (d). By Th. 4.2.3, (b) holds iff every short exact sequence of left R -modules splits, and this is equivalent to each of (c), (d). Now the right-hand analogues follow by the symmetry of (a). ■

Let R be a simple Artinian ring, say $R \cong \mathfrak{M}_r(K)$. Every left R -module is a direct sum of copies of the unique simple left R -module. If the number of copies in the direct sum is α , we may define the *dimension* of M over R as

$$[M : R] = \frac{1}{r} \cdot \alpha. \quad (2)$$

This is either a rational number with denominator r or ∞ . Here the normalization has been chosen so that $[R : R] = 1$. In a similar way the dimension of a module over a semisimple ring could be defined as an m -tuple of rational numbers, but this will not be needed.

Exercises

- (1) Show that in the full matrix ring over a skew field, $\mathfrak{M}_n(K)$, each row is a minimal right ideal and each column is a minimal left ideal.
- (2) For any ring R and any $n \geq 1$, show that the centre of $\mathfrak{M}_n(R)$ is isomorphic to the centre of R .
- (3) Show that the centre of a semisimple ring is a direct product of fields. Show that if the centre of a semisimple ring R is a field, then R is simple.
- (4) Let R be a semisimple ring such that the endomorphism ring of any simple R -module is commutative. Show that $R^\circ \cong R$, even though R° need not equal R (the latter equality holds iff R is commutative).
- (5) Show that the dimension of a module M over a simple ring R , as defined in (2), is a non-negative integer iff M is finitely generated free.
- (6) Let $R = \prod_1^r \mathfrak{M}_{n_i}(K_i)$. Show that every finitely generated R -module M is defined up to isomorphism by an r -tuple of rational numbers $\alpha = (\alpha_1, \dots, \alpha_r)$ such that $n_i \alpha_i \in \mathbb{N}$. What is the condition on α for M to be free?

- (7) Show that every homomorphic image of a semisimple ring is again semisimple, but that this need not hold for every subring.
- (8) Show that every right ideal in a semisimple ring is generated by an idempotent, which is central iff the ideal is two-sided.
- (9) Let R be a simple ring. Show that any two non-zero elements of R have the same additive order λ , which is either 0 or a prime number. Verify that R may be defined as a P -algebra, where P is the prime field of characteristic λ .
- (10) Show that if R is left Artinian, (or left Noetherian) and $n \geq 1$, then so is $\mathfrak{M}_n(R)$.

5.4 The radical

In general an Artinian ring need not be semisimple, but we shall find that there is a uniquely determined ‘largest’ homomorphic image which is semisimple. The kernel of this homomorphism is called the *radical*. To see the form this takes let us consider for a moment the commutative case. From Th. 3.4 it is clear that a commutative Artinian ring is semisimple iff it is a direct product of fields. The simplest case is that where R is a homomorphic image of a polynomial ring $k[x]$, k being a field. This is a principal ideal domain (cf. 10.5 in Vol. 1), so every ideal has the form (f) , where f is a polynomial in x over k . Let

$$f = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$$

be a complete factorization of f into irreducible factors p_i (where the p_i are distinct). By the Chinese remainder theorem, Th. 2.2, we have

$$R = k[x]/(f) = \prod_{i=1}^r k[x]/(p_i^{\alpha_i}),$$

and this is a direct product of fields iff $\alpha_1 = \cdots = \alpha_r = 1$. It follows that the largest semisimple homomorphic image of R is $k[x]/(p_1 \cdots p_r)$. We note that if \bar{u} denotes the residue class $(\bmod f)$ of $u \in k[x]$, then $p_1 \cdots p_r m = 0$, where $m = \max \{\alpha_1, \dots, \alpha_r\}$. Thus $p_1 \cdots p_r$ is nilpotent and we see that R is semisimple whenever it has no nilpotent elements apart from 0. We shall soon see that this condition always holds for commutative Artinian rings, but it certainly does not hold generally in this form, since e.g. the 2×2 matrix ring over a field is semisimple and yet contains the nilpotent element e_{12} . What is required is a nilpotent *ideal*; the ideal generated by e_{12} is not nilpotent because it contains the non-zero idempotent $e_{11} = e_{12}e_{21}$. An ideal \mathfrak{a} is said to be *nilpotent* if $\mathfrak{a}^n = 0$ for some n , where \mathfrak{a}^n is the sum of all terms $a_1a_2 \dots a_n$, $a_i \in \mathfrak{a}$. We shall find that the radical in an Artinian ring can be characterized as the sum of all nilpotent ideals. This will be proved in Th. 4.5 below, but to do so it will be convenient to define the notion of radical in quite general rings. First we consider what happens in the Artinian case.

THEOREM 4.1 Every Artinian ring R contains a unique ideal \mathfrak{N} such that (i) R/\mathfrak{N} is semisimple and (ii) any ideal c of R such that R/c is semisimple has the property that $c \supseteq \mathfrak{N}$.

The ideal \mathfrak{N} determined in this way is called the *radical* of R .

Proof. If a is any maximal ideal of R , then the quotient R/a is simple Artinian, because the ideals of R/a correspond to the ideals of R that contain a . Given any maximal ideals a_1, a_2, \dots of R , we can form the descending chain

$$a_1 \supset a_1 \cap a_2 \supset a_1 \cap a_2 \cap a_3 \supset \dots, \quad (1)$$

where at each stage a_{r+1} is chosen so that $a_1 \cap \dots \cap a_r \not\subseteq a_{r+1}$, to ensure that the inclusion in (1) is proper. Since R is Artinian, the chain (1) must break off, so for some r , $\mathfrak{N} = a_1 \cap a_2 \cap \dots \cap a_r$ contains all maximal ideals of R . Since the a_i are distinct and maximal, they are pairwise comaximal, so by Th. 2.2 we have

$$R/\mathfrak{N} \cong \prod_{i=1}^r R/a_i.$$

Each factor on the right is simple Artinian, hence R/\mathfrak{N} is semisimple. If c is any ideal of R such that R/c is semisimple, then c can be expressed as an intersection of maximal ideals of R ; we need only take the ideals of R corresponding to the maximal ideals of R/c . Thus $c = \bigcap b_i$, where each b_i is maximal. Since $b_i \supseteq \mathfrak{N}$, it follows that $c \supseteq \mathfrak{N}$. Thus \mathfrak{N} has the required properties, and it is clear that \mathfrak{N} is uniquely determined by (i) and (ii). ■

The proof of Th. 4.1 shows that the radical of an Artinian ring R may be defined as the intersection of all maximal ideals of R . From the structure of semisimple rings it is clear that the radical may also be defined as the intersection of all maximal left ideals, or equivalently as the intersection of all maximal right ideals. This property was used by Jacobson in 1945 (taking up an earlier idea of Perlis), to study the radical in general rings. In fact there are several equivalent properties of the radical which will also be needed, and we begin by introducing them.

LEMMA 4.2 Let R be any ring. For an element $a \in R$ the following conditions are equivalent:

- (a) for each simple left R -module M , $aM = 0$,
- (b) a belongs to each maximal left ideal of R ,
- (c) $1 - xa$ has a left inverse for all $x \in R$,
- (d) $1 - xay$ has an inverse for all $x, y \in R$,
- (a°)–(d°) the right-hand analogues of (a)–(d).

Proof. (a) \Rightarrow (b). Let m be a maximal left ideal of R . Then R/m is a simple left R -module, hence $a(R/m) = 0$ by (a), i.e. $aR \subseteq m$, and so $a \in m$.

(b) \Rightarrow (c). Assume that (b) holds but not (c). Then for some $x \in R$, $1 - xa$ has no left inverse, i.e. $1 \notin R(1 - xa)$. By Krull's theorem (Th. 2.2.11) there is a maximal left ideal $m \supseteq R(1 - xa)$, thus $1 - xa \in m$, and by (b), $a \in m$; hence $xa \in m$ and $1 = 1 - xa + xa \in m$, a contradiction.

(c) \Rightarrow (a). Let M be a simple left R -module. Given $u \in M$, if $au \neq 0$, then $Rau = M$ by simplicity, hence $u = xau$ for some $x \in R$, i.e. $(1 - xa)u = 0$, and by (c) it follows that $u = 0$. Hence $au = 0$ for all $u \in M$ and (a) holds.

Now (d) \Rightarrow (c) trivially; we complete the proof by showing that (a) + (c) \Rightarrow (d). By (a), $aM = 0$ for any simple module M , hence for any $y \in R$, $ayM \subseteq aM = 0$, i.e. ay satisfies (a) and hence (c). Thus $1 - xay$ has a left inverse $1 - b$ say:

$$(1 - b)(1 - xay) = 1. \quad (2)$$

This may be written as $b = (b - 1)xay$, hence $bM = 0$ for any simple M ; therefore b satisfies (a) and hence (c), so $1 - b$ has a left inverse $1 - c$:

$$(1 - c)(1 - b) = 1. \quad (3)$$

By (2), (3)

$$1 - c = (1 - c)(1 - b)(1 - xay) = 1 - xay,$$

hence $c = xay$ and now (2), (3) show $1 - b$ to be the inverse of $1 - xay$.

This proves that (a)–(d) are equivalent; now the symmetry follows because (d) is left-right symmetric. ■

We note that if $1 - a$ has the inverse $1 - a'$, then

$$a + a' = aa' = a'a. \quad (4)$$

Such an element a' is uniquely determined by a ; it is called the *quasi-inverse* of a . In any ring R the set $\mathbf{J}(R)$ of all $a \in R$ satisfying one (and hence all) of (a)–(d) is called the *Jacobson radical* of R . By the remarks following Th. 4.1 it is clear that $\mathbf{J}(R)$ coincides with the radical in an Artinian ring. We now give another proof of this fact which helps to clarify the relation between the two ways of defining the radical.

THEOREM 4.3 *Let R be a left (or right) Artinian ring and $J = \mathbf{J}(R)$ its Jacobson radical. Then R/J is semisimple and J is the least ideal with semisimple quotient.*

Proof. Pick maximal left ideals I_1, I_2, \dots in R such that $I_r \not\supseteq I_1 \cap \dots \cap I_{r-1}$ as far as possible. Then

$$R \supset I_1 \supset I_1 \cap I_2 \supset \dots \supset I_1 \cap \dots \cap I_r, \quad (5)$$

and since R is Artinian, the chain eventually breaks off, say at the r th stage. This means that every maximal left ideal contains $I_1 \cap \dots \cap I_r$, which therefore coincides with J . If any I_i , $1 \leq i \leq r$, contains the intersection of the others, we omit

it; so we may take the intersection $I_1 \cap \dots \cap I_r$ to be *irredundant*, i.e. no I_i can be omitted. We now have $I_i \not\supseteq a_i = \bigcap_{j \neq i} I_j$ and it follows that $a_i + I_i = R$, because I_i is maximal. Furthermore, $a_i \cap I_i = \bigcap_{j \neq i} I_j = J$ and

$$I_1 \cap \dots \cap I_{i-1} / I_i \cap \dots \cap I_r \cong (I_1 \cap \dots \cap I_{i-1} + I_i) / I_i = R / I_i,$$

where $I_1 \cap \dots \cap I_{i-1}$ and I_i are comaximal by the maximality of I_i . Since R / I_i is a simple module, this shows that the chain (5) has simple quotients and hence has finite composition length. Moreover, $a_i / J \cong (a_i + I_i) / I_i \cong R / I_i$, hence a_i / J is simple, and the sum $\sum (a_i / J)$ is direct, for if $x_i \in a_i$ satisfies $\sum x_i \equiv 0 \pmod{J}$, then $x_i \in a_1 \cap I_1 = J$, hence $x_i \equiv 0 \pmod{J}$, and similarly for each x_i . This shows $R / J = \bigoplus (a_i / J)$ to be semisimple.

J is the least ideal with this property, for if c is such that R / c is semisimple, then there are left ideals b_i minimal above c such that $R / c = \bigoplus (b_i / c)$. Clearly, if $d_1 = \sum_{j \neq 1} b_j$, then $R / d_1 = (d_1 + b_1) / d_1 \cong b_1 / (d_1 \cap b_1) = b_1 / J$ and this is simple, hence d_1 is maximal and similarly for $d_i = \sum_{j \neq i} b_j$. By definition of J , $d_i \supseteq J$, hence $c = \bigcap d_i \supseteq J$, as we had to show. ■

This theorem describes the structure of R / J in the Artinian case. It remains to examine J itself; in particular we wish to identify J in the Artinian case as the maximal nilpotent ideal. We firstly note that every nilpotent element has a quasi-inverse. For any x satisfies

$$(1 - x)(1 + x + x^2 + \dots + x^{n-1}) = (1 + x + x^2 + \dots + x^{n-1})(1 - x) = 1 - x^n.$$

This shows that if $x^n = 0$, then $1 - x$ has an inverse, i.e. x has a quasi-inverse. By a *nilideal* one understands an ideal in which every element is nilpotent; every nilpotent ideal is clearly a nilideal, but the converse need not hold.

PROPOSITION 4.4 *Every left or right nilideal of a ring R is contained in $J(R)$. In particular, this holds for every nilpotent (left or right) ideal.*

The proof is immediate from the above remarks and condition (d) of Lemma 4.2. ■

We can now give the promised description of the radical in an Artinian ring.

THEOREM 4.5 *In any left (or right) Artinian ring R the sum of all nilpotent ideals is itself a nilpotent ideal, the radical J , and R / J is semisimple. Moreover, R itself is semisimple if and only if it has no nilpotent ideals other than zero.*

Proof. By Prop. 4.4, $J = J(R)$ contains all nilpotent left or right ideals and R / J is semisimple, by Th. 4.3, while no smaller ideal has this property. To complete the proof it therefore only remains to show that J itself is nilpotent. We have $J \supseteq J^2 \supseteq \dots$ and since R is left Artinian, equality must hold at some stage:

$J^k = J^{k+1} = \dots$. We put $I = J^k$, so that $I^2 = I$; our aim will be to show that $I = 0$. If $I \neq 0$, let \mathfrak{n} be a minimal left ideal subject to $I\mathfrak{n} \neq 0$. Then for some $a \in \mathfrak{n}$, $Ia \neq 0$, and $I(Ia) = I^2a = Ia$, hence by the minimality of \mathfrak{n} , $Ia = \mathfrak{n}$. In particular, $a = xa$ for some $x \in I$, so $(1-x)a = 0$, but $x \in J$, hence by Lemma 4.2(c), $a = 0$, which is a contradiction. Therefore $I = J^k = 0$, as claimed. ■

We note an important consequence of Lemma 4.2 which is often useful.

LEMMA 4.6 (Nakayama's lemma) *Let R be any ring and M a finitely generated non-zero left R -module. Then $\mathbf{J}(R)M \neq M$.*

Proof. Since M is non-zero finitely generated, it has a maximal proper submodule M' (Prop. 2.2.10), and M/M' is simple. By Lemma 4.2(a), $a(M/M') = 0$ for all $a \in \mathbf{J}(R)$, hence $\mathbf{J}(R)M \subseteq M'$ and so $\mathbf{J}(R)M \neq M$. ■

We record another form of Nakayama's lemma, which is also used:

COROLLARY 4.7 *Let R be any ring, M a finitely generated left R -module and N a submodule of M . If $JM + N = M$, where $J = \mathbf{J}(R)$, then $N = M$.*

For we need only apply Lemma 4.6 to M/N . By hypothesis $J(M/N) = M/N$, and M/N is finitely generated, hence $M/N = 0$, i.e. $N = M$. ■

We conclude this section by showing that every Artinian ring is Noetherian (Hopkins' theorem).

THEOREM 4.8 *Let R be a left Artinian ring and M a left R -module. Then the following conditions are equivalent:*

- (a) M is Artinian,
- (b) M is Noetherian,
- (c) M has a composition series.

Proof. The quotient $A = R/J$ of R by the radical $J = \mathbf{J}(R)$ is semisimple; moreover J is nilpotent, say $J^k = 0$. We form the chain of submodules

$$M \supseteq JM \supseteq J^2M \supseteq \dots \supseteq J^kM = 0. \quad (6)$$

Each quotient $F_i = J^{i-1}M/J^iM$ is annihilated by J and so may be regarded as an A -module, and R -submodules and A -submodules are the same. As A -module, F_i is semisimple, by Th. 3.7. Now if M is Artinian, then each F_i is Artinian and as semisimple Artinian A -module it has a composition series. By composing all these series we obtain a composition series for M . Hence (a) \Rightarrow (c); similarly (b) \Rightarrow (c), and the reverse implications are clear. ■

Applying the result to R itself, we obtain

COROLLARY 4.9 Every left Artinian ring is left Noetherian. ■

In general a ring may well be left Artinian without being right Artinian, see e.g. Ex. (3).

Exercises

- (1) Find the radical of \mathbf{Z}/n , for different n .
- (2) Find the radical of $\mathfrak{T}_n(k)$, the ring of all upper triangular $n \times n$ matrices over a field k .
- (3) Show that the ring $\begin{pmatrix} \mathbf{R} & \mathbf{R} \\ 0 & \mathbf{Q} \end{pmatrix}$ is left but not right Artinian, and $J^2 = 0$.
- (4) Show that a reduced Artinian ring is a direct product of skew fields.
- (5) Show that $\mathbf{J}(R_n) = \mathbf{J}(R)_n$ for any $n \geq 1$ and any ring R .
- (6) Let R be a ring and α an ideal in R . Show that if $\mathbf{J}(R/\alpha) = 0$, then $\mathbf{J}(R) \subseteq \alpha$.
- (7) Show that every finitely generated module over an Artinian ring has finite composition length.
- (8) Let A be an algebra over a field k . Show that any element of $\mathbf{J}(A)$ algebraic over k is nilpotent. Deduce that if A has countable dimension over an uncountable field, then $\mathbf{J}(A)$ is a nilideal. (*Hint.* Show that if a is transcendental over k and $\lambda_i \in k$ are distinct, then the elements $(a - \lambda_i)^{-1}$ are linearly independent over k .)
- (9) Verify that the ring $k[[x]]$ of formal power series in x over a field k is an integral domain with non-zero Jacobson radical.

5.5 The tensor product of algebras

Let K be a commutative ring. We recall that a K -algebra is a K -module A with a bilinear mapping $\mu_0: A \times A \rightarrow A$, the multiplication in A . By Th. 4.7.1 it comes to the same thing to have a linear mapping

$$\mu: A \otimes A \rightarrow A, \quad (1)$$

and we shall also refer to μ in (1) as the *multiplication* in A . Explicitly we have $(\sum x_i \otimes y_i)\mu = \sum x_i y_i$ for $x_i, y_i \in A$. The associativity of A can be expressed by the commutativity of the diagram

$$\begin{array}{ccc} A \otimes A \otimes A & \xrightarrow{\mu \otimes 1} & A \otimes A \\ \downarrow 1 \otimes \mu & & \downarrow \mu \\ A \otimes A & \xrightarrow{\mu} & A \end{array}$$

and the existence of a unit element e of A is expressed by the commutativity of the diagram

$$\begin{array}{ccccc} & & A \otimes A & & \\ & \nearrow \otimes e & \downarrow \mu & \swarrow e \otimes & \\ A & \xrightarrow{1} & A & \xleftarrow{1} & A \end{array}$$

Further, if the multiplication is commutative, this is expressed by the commutativity of the diagram

$$\begin{array}{ccc} A \otimes A & \xrightarrow{\tau} & A \otimes A \\ \searrow \mu & & \swarrow \mu \\ & A & \end{array}$$

where τ is the transposition

$$\tau: x \otimes y \mapsto y \otimes x. \quad (2)$$

Given two K -algebras A, B , we can define their tensor product as a K -algebra in a natural fashion. Let $\tau: B \otimes A \rightarrow A \otimes B$ be defined as in (2), for $x \in B, y \in A$; this gives rise to the permutation map $\tau_1 = 1 \otimes \tau \otimes 1: A \otimes B \otimes A \otimes B \rightarrow A \otimes A \otimes B \otimes B$, where

$$(a_1 \otimes b_1 \otimes a_2 \otimes b_2) \tau_1 = a_1 \otimes a_2 \otimes b_1 \otimes b_2.$$

If we combine this with the multiplications μ of A and ν of B , we obtain a linear mapping

$$\pi = \tau_1(\mu \otimes \nu): A \otimes B \otimes A \otimes B \rightarrow A \otimes B. \quad (3)$$

We claim that this multiplication is associative whenever μ and ν are. Put $C = A \otimes B$; then (3) can be written $\pi: C \otimes C \rightarrow C$ and for any $a_i \in A, b_i \in B$ we have

$$\begin{aligned} (a_1 \otimes b_1 \otimes a_2 \otimes b_2 \otimes a_3 \otimes b_3)(\pi \otimes 1)\pi &= (a_1 a_2 \otimes b_1 b_2 \otimes a_3 \otimes b_3)\pi \\ &= (a_1 a_2)a_3 \otimes (b_1 b_2)b_3. \end{aligned}$$

Applying $(1 \otimes \pi)\pi$, we obtain $a_1(a_2 a_3) \otimes b_1(b_2 b_3)$, which is the same, by the associativity in A and B . Since the elements on the left span $C \otimes C \otimes C$, the associativity of π follows. In a similar way we can show that C is commutative whenever A and B are. When A has a unit element e , then the mapping

$$b \mapsto e \otimes b \quad (4)$$

is a homomorphism from B to $A \otimes B$; likewise for a one in B , and if A, B have ones e, f say, then $e \otimes f$ is a one for C . We sum up these results in

THEOREM 5.1 *Let K be a commutative ring and A, B any K -algebras. Then $A \otimes B$ is again a K -algebra, which is associative, commutative or has a unit element whenever this is so for A and B . ■*

In what follows we shall assume that our algebras are associative and have a one. Even when A has a one e , (4) need not be an embedding (cf. Ex. (5)), but this is so if K is a field and $A \neq 0$, for in that case A has a basis including e and the subspace spanned by e is a direct summand in A , so we can apply Prop. 4.7.3.

Examples

1. Let K be a field and E an extension field of K . If A is any K -algebra, then $A \otimes E$ is an algebra over E , of dimension $[A:K]$. Explicitly, if u_1, \dots, u_n is a K -basis of A , then the elements $u_1 \otimes 1, \dots, u_n \otimes 1$ form a basis of $A \otimes E$ over E . Regarded as an E -algebra, $A \otimes E$ is denoted by A_E and is called the algebra obtained from A by extension of the ground field to E .

2. Let $A = \mathfrak{M}_r(K)$ be a full matrix ring over K . Then A has a basis e_{ij} ($i, j = 1, \dots, r$) over K (the ‘matrix units’), with the multiplication rule

$$e_{ij}e_{kl} = \delta_{jk}e_{il}.$$

For any K -algebra B , the tensor product $A \otimes B$ is a free B -module with the same basis as A , hence

$$\mathfrak{M}_r(K) \otimes B \cong \mathfrak{M}_r(B).$$

In particular, if $B = \mathfrak{M}_s(K)$, then the right-hand side becomes $\mathfrak{M}_r(\mathfrak{M}_s(K)) \cong \mathfrak{M}_{rs}(K)$, for the elements of $\mathfrak{M}_r(\mathfrak{M}_s(K))$ are $r \times r$ matrices whose entries are $s \times s$ matrices over K . Thus we find

$$\mathfrak{M}_r(K) \otimes \mathfrak{M}_s(K) \cong \mathfrak{M}_{rs}(K). \quad (5)$$

As a module isomorphism this already follows from Prop. 4.7.5.

There is a simple criterion for an algebra to be a tensor product which is often useful. Let C be an algebra over a field k , and let U, V be subspaces of C ; then U and V are said to be *linearly disjoint* over K if for any linearly independent elements u_i in U and v_j in V , the elements $u_i v_j$ in C are linearly independent over k . Clearly this just means that the natural mapping

$$U \otimes V \rightarrow C$$

induced by the mapping $(u, v) \mapsto uv$ is injective. Now the criterion can be stated as follows:

PROPOSITION 5.2 *Let C be an algebra over a field k . Given subalgebras A, B of C , if (i) A and B are linearly disjoint, (ii) $AB = C$ and (iii) A and B commute elementwise, then $C \cong A \otimes B$.*

Proof. The mapping $(x, y) \mapsto xy$ from $A \times B$ to C is bilinear and so induces a k -linear mapping from $A \otimes B$ to C . This is injective by (i), surjective by (ii) and a homomorphism by (iii). ■

Next we shall describe centralizers in a tensor product. In any ring R the *centralizer* of a subset X is defined as the set

$$\mathbf{C}(X) = \{y \in R \mid xy = yx \text{ for all } x \in X\}.$$

As we have seen in 4.3, $\mathbf{C}(X)$ is a subring of R ; when R is a K -algebra, $\mathbf{C}(X)$ is again a K -algebra.

We shall also need a formula for intersections in a tensor product. Let U, V be any K -modules (where K is a commutative ring) and let U', V' be direct summands in U, V respectively, so that $U' \otimes V, U \otimes V', U' \otimes V'$ may be regarded as submodules of $U \otimes V$, by Prop. 4.7.3. Then

$$U' \otimes V \cap U \otimes V' = U' \otimes V'. \quad (6)$$

For we can write $U = U' \oplus U'', V = V' \oplus V''$; hence $U \otimes V = U' \otimes V' \oplus U' \otimes V'' \oplus U'' \otimes V' \oplus U'' \otimes V'' = W_1 \oplus W_2 \oplus W_3 \oplus W_4$ say. Now the left-hand side of (6) is $(W_1 \oplus W_2) \cap (W_1 \oplus W_3)$ and this is clearly W_1 , so (6) holds.

PROPOSITION 5.3 *Let A_1, A_2 be algebras over a field k , let B_i be a subalgebra of A_i and B'_i the centralizer of B_i in A_i . Then the centralizer of $B_1 \otimes B_2$ in $A_1 \otimes A_2$ is $B'_1 \otimes B'_2$.*

Proof. Let us denote by C the centralizer of $B_1 \otimes B_2$ in $A_1 \otimes A_2$. It is clear that

$$B'_1 \otimes B'_2 \subseteq C, \quad (7)$$

where we have identified $B'_1 \otimes B'_2$ with its image in $A_1 \otimes A_2$. By the above remark this is justified because k is a field.

It remains to prove equality in (7). Let $\{v_i\}$ be a basis of A_2 . Then every element of $A_1 \otimes A_2$ can be uniquely written in the form $\sum a_i \otimes v_i$, where $a_i \in A_1$. Given any $b \in B_1$, we have

$$(\sum a_i \otimes v_i)(b \otimes 1) - (b \otimes 1)(\sum a_i \otimes v_i) = \sum (a_i b - b a_i) \otimes v_i. \quad (8)$$

If $\sum a_i \otimes v_i \in C$, then the left-hand side of (8) is 0, hence $a_i b = b a_i$, and this holds for all $b \in B_1$; therefore $a_i \in B'_1$, and so $\sum a_i \otimes v_i \in B'_1 \otimes A_2$. Thus $C \subseteq B'_1 \otimes A_2$ and similarly $C \subseteq A_1 \otimes B'_2$; it follows that

$$C \subseteq (A_1 \otimes B'_2) \cap (B'_1 \otimes A_2) = B'_1 \otimes B'_2,$$

by (6) above. Together with (7) this shows that $C = B'_1 \otimes B'_2$ as claimed. ■

If we take $B_i = A_i$, B'_i is the centre of A_i and we obtain

COROLLARY 5.4 If A_1, A_2 are algebras over a field with centres Z_1, Z_2 respectively, then the centre of $A_1 \otimes A_2$ is $Z_1 \otimes Z_2$. ■

Exercises

- (1) Show that if A, B are finitely generated K -algebras, then so is $A \otimes B$.
- (2) Let A be a K -algebra and x an indeterminate over A . Show that $A[x] \cong A \otimes K[x]$.
- (3) Let $\mathbf{Z}[i]$ be the ring of Gaussian integers (obtained by adjoining a root i of $x^2 + 1 = 0$ to \mathbf{Z}). Show that $\mathbf{R} \otimes_{\mathbf{Z}} \mathbf{Z}[i] = \mathbf{C}$.
- (4) Let B be an algebra with unit element 1. Verify that the map from A to $A \otimes B$ defined by $x \mapsto x \otimes 1$ is a homomorphism.
- (5) Let $A = \mathbf{Z}/2$, as \mathbf{Z} -algebra and $B = \mathbf{Z}$. Verify that the natural homomorphism $B \rightarrow A \otimes B$ is not an embedding.
- (6) Fill in the details in the proof of Th. 5.1.
- (7) Let A, B be algebras with unit elements $\neq 0$ over a field, but not assumed to be associative. Show that if $A \otimes B$ is associative, then so are A and B .
- (8) Let A be a k -algebra, where k is a field, and let E be an extension field of k . Show that $A \otimes E$ may be defined by taking a multiplication table of A in terms of a basis over k and regarding it as a multiplication table over E .
- (9) Let A, B be K -algebras. Given a right A -module U and a left $(A \otimes B)$ -module V , show that $(U \otimes_K B) \otimes_{A \otimes B} V \cong U \otimes_A V$. (Hint. Use (7) of 4.7 and the associative law.)

5.6 The regular representation; norm and trace

Let A be a finite-dimensional k -algebra. We have seen in 5.1 that A has a faithful matrix representation over k , the regular representation ρ . If u_1, \dots, u_n is a basis of A , then the regular representation $a \mapsto \rho(a) = (\rho_{ij}(a))$ is given by the equations

$$u_i a = \sum \rho_{ij}(a) u_j.$$

The matrix of the regular representation is still dependent on the choice of basis in A , but we get an invariant in A by taking the characteristic polynomial

$$\det(xI - \rho(a)) = x^n + \lambda_1 x^{n-1} + \cdots + \lambda_n. \quad (1)$$

Here the first and last coefficients are just the trace and the determinant of $\rho(a)$, apart from sign. We define the *trace* of a , $T_{A/k}(a)$, and the *norm* of a , $N_{A/k}(a)$, as

$$T_{A/k}(a) = -\lambda_1 = \text{tr}(\rho(a)),$$

$$N_{A/k}(a) = (-1)^n \lambda_n = \det(\rho(a)).$$

More briefly we often write $T(a)$, $N(a)$ when no confusion is possible. The properties of the trace and determinant of a matrix lead immediately to the formulae:

$$\begin{aligned} T(a+b) &= T(a) + T(b), & T(\lambda a) &= \lambda T(a), \\ T(ab) &= T(ba), & T(1) &= n, & a, b \in A, \lambda \in k. \\ N(ab) &= N(a)N(b), & N(\lambda a) &= \lambda^n N(a), & N(1) &= 1. \end{aligned} \quad (2)$$

As a first application we obtain a criterion for zerodivisors in A in terms of the norm:

PROPOSITION 6.1 *Let A be a finite-dimensional algebra over a field k and $c \in A$. Then the following conditions are equivalent:*

- (a) c is a non-zerodivisor in A ,
- (b) c is a unit in A ,
- (c) $N(c) \neq 0$.

In particular every finite-dimensional algebra without zerodivisors is a division algebra.

Proof. (a) \Rightarrow (b). If $[A:k] = n$, then the elements $1, c, c^2, \dots, c^n$ are linearly dependent over k , so we have an equation

$$\gamma_0 c^r + \gamma_1 c^{r-1} + \cdots + \gamma_r = 0, \quad \gamma_i \in k, \text{ not all } 0.$$

Choose an equation of least degree r ; then since c is a non-zerodivisor, $\gamma_r \neq 0$, and on dividing by it, we may assume that $\gamma_r = 1$. Then $-(\gamma_0 c^{r-1} + \cdots + \gamma_{r-1})$ is an inverse for c in A .

(b) \Rightarrow (c). If c is a unit, then $1 = N(cc^{-1}) = N(c)N(c^{-1})$, hence $N(c) \neq 0$.

(c) \Rightarrow (a). If $N(c) \neq 0$, then the matrix $\rho(c)$ of the regular representation is non-singular, hence the mapping $x \mapsto xc$ is injective, i.e. c is a non-zerodivisor. ■

To establish the usual transitivity formulae we shall need a lemma.

LEMMA 6.2 *Let F/k be a field extension of degree r and denote by ρ the regular representation of F in k ; thus $\rho: F \rightarrow \mathfrak{M}_r(k)$ is a homomorphism. Denote by ρ_n the induced homomorphism $\mathfrak{M}_n(F) \rightarrow \mathfrak{M}_{nr}(k)$. Then for any $C \in \mathfrak{M}_n(F)$,*

$$\text{tr}(\rho_n(C)) = \text{tr}(\rho(\text{tr}(C))), \quad (3)$$

$$\det(\rho_n(C)) = \det(\rho(\det(C))). \quad (4)$$

Proof. By definition $\rho(c)$ is an $r \times r$ matrix over k ; its entries will be written $\rho_{\lambda\mu}(c)$. Writing $C = (c_{ij})$, we have on the left of (3), $\sum \rho_{\lambda\lambda}(c_{ii})$, and on the right $\sum \rho_{\lambda\lambda}(\sum c_{ii})$, and these two expressions are equal, by the linearity of the functions $\rho_{\lambda\lambda}$.

Next consider (4). Since both sides are unchanged on replacing C by $P^{-1}CP$, we may, on passing to an algebraic closure of F , replace C by a triangular matrix. Now $\rho_n(C)$ is a block triangular matrix, with $r \times r$ blocks $\rho(c_{ii})$ on the main

diagonal and zeros below it; hence on taking a Laplace expansion by the first r rows and using induction on n , we obtain $\det \rho(c_{11}) \cdot \det \rho(c_{22}) \dots \det \rho(c_{nn})$ on the left of (4). On the right we have $\det \rho(c_{11}c_{22}\dots c_{nn})$, and these results agree, because ρ is a homomorphism. Thus (4) holds over an algebraic closure of F , and hence over F itself. ■

Let F/k be a field extension of degree r and A an F -algebra of dimension n . Then we may regard A as a k -algebra of dimension rn , by Prop. 3.1.2. We can therefore define norm and trace for A as an F -algebra and as a k -algebra; they are related by the

TRANSITIVITY FORMULAE

$$T_{F/k}(T_{A/F}(c)) = T_{A/k}(c), \quad (5)$$

$$N_{F/k}(N_{A/F}(c)) = N_{A/k}(c). \quad (6)$$

Proof. If $\rho_{A/F}$ denotes the regular representation of A over F , regarded as a homomorphism $A \rightarrow \mathfrak{M}_n(F)$, then we clearly have

$$\rho_{A/k}(c) = \rho_{F/k}(\rho_{A/F}(c)). \quad (7)$$

In (3) let us take $C = \rho_{A/F}(c)$; the left-hand side can by (7) be written $\text{tr}(\rho_{A/k}(c))$ and this is just $T_{A/k}(c)$. The right-hand side is $\text{tr} \rho_{F/k}(T_{A/F}(c)) = T_{F/k}(T_{A/F}(c))$, and so (5) is established. Now (6) follows in the same way from (4), putting again $C = \rho_{A/F}(c)$. ■

As an example let us calculate the norm and trace in a field extension. Let E/k be a field extension of degree r and take $c \in E$ of degree s over k . The minimal equation for c has the form

$$x^s + \lambda_1 x^{s-1} + \dots + \lambda_s = 0. \quad (8)$$

The extension $k(c)/k$ is of degree s , and the matrix $\rho(c)$ of c in the regular representation satisfies (8), hence a comparison of degrees shows that the minimal polynomial is also the characteristic polynomial of $\rho(c)$. More explicitly we can see this by taking the basis $1, c, \dots, c^{s-1}$ of $k(c)/k$. In terms of this basis we have

$$c^i \cdot c = \begin{cases} c^{i+1} & \text{if } i < s-1, \\ -\lambda_1 c^{s-1} - \dots - \lambda_s & \text{if } i = s-1. \end{cases}$$

Hence the matrix for c is

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -\lambda_s & -\lambda_{s-1} & -\lambda_{s-2} & \dots & -\lambda_2 & -\lambda_1 \end{pmatrix}.$$

This is just the companion matrix of the minimal polynomial. Its trace is $-\lambda_1$ and its determinant is $(-1)^s \lambda_s$. Since $[E:k(c)] = r/s$, we have, by transitivity,

$$T_{E/k}(c) = \frac{r}{s}(-\lambda_1), \quad N_{E/k}(c) = [(-1)^s \lambda_s]^{r/s}.$$

These formulae show that for a separable field extension the definitions of norm and trace given here coincide with those given in 3.9.

More generally, given any representation of a k -algebra, $\sigma: A \rightarrow \mathfrak{M}_r(k)$, we can define its trace and norm as

$$\text{Tr}_\sigma(x) = \text{Tr}(\sigma(x)), \quad \text{Nm}_\sigma(x) = \det(\sigma(x)),$$

and they will again satisfy laws corresponding to (2). The trace defines a quadratic form $\text{Tr}_\sigma(x^2)$ on A , which is non-singular iff

$$\det(\text{Tr}_\sigma(u_i u_j)) \neq 0 \tag{9}$$

for a basis u_1, \dots, u_n of A . Such a representation will sometimes allow us to recognize when the algebra is semisimple.

PROPOSITION 6.3 *Let σ be a representation of an algebra A over k . If the quadratic form Tr_σ is non-singular, then A is semisimple.*

Proof. We first note that if an element c of A is nilpotent, then so is $\sigma(c)$, hence its trace is zero, thus $\text{Tr}_\sigma(c) = 0$. Suppose that A is not semisimple; then its radical is non-zero and if a basis v_1, \dots, v_n of A is chosen so that v_1, \dots, v_r is a basis of $\mathbf{J}(A)(r \geq 1)$, then $v_i v_i$ is nilpotent for all i ; hence the first row of the matrix $\text{Tr}_\sigma(v_1 v_i)$ is zero, and so the form Tr_σ is singular. ■

This sufficient condition for semisimplicity is also necessary for a faithful representation in characteristic 0 (cf. Ex. (4)).

In any commutative algebra the sum and product of two nilpotent elements is again nilpotent, but neither of these assertions holds in the general case. In these circumstances it is somewhat surprising that the nilpotence of an algebra follows from the existence of a nilpotent basis. This is the content of the next result, due to Wedderburn in 1937. Note that, although the proof depends on a trace argument, it holds in any characteristic.

THEOREM 6.4 *Let A be a finite-dimensional algebra over a field k and B a subalgebra with a basis consisting of nilpotent elements. Then B is nilpotent.*

Proof. We remark that the conclusion shows that B cannot contain 1, but this is not obvious at the outset. Let C be the subalgebra of A generated by B and 1; thus if u_1, \dots, u_n is the nilpotent basis of B , then C is spanned by 1, u_1, \dots, u_n . To prove that B is nilpotent it will be enough to show that B is contained in every maximal ideal of C , for then $B \subseteq \mathbf{J}(C)$ and the conclusion follows by the nilpotence of $\mathbf{J}(C)$.

(Th. 4.5). Suppose then that there is a maximal ideal m of C not containing B . Then $m + B = C$ and $C/m \cong B/(B \cap m)$. If $\bar{C} = C/m$ and \bar{x} denotes the residue class of x (mod m), then the simple algebra \bar{C} is spanned by $\bar{u}_1, \dots, \bar{u}_n$ and for suitable renumbering, $\bar{u}_1, \dots, \bar{u}_r$, is a basis. Each \bar{u}_i is again nilpotent, so we have a simple algebra with a nilpotent basis; we shall show that this leads to a contradiction. Let E be an algebraically closed field containing k ; then \bar{C}_E is a simple E -algebra, hence isomorphic to $\mathfrak{M}_t(E)$ for some t . Let Tr be the trace function on the matrix ring $\mathfrak{M}_t(E)$; for any nilpotent element u , $\text{Tr}(u) = 0$, and since there is a basis of nilpotent elements, $\text{Tr}(x) = 0$ for all $x \in \mathfrak{M}_t(E)$, by linearity. But $\text{Tr}(e_{11}) \neq 0$, and so we have reached a contradiction. Hence B is nilpotent. ■

We remark that the regular representation on $\mathfrak{M}_t(E)$ would give $T(e_{11}) = t$, and so cannot be used in finite characteristic. In fact the trace function used here is the ‘reduced trace’ which we shall meet again in Vol. 3.

Exercises

- (1) Show that the norm and trace of the full matrix algebra $\mathfrak{M}_n(k)$ are given by $N(A) = (\det A)^n$, $T(A) = n \cdot \text{tr}(A)$. Find the norm and trace for the algebra $\mathfrak{T}_n(k)$ of upper triangular matrices.
- (2) Find the norm and trace for the group algebra of C_n , the cyclic group of order n . Likewise for the Klein 4-group $V = \langle a, b \mid a^2 = b^2 = (ab)^2 = 1 \rangle$.
- (3) Show that in characteristic 0 a matrix A is nilpotent iff $\text{tr}(A^r) = 0$ for $r = 1, 2, \dots$
- (4) Let σ be a faithful representation of a k -algebra A , where $\text{char } k = 0$ (e.g. the regular representation). Show that the set $\mathfrak{n} = \{x \in A \mid \text{Tr}_\sigma(xa) = 0 \text{ for all } a \in A\}$ is a nilpotent ideal of A . Deduce that if A is semisimple, then Tr_σ is non-singular.
- (5) Let F/k be an inseparable field extension of characteristic p . Given $c \in F$, let c^q , where $q = p^r$, be the least power of c which is separable over k and write $E = k(c^q)$, $[F:E] = s$. Show that $T_{F/k}(c) = sT_{E/k}(c)$, $N_{F/k}(c) = N_{E/k}(c)^s$. (This exercise defines norm and trace for arbitrary finite field extensions in terms of the special case given in 3.9.)
- (6) Show that for a purely inseparable field extension of degree $r > 1$, $N(c) = c^r$, $T(c) = 0$.
- (7) Show that a field extension F/k of finite degree is separable iff there is an element c in F such that $T(c) \neq 0$.

5.7 Composites of fields

We shall apply the results found in 5.4 to study the composites of two fields. Given two fields with a common subfield k , we can always find a common extension field of E and F : the tensor product $R = E \otimes_k F$ is a non-zero commutative algebra,

and if \mathfrak{m} is a maximal ideal, then $R/\mathfrak{m} = K$ is a field, which has homomorphisms of E, F into K , given by $x \mapsto \bar{x} \otimes \bar{1}$, $y \mapsto \bar{1} \otimes \bar{y}$, where $x \in E$, $y \in F$ and the bar denotes the residue class mod \mathfrak{m} . Like every homomorphism between fields, these mappings are embeddings, so E and F have been embedded in K . At this point three questions arise. Firstly we shall want to know how far K is determined by E and F ; secondly we may ask under what conditions $E \otimes F$ is itself a field or at least an integral domain; and thirdly there is the question of the role played by the common subfield k .

The third point is easily dealt with. If two fields E and F are to be subfields of a third, they must have the same characteristic, hence their prime subfields are isomorphic and so may be identified. This shows that it represents no loss of generality to assume that E and F have a common subfield. The second question will be answered in Vol. 3 and we now turn to the first question, concerning the different possible composites; we have to begin by defining this term.

Let E, F be two extension fields of k . By a *composite* of E and F over k we understand a field extension K/k with k -homomorphisms $E \rightarrow K, F \rightarrow K$ such that K is generated (as a field) by the images of E and F . Two composites K and K' are said to be *equivalent* over k if there is a k -isomorphism $f:K \rightarrow K'$ such that the diagram shown commutes.

$$\begin{array}{ccccc} & & E & & \\ & \swarrow & & \searrow & \\ K & & f & & K' \\ & \uparrow & & \downarrow & \\ & & F & & \end{array}$$

Our first result gives a survey of the different composites. Given an integral domain R , we shall write $\mathcal{F}(R)$ for its field of fractions. An ideal \mathfrak{a} in a commutative ring R is said to be *prime* if R/\mathfrak{a} is an integral domain (cf. 9.2). Since an integral domain is non-trivial, this means in particular that a prime ideal must be proper.

THEOREM 7.1 *Let K be any field and E, F two extension fields of k . Then there exists a composite of E and F over k . Moreover, the equivalence classes of composites of E and F over k correspond to the different prime ideals in $R = E \otimes_k F$.*

Proof. We have already seen that composites arise as homomorphic images of R by maximal ideals, which exist by Krull's theorem (Th. 2.2.11). Now let a composite of E and F be given, say $\alpha:E \rightarrow K, \beta:F \rightarrow K$. Then there is a homomorphism $\gamma:R \rightarrow K$, given by $\gamma = \alpha \otimes \beta$. If $\mathfrak{p} = \ker \gamma$, then R/\mathfrak{p} is embedded in K , hence an integral domain, so \mathfrak{p} is a prime ideal in R , and from the definition of K we see that $K = \mathcal{F}(R/\mathfrak{p})$. Conversely, every prime ideal \mathfrak{p} of R gives rise to a composite $\mathcal{F}(R/\mathfrak{p})$ of E and F in this way. For $\mathcal{F}(R/\mathfrak{p})$ is a field and we have k -

homomorphisms from E and F to $\mathcal{F}(R/\mathfrak{p})$ by combining the inclusion in R with the residue class mapping. Moreover, equivalent composites clearly have the same kernel and, conversely, the kernel determines the composite up to equivalence. ■

If one of E, F is finite-dimensional over k , the result can still be simplified. We shall need a result on general commutative algebras:

LEMMA 7.2 *Any finite-dimensional non-trivial commutative k -algebra A has only finitely many prime ideals, all maximal, $\mathfrak{m}_1, \dots, \mathfrak{m}_r$, say; the quotients $A/\mathfrak{m}_i = K_i$ are extension fields of k and if $\mathfrak{N} = \mathfrak{m}_1 \cap \dots \cap \mathfrak{m}_r$ is the radical of A , then*

$$A/\mathfrak{N} \cong K_1 \times \dots \times K_r. \quad (1)$$

Proof. Any finite-dimensional k -algebra which is an integral domain must be a field (e.g. by Prop. 6.1), therefore any prime ideal in A is maximal. The radical \mathfrak{N} of A is the intersection of all maximal ideals and is nilpotent; further, A/\mathfrak{N} is semisimple and so is a direct product of a finite number of fields, by Wedderburn's theorem (Th. 3.4), so we obtain a representation (1). The homomorphism from A to the right-hand side of (1) combined with the projection on K_i is a homomorphism onto K_i with kernel \mathfrak{m}_i , a maximal ideal, and $\mathfrak{N} = \mathfrak{m}_1 \cap \dots \cap \mathfrak{m}_r$ is the kernel of the homomorphism to the product on the right of (1). It remains to show that there are no other prime ideals. Let \mathfrak{m} be any prime ideal of A ; then A/\mathfrak{m} is an integral domain, hence a field, and the natural mapping $A \rightarrow A/\mathfrak{m}$ induces on each K_i either an isomorphism or the zero mapping.

$$\begin{array}{ccc} & & A/\mathfrak{m} \\ & \swarrow & \downarrow \cong \\ A & & \searrow \\ & & A/\mathfrak{m}_i \end{array}$$

It must be an isomorphism for at least one i ; thus we have the commutative triangle shown and it follows that the kernels of the residue class mappings agree; hence $\mathfrak{m} = \mathfrak{m}_i$. Thus $\mathfrak{m}_1, \dots, \mathfrak{m}_r$ are the only prime ideals of A . ■

If $[F:k] < \infty$, then $E \otimes_k F$ is a finite-dimensional E -algebra and we can apply Lemma 7.2 with E for k :

PROPOSITION 7.3 *Let E, F be fields over k , one of which, say F , is finite-dimensional over k . Then the number of composites of E and F is finite, say K_1, \dots, K_r ($r \geq 1$), and*

$$(E \otimes_k F)/\mathfrak{N} \cong K_1 \times \dots \times K_r, \quad (2)$$

where \mathfrak{N} is the radical of $E \otimes_k F$. In particular, E, F have a single composite precisely if $E \otimes F$ has a prime ideal consisting of nilpotent elements. ■

Specializing further, let us take F/k to be a simple extension, say $F = k(a)$, where a has minimal polynomial f over k . Then $F = k[x]/(f)$, so we have an exact sequence

$$0 \rightarrow (f) \rightarrow k[x] \rightarrow F \rightarrow 0.$$

Tensoring with E we obtain

$$0 \rightarrow (f) \rightarrow E[x] \rightarrow E \otimes F \rightarrow 0,$$

and so $E \otimes F \cong E[x]/(f)$. Let the decomposition of f into irreducible factors over E be $f = p_1^{\alpha_1} \dots p_r^{\alpha_r}$; then the maximal ideals in $E \otimes F$ are (p_i) , $i = 1, \dots, r$, and their intersection is $\mathfrak{N} = (p_1 \dots p_r)/(f)$. In particular, if F/k is separable, then all the α 's are 1 and $\mathfrak{N} = 0$. In that case (2) shows that

$$[E \otimes F : E] = \sum_i [K_i : E].$$

Now $[E \otimes F : E] = [F : k]$, hence we obtain

COROLLARY 7.4 *If F/k is a finite separable field extension, E/k is an arbitrary field extension and K_1, \dots, K_r are the inequivalent composites of E and F over k , then the K_i correspond to the irreducible factors over E of the minimal polynomial for a generator of F over k , and*

$$E \otimes_k F \cong K_1 \times \dots \times K_r;$$

moreover, $[F : k] = \sum_i [K_i : E]$.

This follows from what has been said, if we remember that any finite separable field extension is simple (3.9). ■

When F is a Galois extension and $E = F$, then the minimal polynomial for a generator of F over k splits into linear factors over F (because F is normal over k); hence we obtain

COROLLARY 7.5 *If F/k is a finite Galois extension of degree n , then*

$$F \otimes_k F \cong F_1 \times \dots \times F_n, \quad \text{where } F_i \cong F. \quad \blacksquare$$

Exercises

- (1) Show that if $E \otimes_k F$ is a field, then either E or F must be algebraic over k .
- (2) Show that two fields can be embedded in the same field iff they have the same characteristic.

- (3) Let E, F be finite separable field extensions of k , with generators having minimal polynomials f, g respectively. Show that f splits into as many irreducible factors over F as g does over E . Deduce that for any irreducible separable polynomials f, g over k , f remains irreducible over $k[x]/(g)$ iff g remains irreducible over $k[x]/(f)$. What happens when f, g are no longer separable?
- (4) Let F/k be a finite separable field extension of degree > 1 , let f be the minimal polynomial of a generator of F and let E be a minimal splitting field of f over k . Suppose further that the Galois group of f over k is two-fold transitive. Show that E and F have composites that are k -isomorphic as fields, but inequivalent.
- (5) Let F/k be a finite separable field extension of degree > 1 and suppose that the minimal polynomial for a generator of F has a two-fold transitive Galois group. Show that $F \otimes F$ is the direct product of F and another field. Illustrate this by considering $x^3 - 2$ over \mathbf{Q} .
- (6) Show that E/k is purely inseparable and F/k arbitrary, then $E \otimes F$ is primary. Deduce that there is just one composite of E and F over k . If moreover, F/k is separable, then E, F are linearly disjoint in this composite.

Further exercises on Chapter 5

- (1) Let A be an n -dimensional algebra (without 1) over a field k . Show that A has a faithful representation in $\mathfrak{M}_{n+1}(k)$ and give an example where $n+1$ cannot be replaced by n . (Hint. Take A to be nilpotent.)
- (2) Show that there are no division algebras over \mathbf{C} , other than \mathbf{C} itself.
- (3) Find all ideals in $\mathfrak{M}_n(\mathbf{Z})$, for varying n .
- (4) Let K be any commutative ring and $u \in K^n, v \in K^n$. Show that the set of all $n \times n$ matrices A such that $uA = 0 = Av$ is a subalgebra of K^n (possibly without 1). Deduce that the set of ‘semimagic squares’, i.e. matrices whose row sums and column sums are zero, is an algebra. Is this true of the set of all ‘magic squares’ (in which the sums along the two diagonals are also 0)?
- (5) Show that a finite commutative ring with n units has at most $(n+1)^2$ elements.
- (6) Show that a ring with p^3 elements (p prime) is either commutative or isomorphic to $\mathfrak{T}_2(\mathbf{F}_p)$.
- (7) Show that if an element a in a ring R has two right inverses a', a'' , then $a'' + a'a - 1$ is another right inverse. Deduce that any element with more than one right inverse has infinitely many.
- (8) Show that a ring R such that $nR = 0$ for a square-free integer n is a direct product of algebras over fields.

- (9) Show that $\begin{pmatrix} \mathbf{Q} & \mathbf{Q} \\ 0 & \mathbf{Z} \end{pmatrix}$ is left but not right Noetherian.
- (10) Let A be a k -algebra, where k is a field of prime characteristic p , and define a trace function on A as a linear function $f:A \rightarrow k$ such that $f(a^p) = f(a)^p$. Show that any trace function vanishes on the radical.
- (11) Show that if R_i is a ring with centre C_i , then $\prod R_i$ has centre $\prod C_i$.
- (12) Let A be a commutative algebra (not necessarily with 1) and define its *duplicate* as $A' = A \otimes A/D$, where $D = \{x \otimes y - y \otimes x \mid x, y \in A\}$. Show that $A' = A^2 \oplus C$, where C is isomorphic to the annihilator of A ; in particular for algebras with 1, $A' \cong A$. What happens if this definition is applied to a non-commutative algebra?
- (13) Show that the tensor product of augmented algebras is again augmented.
- (14) Show that an algebra (possibly without 1) is augmented iff it has an ideal of codimension 1 not containing A^2 .
- (15) A vector $a = (a_1, \dots, a_n) \in R^n$ is called *stochastic* if $0 \leq a_i \leq 1$ and $\sum a_i = 1$. A matrix is *stochastic* if its rows are stochastic and an algebra A is *stochastic* if it has a basis (called a *natural basis*) u_i such that in the regular representation $\rho(u_i)$ is stochastic. Show that a basis is natural iff it is obtained from a natural basis by transformation by a stochastic matrix. Show that an algebra is stochastic iff it is augmented and the convex hull of the elements satisfying $a(x) = 1$ is a simplex closed under multiplication.
- (16) Show that an algebra A is semisimple provided that the trace with respect to some representation σ satisfies $\det(\text{Tr}_\sigma(u_i v_j)) \neq 0$, where $\{u_i\}, \{v_j\}$ are any two bases of A . By taking as bases for the group algebra kG of a finite group G the elements u_1, \dots, u_n of G and $u_1^{-1}, \dots, u_n^{-1}$, show that kG is semisimple whenever $\text{char } k \nmid |G|$ (Maschke's theorem, cf. 7.2).

6

Quadratic forms and ordered fields

In Ch. 8 of Vol. 1 we met inner product spaces and obtained a classification of such spaces over the real and complex numbers. We now take up the subject again in a somewhat more general context and look at the properties of quadratic forms over a general field, and its group of isometries. We also briefly discuss the related topic of ordered fields, leading to a construction of the real numbers, but the more detailed study of quadratic forms, in particular their invariants, the Clifford algebra of a form and the Witt ring of a field, will be left to Vol. 3.

6.1 Inner product spaces

Let k be a field. By an *inner product space* we understand a pair (V, b) consisting of a finite-dimensional vector space V over k and a symmetric bilinear form b on V , i.e. a mapping

$$b: V \times V \rightarrow k,$$

which is *symmetric*: $b(x, y) = b(y, x)$; and *bilinear*:

$$b(\alpha x + \beta y, z) = \alpha b(x, z) + \beta b(y, z), \quad b(x, \alpha y + \beta z) = \alpha b(x, y) + \beta b(x, z).$$

There is a second way of defining inner product spaces, in terms of quadratic forms. By a *quadratic form* on a given vector space V over k we understand a mapping $q: V \rightarrow k$ such that

$$\mathbf{Q.1} \quad q(\alpha x) = \alpha^2 q(x) \quad \text{for all } x \in V \text{ and } \alpha \in k,$$

$$\mathbf{Q.2} \quad q(x + y) - q(x) - q(y) = b_q(x, y) \quad \text{is a bilinear form on } V.$$

Clearly the function b_q in Q.2 is symmetric; it is called the *associated bilinear form*, and it can be used to define an inner product structure on V . If $\text{char } k \neq 2$, then every symmetric bilinear form b is associated with a quadratic form, namely $b = b_q$, where $q = \frac{1}{2}b(x, x)$. This follows because any bilinear form b on V satisfies

$$b(x + y, x + y) - b(x, x) - b(y, y) = b(x, y) + b(y, x).$$

When b is symmetric, this reduces to $2b(x, y)$, so in that case $b_q(x, y) = q(x + y) - q(x) - q(y) = b(x, y)$. So in characteristic not 2 there is complete equivalence

between quadratic forms and symmetric bilinear forms, and we shall use both interchangeably. To be precise, we shall use $b(x, y)$ and $q(x) = b(x, x)$, so that

$$b(x, y) = \frac{1}{2}[q(x + y) - q(x) - q(y)]. \quad (1)$$

In characteristic 2 we can still associate with each quadratic form a bilinear form (as in Q.2) and each quadratic form can be expressed as $b(x, x)$ for some bilinear form b (cf. Ex. (3)), but now b cannot in general be chosen to be symmetric. In what follows we shall mainly concentrate on the case $\text{char } k \neq 2$.

Let V be an inner product space. Relative to a basis e_1, \dots, e_n of V , its form b is determined by the coefficients $a_{ij} = b(e_i, e_j)$, for if $x = \sum \xi_i e_i, y = \sum \eta_i e_i$, then by linearity we have $b(x, y) = \sum a_{ij} \xi_i \eta_j$. In matrix form this can be written as

$$b(x, y) = \xi A \eta^T, \quad (2)$$

where $\xi = (\xi_1, \dots, \xi_n), \eta = (\eta_1, \dots, \eta_n)$ and T denotes transposition. Since b is symmetric, A is a symmetric matrix, i.e. $A^T = A$. Conversely, any symmetric matrix A defines a symmetric bilinear form on V by the formula (2).

If the basis of V is changed, let the new coordinates of x, y be the rows ξ', η' , where $\xi = \xi' P, \eta = \eta' P$ for some invertible matrix P (describing the change of basis in V). Then $b(x, y) = \xi' A' \eta'^T = \xi' A \eta^T = \xi' P A P^T \eta'^T$. Since this holds for all ξ', η' we conclude that

$$A' = P A P^T, \quad \text{where } P \text{ is invertible.} \quad (3)$$

Two matrices A, A' related as in (3) are said to be *congruent*; what we have said shows that the matrices of the form b in different coordinate systems are congruent. Conversely, if b is represented by the matrix A in one coordinate system and A' is congruent to A , say (3) holds, then transformation by P will define a new basis, relative to which b has the matrix A' . This proves

PROPOSITION 1.1 *Two matrices represent the same bilinear form in different coordinate systems if and only if they are congruent.* ■

The result holds for any bilinear forms, symmetric or not; for symmetric forms in characteristic not 2, the matrices can also be chosen to be symmetric. Moreover, given two matrices A, A' related as in (3), if one of them is symmetric, then clearly so is the other.

An *isometry* between two inner product spaces V, V' (or also between their forms) is an isomorphism $f: V \rightarrow V'$ which transforms the form of V into that of V' ; thus if the forms are b, b' respectively, then

$$b'(xf, yf) = b(x, y) \quad \text{for all } x, y \in V.$$

Equivalently (in characteristic not 2), if the quadratic forms in V, V' are q, q' respectively, then $q'(xf) = q(x)$ for all $x \in V$. We shall then say that the spaces V, V' are *isometric* and write $V \cong V'$. Relative to suitable bases the bilinear forms in

spaces that are isometric are represented by the same matrix; hence two spaces with given bases are isometric iff the matrices of the forms relative to these bases are congruent.

Let (V, b) be an inner product space. Any subspace U of V is again an inner product space, the form being $b|_U$, the restriction of the form b to U .

Two vectors $x, y \in V$ are said to be *orthogonal* if $b(x, y) = 0$; by the symmetry of b this is a symmetric relation. For any subset S of V we define its *orthogonal space* as

$$S^\perp = \{x \in V \mid b(x, y) = 0 \quad \text{for all } y \in S\}. \quad (4)$$

It is easily seen that S^\perp is a subspace of V ; this holds for any subset S of V , although we shall mainly use (4) when S is itself a subspace. In particular, the subspace V^\perp is called the *radical* of V ; it consists of all vectors orthogonal to all of V . If $V^\perp \neq 0$, then the form b (or also the space V) is called *singular*; otherwise V is *non-singular* or *regular*. In Vol. 1, inner product spaces were regular by definition, but here it is more appropriate not to make this restriction.

Let V be an inner product space whose form relative to a basis is given by (2). The radical of V is obtained by solving the equations $b(x, e_i) = 0$, i.e. $\xi A = 0$. These equations have a non-trivial solution precisely when A is singular, so we obtain

PROPOSITION 1.2 *An inner product space is singular if and only if the matrix of its form (relative to any basis of the space) is singular.* ■

Let V be a regular space and U any subspace. If v_1, \dots, v_n is a basis of V , chosen so that v_1, \dots, v_r is a basis of U , then a vector $x = \sum \alpha_i v_i$ is orthogonal to U iff

$$0 = b(x, v_j) = \sum \alpha_i b(v_i, v_j) \quad j = 1, \dots, r. \quad (5)$$

By hypothesis the $n \times n$ matrix $(b(v_i, v_j))$ is regular, hence the $n \times r$ matrix consisting of the first r columns has rank r . Therefore the system (5) has rank r and the space of solutions has dimension $n - r$. This proves the first part of

PROPOSITION 1.3 *If V is a regular inner product space and U is any subspace of V , then*

$$\dim U + \dim U^\perp = \dim V, \quad (6)$$

and

$$U^{\perp\perp} = U. \quad (7)$$

To establish (7) we note that by (6), $\dim U + \dim U^\perp = \dim U^\perp + \dim U^{\perp\perp}$, hence $\dim U^{\perp\perp} = \dim U$. Since clearly $U \subseteq U^{\perp\perp}$, it follows that $U^{\perp\perp} = U$. ■

We note that $U \cap U^\perp$ need not be 0; whether it is will depend on whether U is regular in the induced metric. We shall deal with this case in the next section.

For a regular space the determinant of the matrix, though not itself invariant,

provides an invariant of the space. By (3) we have the relation

$$\det A' = \det A \cdot (\det P)^2. \quad (8)$$

Here $(\det P)^2$ can assume the value of any non-zero square in the field k . If we denote this set of squares by $k^{\times 2}$, then by (8) each regular inner product space V determines a unique element of the factor group $k^{\times}/k^{\times 2}$, namely the residue class of $\det A$. This is called the *determinant* of the space V , or also of the form b .

Finally we remark that by fixing one of the arguments in an inner product, we obtain a linear functional on the space, i.e. an element of $V^* = \text{Hom}_k(V, k)$. The inner product b of V thus defines a mapping

$$\varphi_b: V \rightarrow V^*, \quad \text{given by } x \mapsto b(x, -).$$

This mapping is linear, as is easily seen. It is injective iff b is regular; since V^* and V have the same dimension, φ_b is then an isomorphism.

Exercises

- (1) Let V be an inner product space. Show that for each $u \in V$ the mapping $\lambda_u: x \mapsto b(x, u)$ is an element of V^* and that the mapping $\varphi: u \mapsto \lambda_u$ is an isomorphism iff V is regular. Examine the case where the coefficient ring is not a field.
- (2) Let V be an inner product space and U_1, U_2 subspaces. Show that
 (i) $U_1 \subseteq U_2 \Rightarrow U_1^\perp \supseteq U_2^\perp$, (ii) $U_1 \subseteq U_1^{\perp\perp}$, (iii) $U_1^\perp = U_1^{\perp\perp\perp}$.
- (3) Given a quadratic form q on a space V in characteristic 2, find a bilinear form b on V such that $q(x) = b(x, x)$. Find the conditions on q for which b can be chosen symmetric (as well as bilinear).
- (4) Show that any quadratic form q satisfies $q(x + y) + q(x - y) = 2[q(x) + q(y)]$.
- (5) Let V be a regular inner product space and U_1, U_2 subspaces. Show that $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$, $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.

6.2 Orthogonal sums and diagonalization

Let V and V' be inner product spaces with quadratic forms q and q' respectively; their *orthogonal sum* is defined as the direct sum of the spaces V and V' with the quadratic form Q defined by

$$Q(x + x') = q(x) + q'(x'), \quad x \in V, x' \in V'. \quad (1)$$

It is easily checked that Q is a quadratic form on $V \oplus V'$; the inner product space so defined is written $V \perp V'$. Clearly V' is orthogonal to V ; moreover, if q, q' have matrices A, A' , then Q has the matrix

$$\begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix}.$$

It follows that $\det Q = \det q \cdot \det q'$.

The next result is useful in decomposing a space into an orthogonal sum. Note that the space itself is not required to be regular, only the subspace.

LEMMA 2.1 *Let V be an inner product space and U a subspace which is regular (under the inner product induced from V). Then*

$$V = U \perp U^\perp. \quad (2)$$

Proof. Let e_1, \dots, e_r be a basis of U ; given $x \in V$, we write

$$x = \sum \xi_i e_i + y, \quad (3)$$

and try to solve the equations

$$\sum \xi_i b(e_i, e_j) = b(x, e_j), \quad (4)$$

which express that $b(y, e_j) = 0$ for all j , and so $y \in U^\perp$. Since U is regular, the matrix $(b(e_i, e_j))$ is invertible, hence the system (4) has a unique solution ξ_1, \dots, ξ_r . Since x in (3) was arbitrary in V , this proves that $V = U + U^\perp$; the sum is direct because U is regular and so $U \cap U^\perp = 0$, and it is clearly orthogonal. ■

When we have a decomposition of V of the form (2), U^\perp is called the *orthogonal complement* of U .

In characteristic not 2, every inner product space is an orthogonal sum of one-dimensional spaces. In essence this was already proved in Vol. 1 (Th. 2, 8.2). Below we state this result in the present terminology and, in view of its importance, briefly recall the proof.

THEOREM 2.2 *Every inner product space over a field of characteristic not 2 is an orthogonal sum of one-dimensional spaces.*

Proof. Let V be the space, with quadratic form q . If $q = 0$, the result is clear; otherwise take $e_1 \in V$ such that $q(e_1) \neq 0$ and let U be the subspace spanned by e_1 . Then U is regular and by Lemma 2.1, $V = U \perp U^\perp$. Now the result follows by induction on $\dim V$, because $\dim U^\perp < \dim V$. ■

In terms of matrices this means that every symmetric matrix in characteristic not 2 is congruent to a diagonal matrix. For example, for the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ we have

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

We see that $\text{char } k \neq 2$ is essential for this example.

If we write V as an orthogonal sum of one-dimensional spaces and take an

adapted basis e_1, \dots, e_n , this is called an *orthogonal basis* of V ; relative to this basis the quadratic form on V has a diagonal matrix $\text{diag}(a_1, \dots, a_n)$. Let us write

$$\langle a_1, \dots, a_n \rangle \quad (5)$$

for the inner product space with a matrix of this form; by Th. 2.2 every inner product space (in $\text{char} \neq 2$) can be expressed in the form (5). In particular, $\langle a \rangle$ denotes a one-dimensional space with the form ax^2 ; the n -dimensional space with the form $\sum ax_i^2$ is $\langle a, a, \dots, a \rangle$. This will occasionally be written as $\langle a^n \rangle$; the notation will mainly be used for $a = 0$ or 1, so the risk of ambiguity is slight. We observe that for any $c_1, \dots, c_n \in k^\times$, $\langle a_1, \dots, a_n \rangle \cong \langle c_1^2 a_1, \dots, c_n^2 a_n \rangle$.

From the proof of Th. 2.2 we see that if $q(e_1) = a \neq 0$, then $V = \langle a \rangle \perp V'$, for some space V' . Let us say the quadratic form q represents $a \in k$ if $q(x) = a$ for some $x \neq 0$. Then we can express the conclusion as follows:

COROLLARY 2.3 *Let V be an inner product space with quadratic form q . If q represents $a \neq 0$, then $V = \langle a \rangle \perp V'$, for some subspace V' of V .* ■

It is often easier to work with regular spaces; this is achieved by the following reduction.

PROPOSITION 2.4 *Let V be an inner product space and let V_0 be a complement of V^\perp in V (as vector space). Then*

$$V = V_0 \perp V^\perp. \quad (6)$$

The subspace V_0 is regular and is determined up to isometry by V .

Proof. Clearly V_0 is orthogonal to V^\perp , hence (6) follows. Moreover, $V_0^\perp = V^\perp$ and this meets V_0 in 0, so V_0 is regular. Now since $b(x, y)$ vanishes for x or y in V^\perp , we can define the quotient V/V^\perp with the natural homomorphism $x \mapsto \bar{x}$ from V to V/V^\perp as an inner product space by putting $b(\bar{x}, \bar{y}) = b(x, y)$. With this definition it is clear that $V/V^\perp \cong V_0$, so the latter is unique up to isometry. ■

The space V_0 in Prop. 2.4 is called the *regular part* of V ; its dimension is the *rank* of V or also of the quadratic form q , written $\text{rk } q$. In view of this result we can mainly restrict ourselves to regular forms. We note the following criterion for isometry:

COROLLARY 2.5 *Let k be a field of characteristic not 2, in which every element is a square. Then two inner product spaces over k are isometric if and only if they have the same dimension and the same rank.*

The condition is clearly necessary; since $\langle a \rangle \cong \langle 1 \rangle$ for any $a \in k^\times$, any space of dimension n and rank r is isometric to $\langle 1^r, 0^{n-r} \rangle$, hence the condition is also sufficient. ■

In particular this solves the classification problem for any algebraically closed field of characteristic not 2.

The last result can be generalized as follows. A quadratic form is said to be *universal* if it represents every non-zero element of the field.

LEMMA 2.6 *Let k be a field of characteristic not 2. If every inner product space of rank v is universal (for some $v \geq 1$), then every inner product space of dimension n is isometric to $\langle \alpha_1, \dots, \alpha_r, 1^s, 0^{n-r-s} \rangle$, where $\alpha_1, \dots, \alpha_r$ are non-squares in k and $r < v$.*

Proof. Clearly $\langle c \rangle \cong \langle 1 \rangle$ for any non-zero square c in k , therefore any inner product space V is isometric to $\langle \alpha_1, \dots, \alpha_r, 1^s, 0^{n-r-s} \rangle$, where the α 's are non-squares. We shall use induction on r . For $r < v$ there is nothing to prove, so assume that $r \geq v$. Then $\langle \alpha_1, \dots, \alpha_r \rangle$ represents 1, by hypothesis, so by Cor. 2.3, we have $V \cong \langle \alpha'_2, \dots, \alpha'_r, 1^{s+1}, 0^{n-r-s} \rangle$ and we can now apply induction on r to complete the proof. ■

As an application let us take a finite field k of odd characteristic. In this case the hypothesis of Lemma 2.6 holds for $v = 2$. To establish this fact we must show that any binary quadratic form over k is universal; thus we have to solve the equation

$$ax^2 + by^2 = c \quad (7)$$

for any $a, b, c \in k^\times$. Let A be the set of all elements ax^2 ($x \in k$) and B the set of all elements $c - by^2$ ($y \in k$). If the field k has q elements, it contains $(q+1)/2$ squares, because the group endomorphism $x \mapsto x^2$ of k^\times has a kernel of order 2. Hence A, B have $(q+1)/2$ elements each and so $A \cap B \neq \emptyset$. This provides a solution for (7).

Applying Lemma 2.6, we see that every quadratic form over a finite field of characteristic not 2 has the form $\langle \alpha, 1^r, 0^s \rangle$. For a regular form this becomes $\langle \alpha, 1^{n-1} \rangle$. Now either α is a square; then we just have $\langle 1^n \rangle$. Or α is not a square; then the previous form cannot be simplified. We note that the determinant of $\langle \alpha, 1^{n-1} \rangle$ is α , so we can decide to which case our form belongs by looking at the determinant. We sum up the results as

THEOREM 2.7 *Two regular forms over a finite field of odd characteristic are isometric if and only if they have the same rank and determinant. More precisely, if λ is any non-square in k , then any regular form of rank n is isometric to $\langle 1^n \rangle$ or $\langle \lambda, 1^{n-1} \rangle$ according as the determinant is or is not 1.* ■

Exercises

- (1) Verify that the function Q defined in (1) is a quadratic form on $V \oplus V'$.

- (2) Show that (in char $\neq 2$) a regular quadratic form which represents 0 is universal. (Hint. If $\sum a_i \alpha_i^2 = 0$, put $x_1 = \alpha_1(1+t)$, $x_i = \alpha_i(1-t)$, $i = 2, \dots, n$.)
- (3) Show that a regular quadratic form $\langle a_1, \dots, a_n \rangle$ represents $a \neq 0$ iff $\langle a_1, \dots, a_n, -a \rangle$ represents 0.
- (4) Show that two binary quadratic forms (i.e. of dimension 2) in characteristic $\neq 2$ are isometric iff they have the same determinant and there exists $a \in k^\times$ which is represented by both.
- (5) Let V be a three-dimensional inner product space of characteristic 2 whose bilinear form has the matrix

$$\begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

where $a \neq 0$. Find an orthogonal basis. (Hint. Take a basis including $(1, 1, 1)^T$.)

- (6) An inner product space V in characteristic 2 with form b is said to be *alternating* if $b(x, x) = 0$ for all $x \in V$. Use the proof of Th. 2.2 and Ex. (5) to show (by induction on $\dim V$) that every inner product space whose form is not alternating has an orthogonal basis.
- (7) Let V be an inner product space of characteristic 2. Show that the set $I = \{x \in V | q(x) = 0\}$ is a subspace containing the radical and give examples where these two subspaces are distinct. Prove Th. 2.2 in the case $I = V^\perp$.

6.3 The orthogonal group of a space

Let V be an inner product space; the isometries of V with itself are called *orthogonal transformations*. Clearly they form a group, called the *orthogonal group* and denoted by $\mathbf{O}(V)$, or $\mathbf{O}(V, q)$ if the quadratic form q is to be emphasized. The elements of $\mathbf{O}(V)$ are the vector space automorphisms θ of V such that $q(\theta(x)) = q(x)$ for all $x \in V$. In terms of matrices, if q has the matrix A (relative to some fixed basis of V), then $\mathbf{O}(V)$ consists of all invertible matrices X such that

$$XAX^T = A. \tag{1}$$

When V is regular, we see from (1) by taking determinants that $(\det X)^2 = 1$, hence $\det X = \pm 1$. The *special orthogonal group* $\mathbf{SO}(V)$ is defined as the subgroup of $\mathbf{O}(V)$ consisting of all isometries of determinant 1. Its elements are the *proper* orthogonal transformations, also called *rotations*; the elements of $\mathbf{O}(V)$ not in $\mathbf{SO}(V)$ are called *improper*. By taking A diagonal in (1) we see that when -1 is not a square, there are always improper orthogonal transformations, so that $\mathbf{SO}(V)$ is then a subgroup of index 2 in $\mathbf{O}(V)$.

In the special case of a space with the standard quadratic form $\sum x_i^2$ the

condition (1) for a matrix X to be orthogonal reduces to

$$XX^T = I.$$

The set of all $n \times n$ matrices over k satisfying this condition is usually denoted by $\mathbf{O}_n(k)$ and is also called the *orthogonal group* of degree n over k , with the subgroup $\mathbf{SO}_n(k)$ of proper orthogonal matrices.

It is clear that isometric spaces have isomorphic orthogonal groups. Explicitly, if $\alpha: V \rightarrow V'$ is an isometry, then the map $\theta \mapsto \alpha^{-1}\theta\alpha$ is an isomorphism from $\mathbf{O}(V)$ to $\mathbf{O}(V')$.

A vector $x \neq 0$ in an inner product space V is said to be *isotropic* if $q(x) = 0$; otherwise x is *anisotropic*. A subspace U of V is *isotropic* if it contains an isotropic vector, *anisotropic* otherwise. If every non-zero vector in U is isotropic, this means (in characteristic not 2) that the inner product restricted to U vanishes identically; in that case U is said to be *totally isotropic*.

Let us now take a closer look at the orthogonal group $\mathbf{O}(V)$ of an inner product space V . Throughout we shall suppose that $\text{char } k \neq 2$ and that V is regular.

Any $\alpha \in \mathbf{O}(V)$ such that $\alpha^2 = 1$ is called an *involution*. Such transformations may be obtained as follows. Let $V = U \perp U'$ be an orthogonal sum decomposition; we define the *reflexion* with respect to (U, U') as the linear transformation of V which maps each $x \in U$ to $-x$ and leaves U' pointwise fixed. Clearly this is an orthogonal transformation whose square is 1, i.e. it is an involution.

Conversely, let α be an involution on V and put

$$U_+ = \{x \in V | x\alpha = x\}, \quad U_- = \{x \in V | x\alpha = -x\}.$$

Then $U_+ \cap U_- = 0$, because $\text{char } k \neq 2$, and $V = U_+ + U_-$, since every $x \in V$ can be written as $x = u + v$, where $u = (x + x\alpha)/2$, $v = (x - x\alpha)/2$, and clearly $u \in U_+$, $v \in U_-$. Finally, if $x \in U_+$, $y \in U_-$ then $b(x, y) = b(x\alpha, y\alpha) = -b(x, y)$; hence $b(x, y) = 0$ and so $V = U_- \perp U_+$. This shows that every involution is in fact a reflexion.

A reflexion with respect to (U, U') is said to be of *type p* if $\dim U = p$. If we use a basis of V adapted to the decomposition $V = U \perp U'$, then α is represented by a diagonal matrix with $p - 1$'s and $n - p$ 1's, where $n = \dim V$. For example, a reflexion of type 2 in Euclidean 3-space is a rotation through an angle π .

The reflexions of type 1 are particularly important: they consist of reflexions in a hyperplane and the general reflexion can be expressed as a product of reflexions of type 1. To obtain an explicit formula for the reflexion with respect to (u, u^\perp) , where $u \in V$ is an anisotropic vector, consider the mapping

$$\sigma_u: x \mapsto x - \lambda(x)u,$$

where λ is a linear function of x . This is a linear transformation of V and it will be orthogonal iff $q(x) = q(x - \lambda u) = q(x) - 2\lambda b(x, u) + q(u)\lambda^2$. Excluding the trivial case $\lambda = 0$, we see that the condition for an isometry is $\lambda q(u) - 2b(x, u) = 0$; hence

$$\sigma_u: x \mapsto x - \frac{2b(x, u)}{q(u)}u \tag{2}$$

is an orthogonal transformation. It is called the *symmetry* with respect to u . By using a basis adapted to the decomposition $V = (u) \perp u^\perp$ we see that it is the reflexion with respect to (u, u^\perp) : the vectors along u are reversed while the hyperplane of vectors orthogonal to u remains fixed. This makes it clear that the symmetries are just the reflexions of type 1. A reflexion of type p is described in a suitable orthogonal basis e_1, \dots, e_n by $e_i \mapsto -e_i$ ($i = 1, \dots, p$), $e_i \mapsto e_i$ ($i = p+1, \dots, n$). But this can be written as $\sigma_{e_1} \dots \sigma_{e_p}$, hence every reflexion can be written as a product of symmetries. The reflexions in turn generate $\mathbf{O}(V)$; to establish this fact we first show that any anisotropic vector can be transformed into any other vector of the same length by a reflexion.

LEMMA 3.1 *In an inner product space V (in characteristic $\neq 2$) let u, v be vectors such that $q(u) = q(v) \neq 0$. Then there is a reflexion of V which maps u to v .*

Proof. Since $q(u) = q(v)$, the vectors $x = (u + v)/2$, $y = (u - v)/2$ are orthogonal, as is easily checked, and they cannot both be isotropic, because $q(u) \neq 0$. Let X, Y be the spaces spanned by x, y respectively; if $q(x) \neq 0$, then $V = X \perp X^\perp$ and the reflexion with respect to (X^\perp, X) maps $u = x + y$ to $v = x - y$; if $q(y) \neq 0$, then $V = Y \perp Y^\perp$ and the reflexion with respect to (Y, Y^\perp) maps u to v . ■

THEOREM 3.2 *Let V be a regular inner product space (in characteristic $\neq 2$) of dimension n . Then any orthogonal transformation in V can be written as a product of at most n reflexions.*

Proof. Let $\alpha \in \mathbf{O}(V)$ and choose an anisotropic vector $e_1 \in V$. Then $q(e_1) = q(e_1\alpha)$, hence by Lemma 3.1 there is a reflexion σ_1 such that $e_1\sigma_1 = e_1\alpha$, and so $\alpha\sigma_1$ leaves e_1 fixed. It follows that $\alpha\sigma_1$ also maps $V' = e_1^\perp$ into itself, but $\dim V' = n - 1$, so by induction on n we can write $\alpha\sigma_1|_{V'} = \sigma'_n \dots \sigma'_2$, where σ'_i is a reflexion in V' . Each σ'_i can be extended to a reflexion σ_i of V which leaves e_1 fixed. Then $\alpha\sigma_1\sigma_2 \dots \sigma_n$ leaves e_1 fixed, as well as every vector in V' . Hence it must be the identity, and so $\alpha = \sigma_n \dots \sigma_1$, as we wished to show. ■

Since every reflexion is a product of at most n symmetries, we have the

COROLLARY 3.3 *Every orthogonal transformation on an n -dimensional regular space (in characteristic not 2) is a product of at most n^2 symmetries.* ■

This bound can still be improved: if we examine the proof of Lemma 3.1 we see that the reflexion there can be taken to be of type 1 or 2, so the bound n^2 can be replaced by $2n$. But there is a more precise result: the Cartan–Dieudonné theorem states that every orthogonal transformation on an n -dimensional space can be written as a product of at most n symmetries, cf. e.g. Artin (1957). For example, in three dimensions every orthogonal transformation is a product of at most three

symmetries and a proper orthogonal transformation is a product of at most two symmetries; each leaves a plane fixed and these two planes meet in a line, so the transformation must leave a line fixed, i.e. it is a rotation about that line.

Let α be any orthogonal transformation and u an anisotropic vector in V . Then for any $x \in V$, x and $x\sigma_u$ are symmetric with respect to u , hence $x\alpha$ and $x\alpha\sigma_{u\alpha}$ are symmetric with respect to $u\alpha$, and so $(x\sigma_u)\alpha = x\alpha\sigma_{u\alpha}$. Thus we obtain

$$\sigma_{u\alpha} = \alpha^{-1}\sigma_u\alpha. \quad (3)$$

Since $\mathbf{O}(V)$ is generated by the symmetries σ_u , its derived group $\mathbf{O}(V)'$ is generated by all commutators of symmetries $\sigma_u^{-1}\sigma_v^{-1}\sigma_u\sigma_v = (\sigma_u\sigma_v)^2$. For the subgroup H generated by all $(\sigma_u\sigma_v)^2$ is normal in $\mathbf{O}(V)$, by (3), and any two symmetries commute (mod H), hence $\mathbf{O}(V)/H$ is abelian, while clearly $H \subseteq \mathbf{O}(V)'$; therefore $\mathbf{H} = O(V)'$. We can also express $\mathbf{O}(V)'$ in terms of rotations as long as $\dim V$ is not too small:

THEOREM 3.4 *Let V be a regular inner product space (in characteristic $\neq 2$). Then the derived group $\mathbf{O}(V)'$ of the orthogonal group is generated by all $(\sigma_u\sigma_v)^2$ ($u, v \in V$). If $\dim V > 2$, then $\mathbf{O}(V)' = \mathbf{SO}(V)'$.*

Proof. The first part has been shown and it is clear that $\mathbf{SO}(V)' \subseteq \mathbf{O}(V)'$, so it only remains to show that $(\sigma_u\sigma_v)^2$ is a product of commutators of rotations. When $\dim V$ is odd, -1 is an improper orthogonal transformation and $(\sigma_u\sigma_v)^2 = [(-\sigma_u)(-\sigma_v)]^2$ is the required representation. There remains the case when $\dim V$ is even, and so $\dim V \geq 4$. We note that if w is a vector orthogonal to u , then $\sigma_u\sigma_w\sigma_u = \sigma_w$, hence σ_u commutes with σ_w . Thus if we can find an anisotropic vector w orthogonal to u and v , then $(\sigma_u\sigma_v)^2 = (\sigma_u\sigma_w\cdot\sigma_v\sigma_w)^2$, and this is the desired expression of $(\sigma_u\sigma_v)^2$ as commutator of two rotations. It remains to show that such a vector w can always be found. If this is not the case, then every vector orthogonal to u and v is isotropic. Let U be the subspace spanned by u and v , then U^\perp is totally isotropic, hence $U^\perp \subseteq U^{\perp\perp} = U$, and $\dim U^\perp = n - \dim U \geq n - 2 \geq 2$, so $U = U^\perp$ is totally isotropic, which contradicts the fact that $q(u) \neq 0$. ■

We shall see later that when $\dim V = 2$, then $\mathbf{SO}(V)$ is abelian, but $\mathbf{O}(V)$ need not be (cf. Ex. (4)).

Exercises

- (1) Show that a product of r symmetries on an n -dimensional space leaves fixed a subspace of dimension at least $n - r$. Deduce that there are orthogonal transformations which cannot be written as a product of fewer than n symmetries.
- (2) Show that when $\dim V = n$, then $\mathbf{O}(V)$ can be generated by all reflexions of type $n - 1$. What can be said about other types?

- (3) Show that the reflexion in Lemma 3.1 can be taken to be of type 1 or 2. (*Hint.* If $q(y) = 0$, take a reflexion with respect to (H, H^\perp) , where H is a two-dimensional anisotropic subspace containing y and orthogonal to x .)
- (4) Find all fields k and dimensions n for which $\mathbf{O}_n(k)$ can be commutative.
- (5) Show that any orthogonal transformation in the centre of $\mathbf{O}(V)$ leaves any anisotropic vector fixed or reverses it. Deduce that the centre of $\mathbf{O}(V)$ is $\{1, -1\}$.

6.4 Witt's cancellation theorem and the Witt group of a field

We have seen in Th. 2.2 that the matrix of a quadratic form can always be taken in diagonal form, for a suitable choice of basis. However, this is not always the most convenient form, particularly as the diagonal elements of the matrix are not invariants, and the relation between different diagonal matrices is not straightforward. We shall return to this point in Vol. 3, but for the moment consider a different way of decomposing an inner product space. This is the theory developed by E. Witt in 1937, which leads to an invariant of the space; moreover, these invariants form an abelian group, the Witt group associated with the field k . We begin with a cancellation theorem which is often useful.

THEOREM 4.1 (Witt's cancellation theorem) *Let V be an inner product space (in characteristic $\neq 2$) and V_1, V_2 regular subspaces. If $V_1 \cong V_2$, then $V_1^\perp \cong V_2^\perp$.*

Proof. Assume first that V_1, V_2 are one-dimensional, spanned by v_1, v_2 respectively, where $q(v_1) = q(v_2) \neq 0$. By Lemma 3.1 there is an orthogonal transformation θ mapping v_1 to v_2 ; hence θ also maps V_1^\perp to V_2^\perp , so these spaces are isometric, as claimed.

In general we have $V = V_1 \perp V_1^\perp = V_2 \perp V_2^\perp$, by Lemma 2.1. Now we may take $V_1 \cong V_2 \cong \langle a_1, \dots, a_r \rangle$, by Th. 2.2, and applying the case just proved to $\langle a_1 \rangle$ we obtain

$$\langle a_2, \dots, a_r \rangle \perp V_1^\perp \cong \langle a_2, \dots, a_r \rangle \perp V_2^\perp,$$

and now the result follows by induction on $\dim V_1$. ■

We observe that the result may be stated in the form

$$U \perp W_1 \cong U \perp W_2 \Rightarrow W_1 \cong W_2.$$

This form accounts for the name ‘cancellation theorem’. In characteristic 2 the result does not hold: let H be a two-dimensional inner product space with matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; then $H \not\cong \langle 1, 1 \rangle$ in characteristic 2, but $\langle 1 \rangle \perp H \cong \langle 1, 1, 1 \rangle \cong \langle 1 \rangle \perp \langle 1, 1 \rangle$, by Ex. (5) of 6.2. However, the cancellation

theorem does hold in characteristic 2 for alternating forms (cf. Ex. (5) of Further exercises).

For any field k we can define a monoid $M(k)$ as follows: the elements of $M(k)$ are the isometry classes of regular inner product spaces over k . If the class of V is written $[V]$, we can define an addition by the formula

$$[U] + [V] = [U \perp V]. \quad (1)$$

If $U \cong U'$, $V \cong V'$, then $U \perp V \cong U' \perp V'$; this shows the right-hand side of (1) to depend only on $[U]$, $[V]$ and not on U , V themselves; so the addition of classes is well-defined. The operation is clearly associative, so we obtain a monoid in this way, with the class of the zero space as the neutral element. Now Th. 4.1 shows that this monoid $M(k)$ satisfies cancellation.

Th. 4.1 leads to a further reduction; to describe it we must first look at the spaces which are to be regarded as 'trivial'. Let us define a *hyperbolic plane* as a two-dimensional regular isotropic space. Such a space may also be characterized as follows:

PROPOSITION 4.2 *For any two-dimensional inner product space V in characteristic not 2, the following conditions are equivalent:*

- (a) V is a hyperbolic plane,
- (b) V has a basis u, v such that $b(u, v) = 1$, $q(u) = q(v) = 0$,
- (c) $V \cong \langle 1, -1 \rangle$,
- (d) V has determinant $-1 \pmod{k^{\times 2}}$.

A basis u, v as in (b) is called a *hyperbolic pair* for V .

Proof. (a) \Rightarrow (b). Let u be an isotropic vector and complete it to a basis u, v' of V . Then $b(u, v') \neq 0$, because V is regular, and on multiplying v' by a suitable scalar we may assume that $b(u, v') = 1$. Now for any $\alpha \in k$, $b(v' - \alpha u, v' - \alpha u) = q(v') - 2\alpha$ and we can choose α so that the right-hand side is 0; then u and $v = v' - \alpha u$ form a hyperbolic pair for V .

(b) \Rightarrow (c). If u, v is a hyperbolic pair, then $e = (u + v)/2$, $f = (u - v)/2$ is an orthogonal basis and $q(e) = -q(f) = 1$.

(c) \Rightarrow (d) is clear; to prove (d) \Rightarrow (a) we have by Th. 2.2, $V \cong \langle a, b \rangle$, where $-ab$ is a square, say $-ab = c^2$. Then $-a^2b = c^2a$ and now $ax^2 + by^2 = 0$ for $x = c$, $y = a$, so V is isotropic. ■

From (b) or (c) it is clear that a hyperbolic plane is unique up to isometry.

In studying the monoid $M(k)$ we shall regard orthogonal sums of hyperbolic planes as trivial. The reason for this will become clear when we come to deal with Clifford algebras (in Vol. 3): hyperbolic planes are the spaces whose Clifford algebras are full matrix rings. Orthogonal sums of hyperbolic planes may be defined more simply as split spaces. An inner product space V is said to be *split* if it

is regular and has a subspace U (the splitting space) such that $U^\perp = U$. It is clear that an orthogonal sum of hyperbolic planes is split: if

$$V \cong H_1 \perp \cdots \perp H_r, \quad (2)$$

where each H_i is a hyperbolic plane, with hyperbolic pair u_i, v_i say, write U for the subspace spanned by u_1, \dots, u_r . Then it is clear that U is totally isotropic, hence $U \subseteq U^\perp$. If $x = \sum \alpha_i u_i + \sum \beta_i v_i \in U^\perp$, then $0 = b(x, u_i) = \beta_i$, hence $x = \sum \alpha_i u_i \in U$. This shows that $U^\perp \subseteq U$, so that U splits V .

The fact that, conversely, every split space is of the form (2), follows from the next result, which is slightly more general.

THEOREM 4.3 *Any regular inner product space V (in characteristic $\neq 2$) can be written in the form*

$$V \cong H_1 \perp \cdots \perp H_r \perp U, \quad (3)$$

where H_1, \dots, H_r are hyperbolic planes and U is an anisotropic space. Here r is uniquely determined as the maximal dimension of a totally isotropic subspace of V and U is determined up to isometry.

This decomposition (3) is called the *Witt decomposition* of V , and the integer r is called the *Witt index* of V . The subspace U is called the *anisotropic part* of V .

Proof. Let W be any totally isotropic subspace of V , with basis u_1, \dots, u_s and choose $v_1 \in V$ such that $b(u_1, v_1) \neq 0$, $b(u_i, v_1) = 0$ for $i > 1$; since V is regular, this is always possible. The subspace H_1 spanned by u_1, v_1 is regular and it contains the isotropic vector u_1 , hence it is a hyperbolic plane, and by Lemma 2.1 we have the decomposition

$$V = H_1 \perp V',$$

where $V' = H_1^\perp$. Now V' contains u_2, \dots, u_s ; hence by induction on s there exist $v_2, \dots, v_s \in V'$ such that u_i, v_i span a hyperbolic plane H_i and we have a decomposition

$$V = H_1 \perp H_2 \perp \cdots \perp H_s \perp U. \quad (4)$$

If U is isotropic, we can split off more hyperbolic planes, and the process ends when we reach a component U which is anisotropic. Thus we always reach a decomposition (3) where U is anisotropic, and by construction $r \geq s$. It is clear that V given by (3) contains a totally isotropic subspace of dimension r , hence r is uniquely determined as the maximal dimension of a totally isotropic subspace of V . Finally, U is unique up to isometry by the Witt cancellation theorem. ■

If V itself is split, by a subspace W of dimension r , and we take a decomposition (3) of V , then since $W^\perp = W$, we have $U = 0$ and hence we obtain

COROLLARY 4.4 Any split space in characteristic $\neq 2$ is an orthogonal sum of hyperbolic planes. ■

For any field k of characteristic not 2 we now define the *Witt group* $W(k)$ as follows. In the monoid $M(k)$ let us identify any two classes $[V], [V']$ whose spaces have anisotropic parts that are isometric. Writing again $[V]$ for the class of V , we now have

$$[V] = [V'] \Leftrightarrow V \perp mH \cong V' \perp nH,$$

where nH stands for the orthogonal sum of n hyperbolic planes. We thus obtain a monoid $W(k)$, which is in fact a group. For if V is an n -dimensional space with regular form q , and V' is the space with form $-q$, then we can by Th. 2.2 write $V \cong \langle a_1, \dots, a_n \rangle$, $V' \cong \langle -a_1, \dots, -a_n \rangle$; therefore $V \perp V' \cong \langle a_1, -a_1, \dots, a_n, -a_n \rangle$. Now $\langle a_i, -a_i \rangle$ is regular isotropic, hence a hyperbolic plane, and so $V \perp V' \cong nH$. It follows that $[V]$ has the inverse $-[V] = [V']$, so $W(k)$ is indeed a group, called the *Witt group*. Clearly it is an invariant of the field k . For example, for an algebraically closed field the Witt group is of order 2; for the real field \mathbf{R} , quadratic forms can be classified by $n - 2v$, where v is the index, hence we have $W(\mathbf{R}) \cong \mathbf{Z}$. In Vol. 3 we shall see that a ring structure can be defined on $W(k)$, giving the *Witt ring* of k .

By examining the proof of Th. 4.3 we obtain a useful result on extending partial isometries:

THEOREM 4.5 (Witt's extension theorem) Let V be a regular inner product space (in characteristic $\neq 2$) and W a subspace of V . Then any isometric mapping $\theta: W \rightarrow V$ can be extended to an orthogonal transformation of V .

Proof. Let W_0 be the radical of W and write $W = W_0 \perp X$. Here X is regular, so by Lemma 2.1, $V = X \perp X^\perp$. Now $X^\perp \supseteq W_0$ and as in the proof of Th. 4.3 we can write $X^\perp = Z \perp Y$, where Z is a regular subspace split by W_0 . Since $X \cong X^\theta$, we have by Th. 4.1, $X^\perp \cong (X^\theta)^\perp$; but $(X^\theta)^\perp \supseteq W_0^\theta$ because W_0 is orthogonal to X . Therefore we can write $(X^\theta)^\perp = Z' \perp Y'$, where Z' is a regular subspace split by W_0^θ . Let u_1, \dots, u_s be a basis of W_0 and write $Z = H_1 \perp \dots \perp H_s$, $Z' = H'_1 \perp \dots \perp H'_s$, where H_i has the hyperbolic pair (u_i, v_i) and H'_i has the hyperbolic pair (u_i^θ, v_i) . The mapping $u_i \mapsto u_i^\theta$, $v_i \mapsto v_i'$ is an isometry $Z \rightarrow Z'$ whose restriction to W_0 agrees with θ ; thus we have extended θ to the regular space $U = X \perp Z$. Again we have $U \cong U^\theta$, so $U^\perp \cong (U^\theta)^\perp$ and this latter isomorphism gives the desired extension of θ to all of $V = U \perp U^\perp$. ■

Exercises

- (1) Show that a space V is split iff it is regular and has a linear mapping α into itself such that $q(x) + q(x\alpha) = 0$ for all $x \in V$.

- (2) Show that an orthogonal transformation of a split space which is the identity on a maximal totally isotropic subspace is a rotation.
- (3) Show that a hyperbolic plane has exactly two isotropic lines (i.e. every isotropic vector is proportional to one of exactly two vectors). Show also that each vector of a hyperbolic pair uniquely determines the other.
- (4) Show that an orthogonal transformation of a hyperbolic plane is a symmetry iff it interchanges the two isotropic lines.
- (5) For any hyperbolic plane H show that $\mathbf{SO}(H) \cong k^\times$. Describe the orthogonal group of H . What is the orthogonal group of $H \perp \langle 0 \rangle$?
- (6) Prove Th. 4.5 in the special case where W (as well as V) is regular. When can the orthogonal transformation in Th. 4.5 be chosen to be a rotation?
- (7) Give a direct proof (not using Th. 4.3) that any two split spaces of the same dimension are isomorphic.
- (8) Give a proof of Th. 4.1 using Th. 4.5.

6.5 Ordered fields

By an *ordered ring* we understand a non-trivial ring R with a total ordering ' $>$ ' which is preserved by the ring operations, in the sense that the following rules hold:

- O.1** $x > x', y > y' \Rightarrow x + y > x' + y',$
O.2 $x > 0, y > 0 \Rightarrow xy > 0.$

If in **O.2** we replace x, y by $x - x', y - y'$ and rearrange the result, using **O.1**, we obtain the more general rule

$$\mathbf{O.2}' \quad x > x', y > y' \Rightarrow xy + x'y' > xy' + x'y.$$

In discussing the ordering on R we shall, as usual, write ' $x \geq y$ ' to mean ' $x > y$ or $x = y$ ' and use $\leq, <$ for the opposite ordering. We shall call x *positive* or *negative* according as $x > 0$ or $x < 0$.

For any element a of an ordered ring we define the *absolute value* by

$$|a| = \begin{cases} a & \text{if } a \geq 0, \\ -a & \text{if } a < 0. \end{cases}$$

Clearly $|a| \geq 0$ with equality iff $a = 0$, and we have the *triangle inequality*

$$|a + b| \leq |a| + |b|, \tag{1}$$

as well as the multiplicativity rule

$$|ab| = |a| \cdot |b|. \quad (2)$$

Of these, (2) is an immediate consequence of the formula $|a| = \sqrt{a^2}$. To prove (1) we remark that this clearly holds (with equality) when a, b have the same sign. If $a \geq 0 \geq b$ say, then $a + b < a < a - b = |a| + |b|$ and $-a - b \leq -b \leq a - b = |a| + |b|$, hence (1) follows in this case. By symmetry (1) also holds when $b \geq 0 \geq a$.

An ordered ring can also be described by its positive cone. In any ring R , a *cone* is a subset P containing 1 but not 0, and containing with any x, y also $x + y, xy$. For example, in an ordered ring R , the set of all positive elements

$$P = \{x \in R \mid x > 0\}$$

is a cone, called the *positive cone* of R . Moreover, since R is totally ordered, it has the property that

$$\text{each } x \in R \text{ lies in just one of the sets } \{0\}, P, -P = \{-x \mid x \in P\}. \quad (3)$$

Conversely, every cone satisfying (3) defines an ordering on R : we put $x > y$ iff $x - y \in P$.

We note that any ordered ring, and more generally any ring with a cone, is non-trivial, i.e. $1 \neq 0$.

PROPOSITION 5.1 *Any ordered ring is an integral domain, and the square of any non-zero element is positive.*

Proof. Let R be an ordered ring; then $1 \neq 0$ by definition. Given $x, y \in R, x, y \neq 0$, if $x, y > 0$, then $xy > 0$. Similarly if $x, y < 0$, then $-x, -y > 0$ and so $xy = (-x)(-y) > 0$. There remains the case when x, y have opposite signs, say $x > 0 > y$. Then $x, -y > 0$, hence $-xy = x(-y) > 0$, and so $xy < 0$; similarly if $x < 0 < y$, then $xy < 0$. In each case $xy \neq 0$, hence R is an integral domain. In particular, when $x = y$, the last two cases cannot occur and $x^2 > 0$. ■

For the rest of this section all our rings will be commutative. Thus an ordered commutative ring R is an integral domain, and hence has a field of fractions K . We claim that there is just one way of ordering K so as to extend the ordering of R , namely by the rule:

$$\text{For any } a, b \in R, \text{ where } b \neq 0, a/b > 0 \text{ in } K \text{ iff } ab > 0. \quad (4)$$

In the first place, since $a/b \cdot b^2 = ab$ and $b^2 > 0$, any ordering on K which extends that on R must satisfy (4). It remains to show that (4) defines an ordering. It is well-defined, for if $a/b = a'/b'$ then $ab' = ba'$ and hence $abb'^2 = a'b^2b'$. It follows that $ab > 0 \Leftrightarrow abb'^2 > 0 \Leftrightarrow a'b^2b' > 0 \Leftrightarrow a'b' > 0$. Further, it clearly defines a trichotomy on K as in (3), and if $a_1b_1 > 0, a_2b_2 > 0$, then $(a_1b_2 + a_2b_1)b_1b_2 = a_1b_1b_2^2 + a_2b_2b_1^2 > 0$.

$> 0, a_1 a_2 b_1 b_2 = a_1 b_1 \cdot a_2 b_2 > 0$. This shows that if the fractions $f_i = a_i/b_i$ satisfy $f_i > 0$ ($i = 1, 2$), then $f_1 + f_2 > 0$, $f_1 f_2 > 0$.

We also note that $1, 1+1, \dots$ are all > 0 , hence R and with it K must be of characteristic 0. The result may be summed up as

THEOREM 5.2 *Any ordered commutative ring is an integral domain of characteristic 0. The ordering can be extended to an ordering of its field of fractions in just one way.* ■

As an example consider the integers \mathbf{Z} ; in any ordering $1, 1+1, \dots$ must all be positive, and since these numbers and their negatives, with 0, exhaust \mathbf{Z} , there is only one way of ordering \mathbf{Z} , namely the familiar ‘natural’ ordering of the integers. By Th. 5.2 the same applies to \mathbf{Q} and we have

COROLLARY 5.3 *There is only one way of making \mathbf{Q} into an ordered field, namely by the usual ordering.* ■

The real numbers \mathbf{R} also admit only one ordering, but for a different reason: every square must be positive (or 0), and since the squares, their negatives and 0 exhaust \mathbf{R} , no other ordering is possible apart from the usual one. We shall return to this point in Vol. 3. An example of a field with different orderings is $\mathbf{Q}(\alpha)$, where α is a root of $x^2 = 2$. There are two ways of embedding this field into \mathbf{R} , mapping α to $\sqrt{2}$ or to $-\sqrt{2}$, and correspondingly two ways of ordering it.

Two ordered rings R, R' are said to be *order-isomorphic* if there is a ring isomorphism $x \mapsto x'$ from R to R' which preserves the ordering, i.e. $x > 0 \Leftrightarrow x' > 0$; clearly it is enough to require that $x > 0 \Rightarrow x' > 0$. For example, the prime subfield of any ordered field K is order-isomorphic to \mathbf{Q} ; we also say that there is an *order-embedding* of \mathbf{Q} in K .

Exercises

- (1) Show that in any ordered field, $a > b > 0$ implies $b^{-1} > a^{-1}$.
- (2) Let K be an ordered field. Given $\alpha, \beta \in K$, show that if $\alpha < \beta$, then there exists $\gamma \in K$ such that $\alpha < \gamma < \beta$. (This property, that between any two elements of K a third can be found, is expressed by saying that K is *dense in itself*.)
- (3) Let K be an ordered field. Show that the polynomial ring $K[x]$ can be ordered by taking as positive cone the set of all polynomials with positive highest coefficient; hence obtain an ordering of the rational function field $K(x)$. Compare this with the ordering on $K(x)$ obtained by taking the polynomials with positive coefficient of lowest term as positive cone.

- (4) Show that the mapping $\mathbf{Q}(x) \rightarrow \mathbf{R}$ defined by $x \mapsto \pi$ provides an ordering of $\mathbf{Q}(x)$; compare this with the orderings obtained in Ex. (3).
- (5) Show that a skew field D can be ordered iff D^\times has a subgroup of index 2 closed under addition.

6.6 The field of real numbers

There are essentially two distinct ways of constructing the real numbers: (a) Dedekind's method of completion by cuts and (b) Cantor's method of sequences. Of these (a) is logically the simpler, while (b) corresponds more closely to the practical way of approximating real numbers. For us (b) has the further advantage of making it easier to define the algebraic operations; we shall therefore concentrate on (b). Later on, when we come to study valuations (in Ch. 8), we shall find that the same construction can be used for the p -adic numbers.

Let K be an ordered field. By a *null sequence* in K we understand a sequence $\{a_n\} (a_n \in K)$ such that for any $\varepsilon > 0$ in K there exists $n_0 \in \mathbb{N}$ such that $|a_n| < \varepsilon$ for all $n > n_0$. As usual in analysis we also say ' a_n converges to 0 as n tends to ∞ ' and write ' $a_n \rightarrow 0$ ' or ' $\lim a_n = 0$ '. A sequence $\{a_n\}$ is said to *converge to* $a \in K$, $a_n \rightarrow a$, if $a_n - a \rightarrow 0$, and we shall call a the *limit* of the sequence; it is clear that the limit, if it exists, is unique.

The following is a familiar necessary condition for the convergence of a sequence, which does not mention the value of the limit:

CAUCHY'S CONDITION *If a sequence $\{a_n\}$ converges to a limit, then the double sequence $\{c_{mn}\}$, where $c_{mn} = a_m - a_n$, converges to 0.*

In detail this means: given $\varepsilon > 0$, there exists n_0 such that $|a_m - a_n| < \varepsilon$ for all $m, n > n_0$. This condition follows easily from the definitions by the triangle inequality.

We shall call $\{a_n\}$ a *Cauchy sequence* if $|a_m - a_n| \rightarrow 0$; thus Cauchy's condition states that every convergent sequence is a Cauchy sequence. The converse need not hold, e.g. the rational numbers 1, 1.4, 1.41, 1.414, ... obtained in calculating successive approximations to $\sqrt{2}$ form a Cauchy sequence, but they are not convergent within \mathbf{Q} . The case where the converse holds is an important property of fields, and we make the

DEFINITION An ordered field is said to be *complete* if every Cauchy sequence is convergent.

Although an ordered field need not be complete, it always has a completion, constructed in a canonical fashion, as in the process that leads from \mathbf{Q} to \mathbf{R} . If K is

any ordered field, a subfield K' is called *dense in K* if between any two elements of K there is an element of K' . An embedding as a dense subfield is called a *dense embedding*.

THEOREM 6.1 *Let K be an ordered field. Then there exists a complete ordered field \tilde{K} and a dense order-embedding $\lambda: K \rightarrow \tilde{K}$ such that to each order-embedding $f: K \rightarrow L$ into a complete ordered field L there corresponds a unique order-embedding $f': \tilde{K} \rightarrow L$ such that $f = \lambda f'$.*

This is just the universal property of completions; to prove it in detail is somewhat lengthy (although not difficult). We therefore outline the main steps and leave some of the verifications to the reader.

Let R be the set of all Cauchy sequences in K ; this is a subset of $K^{\mathbb{N}}$, the set of all sequences in K . Now $K^{\mathbb{N}}$ is a ring, the direct power of K , and R is clearly a subring, the operations being carried out componentwise: $\{a_n\} \pm \{b_n\} = \{a_n \pm b_n\}$. Since every constant sequence is clearly a Cauchy sequence, we can regard K as a subfield of R by mapping $c \in K$ to the constant sequence c, c, \dots . Consider the set \mathfrak{n} of all null sequences in R ; we claim that this is an ideal in R . Let us show e.g. if $\{a_n\} \in R, \{b_n\} \in \mathfrak{n}$, then $\{a_n b_n\} \in \mathfrak{n}$. Since $\{a_n\}$ is a Cauchy sequence, there exists n_0 such that $|a_m - a_n| < 1$ for all $m, n > n_0$, hence $|a_m| < |a_{n_0}| + 1$ for $m > n_0$. It follows that $|a_m| \leq M$, where

$$M = \max \{|a_1|, |a_2|, \dots, |a_{n_0-1}|, |a_{n_0}| + 1\};$$

thus every Cauchy sequence is bounded. Now $b_n \rightarrow 0$, hence for any $\varepsilon > 0$, $|b_n| < \varepsilon/M$ for $n > n_1$, so $|a_n b_n| < \varepsilon$ for $n > n_1$, i.e. $a_n b_n \rightarrow 0$. This proves that $\{a_n b_n\} \in \mathfrak{n}$, and the remaining properties of \mathfrak{n} are established similarly.

We claim that \mathfrak{n} is a maximal ideal in R . For \mathfrak{n} is proper, e.g. the constant sequence $1, 1, \dots$ is a Cauchy sequence but not a null sequence. Moreover, if $\{a_n\}$ is a Cauchy sequence which is not null, then it has an inverse (mod \mathfrak{n}). To prove this fact let $\{a_n\}$ be not null; then by definition there exists $p \in K, p > 0$, such that to each n there corresponds $n' > n$ with $|a_{n'}| \geq p$. Since $\{a_n\}$ is Cauchy, there exists n_0 such that $|a_m - a_n| < p/2$ for all $m, n > n_0$. Choose any $n > n_0$ and take $n' > n$ as before; then

$$|a_n| \geq |a_{n'}| - |a_n - a_{n'}| > p/2.$$

Therefore the sequence a_n^{-1} is bounded for $n > n_0$. If we define

$$b_n = \begin{cases} 1 & \text{for } n \leq n_0, \\ a_n^{-1} & \text{for } n > n_0, \end{cases}$$

then $\{b_n\}$ is a Cauchy sequence and $\{a_n b_n\}$ converges to 1, i.e. $\{a_n\} \cdot \{b_n\} \equiv 1 \pmod{\mathfrak{n}}$. This shows \mathfrak{n} to be maximal.

Hence R/\mathbf{n} is a field, which we denote by \tilde{K} . We write λ for the natural homomorphism $K \rightarrow R \rightarrow R/\mathbf{n}$ obtained by mapping $a \in K$ to the residue class of the constant sequence a, a, \dots . Like every homomorphism between fields, this is an embedding, and we shall identify K with its image in \tilde{K} .

It is clear how to extend the ordering to \tilde{K} : given $\alpha \in \tilde{K}$, represented by a Cauchy sequence $\{\alpha_n\}$, either this is a null sequence, or there exists $\varepsilon > 0$ and n_0 such that $\alpha_n > \varepsilon$ for $n > n_0$, or there exists n_1 such that $\alpha_n < -\varepsilon$ for $n > n_1$. Moreover, all Cauchy sequences representing α have the same property and accordingly we set $\alpha = 0$, $\alpha > 0$ or $\alpha < 0$. We leave the reader to verify that K is dense in \tilde{K} . Now given a Cauchy sequence $\{\alpha_n\}$ in \tilde{K} , either α_n is constant from some n onwards; then it is a Cauchy sequence in K , and so it converges to a limit in \tilde{K} . Or $\{\alpha_n\}$ contains infinitely many distinct terms; in that case, omitting repetitions, we may assume all the α_n to be distinct. For each n we can choose $a_n \in K$ to lie between α_n and α_{n+1} ; the resulting sequence $\{a_n\}$ is easily seen to be a Cauchy sequence in K which has a limit α in \tilde{K} . Clearly $\lim \alpha_n = \lim a_n = \alpha$, and this shows \tilde{K} to be complete.

Finally let $f: K \rightarrow L$ be an order-embedding in a complete field L . Any element α of \tilde{K} is obtained as the limit of a Cauchy sequence $\{a_n\}$ in K ; it is easily seen that $\{a_n f\}$ is a Cauchy sequence in L , and so has a limit b , say. Any other Cauchy sequence tending to α differs from $\{\alpha_n\}$ by a null sequence, hence its image in L again tends to b ; therefore b depends only on α and we may put $\alpha f' = b$. Now the reader may verify without difficulty that f' is an order-embedding such that $f = \lambda f'$ and that it is the only mapping satisfying this equation. ■

The field \tilde{K} whose existence is proved in Th. 6.1 is called the *completion* of K ; since it is obtained as the solution of a universal problem, it is determined up to order-isomorphism by the properties listed in Th. 6.1. For example, if we take $K = \mathbf{Q}$, then $\tilde{K} = \mathbf{R}$.

The real numbers, as ordered field, have other important properties which can also be used to characterize them. Below we briefly look at two of them.

An ordered field K is said to be *archimedean* if for any $a \in K$ there exists $n \in \mathbf{N}$ such that $n > a$. We note that this is so precisely when \mathbf{Q} is dense in K . For if \mathbf{Q} is dense in K and $a \leq 0$ in K , then $1 > a$. If $a > 0$ in K , then $0 < (2a)^{-1} < a^{-1}$, hence there exists $\alpha \in \mathbf{Q}$ such that $(2a)^{-1} < \alpha < a^{-1}$, so if $\alpha = m/n$, then $a < n/m \leq n$; this shows K to be archimedean. Conversely, assume that K is archimedean and let $0 < a < b$ in K . Then $(b-a)^{-1} < n$ for some $n \in \mathbf{N}$, hence $0 < 1/n < b-a$; further, $na < m$ for some $m \in \mathbf{N}$. With the least such m we have $(m-1)/n \leq a$, hence $m/n \leq a + n^{-1} < a + (b-a) = b$, and so we have $a < m/n < b$. If $a < b < 0$, we can find $\alpha \in \mathbf{Q}$ between $-b$ and $-a$ and so $a < -\alpha < b$, while for $a < 0 < b$ we can take $\alpha = 0$. Thus

PROPOSITION 6.2 *An ordered field K is archimedean if and only if the prime subfield \mathbf{Q} is dense in K .* ■

In particular this result shows \mathbf{R} to be archimedean. If K is an ordered subfield of \mathbf{R} , then \mathbf{Q} is again dense in K , so K is archimedean. Conversely, if K is any archimedean ordered field, then \mathbf{Q} is dense in K , and now the proof of Th. 6.1 shows that there is an order-embedding of K in \mathbf{R} . Thus we have proved

THEOREM 6.3 *Any ordered subfield of \mathbf{R} is archimedean, and conversely, any archimedean ordered field is order-isomorphic to a subfield of \mathbf{R} .* ■

Clearly no proper subfield of \mathbf{R} is complete; we therefore have

COROLLARY 6.4 *Any complete archimedean ordered field is isomorphic to \mathbf{R} .* ■

A second important property of the real numbers is the upper bound property. In any partially ordered set S the supremum of any subset X , when it exists, is unique. It may or may not be a member of X , e.g. $\{0, -1, -2, \dots\}$ and $\{-1, -1/2, 1/3, \dots\}$ both have the supremum 0, which belongs to the first set but not to the second. A set is said to be *bounded above* if it has an upper bound.

In terms of upper bounds we can characterize the real numbers by the

UPPER BOUND PROPERTY *Every non-empty subset that is bounded above has a least upper bound.*

We observe that any ordered field possessing the upper bound property is archimedean. For if $\alpha \in K$ but $n \leq \alpha$ for all $n \in \mathbf{N}$, then the set $N = \{1, 2, \dots\}$ is bounded above and so has a supremum γ , say. It follows that $\gamma - 1 < \gamma$, hence $\gamma - 1 < n$ for some $n \in \mathbf{N}$, and so $\gamma < n + 1$, which contradicts the definition of γ . It follows that $n > \alpha$ for some $n \in \mathbf{N}$ and so K is archimedean, as asserted.

We can now give a characterization of the real numbers in terms of the upper bound property.

THEOREM 6.5 *The field \mathbf{R} of real numbers possesses the upper bound property, and any ordered field with the upper bound property is order-isomorphic to \mathbf{R} .*

Proof. Let X be any set of real numbers, bounded above. If X has a greatest element α , then $\alpha = \sup X$. Otherwise let Y be the set of all its upper bounds and X' the complement of Y in \mathbf{R} . Then $X \subseteq X'$ and the members of X' may be characterized as the numbers that are exceeded by some member of X . Moreover, every number in X' is less than every number in Y . Since they are complementary, we can for each $n = 1, 2, \dots$ find $a_{2n-1} \in X', a_{2n} \in Y$ such that

$$|a_{2n-1} - a_{2n}| < 1/n, \quad a_{2n-3} < a_{2n-1} < a_{2n} < a_{2n-2}.$$

Clearly $\{a_n\}$ is a Cauchy sequence; let α be its limit. If $n \in \mathbf{N}$, then $\alpha + 1/n$ is an

upper bound for X , for it exceeds some a_{2r} and so is in Y , while $\alpha - 1/n$ is less than some a_{2r-1} and so lies in X' . It follows that $\alpha = \sup X$, so the upper bound property holds for \mathbf{R} .

Now let K be any ordered field possessing the upper bound property. By the remark preceding the theorem, we see that K is archimedean, and hence, by Th. 6.3, order-isomorphic to a subfield of \mathbf{R} . We may therefore identify K with a subfield of \mathbf{R} ; clearly K contains \mathbf{Q} as subfield. But every element α of \mathbf{R} may be expressed as supremum of some subset A of \mathbf{Q} , and A also has a supremum α' say in K , by the upper bound property. Clearly $\alpha \leq \alpha'$; if the inequality were strict, we could find $a \in \mathbf{Q}$ such that $\alpha < a < \alpha'$, and this would contradict the fact that $\alpha' = \sup A$ in K . Hence $\alpha' = \alpha$, and it follows that $K = \mathbf{R}$. ■

Now we can establish the intermediate value property of the real numbers used in **6.8** of Vol. 1.

PROPOSITION 6.6 *Given a polynomial f in $\mathbf{R}[x]$ and $\alpha, \beta \in \mathbf{R}$ such that $\alpha < \beta$, if $f(\alpha)f(\beta) < 0$, then there exists $\gamma \in \mathbf{R}$ such that $\alpha < \gamma < \beta$ and $f(\gamma) = 0$.*

Proof. By hypothesis $f(\alpha), f(\beta)$ have opposite signs, and we may assume that $f(\alpha) < 0 < f(\beta)$, replacing f by $-f$ if necessary. Let A be the set of real numbers t such that $t < \beta$ and $f(t) < 0$. This set contains α , is bounded above (by β) and so has a supremum, γ say, where $\alpha \leq \gamma < \beta$. This means that for $\gamma < t < \beta$, $f(t) \geq 0$, but if $t < \gamma$, we can find t_1 such that $t < t_1 < \gamma$ and $f(t_1) < 0$. Now express $f(\gamma + x)$ as a polynomial in x :

$$f(\gamma + x) = a_0 x^n + \cdots + a_n, \quad \text{where } a_n = f(\gamma).$$

If $a_n \neq 0$, then the right-hand side has the same sign as a_n for sufficiently small x . But we have seen that $f(t)$ changes sign in every interval about γ , hence $a_n = 0$, i.e. $f(\gamma) = 0$. ■

The proof shows incidentally that f is continuous; this can of course also be established directly.

Exercises

- (1) Show that an archimedean ordered field has no order-preserving automorphism other than the identity.
- (2) Let K be an ordered field, algebraic over a subfield E . Show that if E is archimedean, then so is K .
- (3) Fill in the details omitted in the proof of Th. 6.1.

- (4) Adapt the proof of Th. 6.3 to prove that every archimedean ordered division ring is commutative.
- (5) Show that if an ordered field K has an archimedean subfield dense in K , then K is itself archimedean.
- (6) Show that the ordered fields constructed in Ex. (3) of 6.5 are not archimedean. Use the method of Ex. (4) of 6.5 to find uncountably many distinct archimedean orderings on $\mathbf{Q}(x)$.

Further exercises on Chapter 6

- (1) Let V be an n -dimensional inner product space over a field k of characteristic $\neq 2$. The *rank* of the quadratic form q on V is defined as $\dim V/V^\perp$. Show that q has rank r iff over a suitable extension of k , q can be reduced to the form $\langle a_1, \dots, a_r, 0^{n-r} \rangle$, where $a_i \neq 0$.
- (2) Show that the reflexion with respect to (U, U^\perp) , followed by that for (U^\perp, U) , where U is any regular subspace, is multiplication by -1 .
- (3) Show that the quadratic form on a hyperbolic plane (in characteristic $\neq 2$) is universal.
- (4) Show that a hyperbolic plane in characteristic 2 is totally isotropic. Determine the orthogonal group in this case.
- (5) Prove the analogue of Witt's cancellation theorem (Th. 4.1) for alternating bilinear forms.
- (6) Let k be a field of characteristic $\neq 2$ and a an element of k not a square. Write $E = k(\alpha)$, where α is a root of $x^2 = a$. Show that the proper orthogonal group of the k -space $\langle 1, a \rangle$ is isometric to the group of elements of E of norm 1, modulo the group of squares of k . Deduce that $\mathbf{SO}(V)$ is abelian. Show that $\mathbf{O}_2(k)$ is non-abelian.
- (7) Let V be a regular inner product space and u any non-zero vector in V . Show that the vectors $u\theta(\theta \in \mathbf{O}(V))$ span V except when $k = \mathbf{F}_3$ and $\dim V = 2$.
- (8) Prove the uniqueness of the limit of a convergent sequence from the definitions.
- (9) In any ordered field K the subset $A = \{a \in K \mid |a| \leq n \text{ for some } n \in \mathbf{N}\}$ is a subring whose non-units form an ideal \mathfrak{m} , the set of inverses of elements of $K \setminus A$. Show that A/\mathfrak{m} has a natural ordering and is isomorphic to a subfield of \mathbf{R} .
- (10) Given a polynomial f over \mathbf{R} and $a < b$ in \mathbf{R} , show that there exists $c \in \mathbf{R}$, $a < c < b$, such that $f(b) - f(a) = (b - a)f'(c)$.

7

Representation theory of finite groups

Although much of the theory of finite-dimensional algebras had its origins in the theory of group representations, it seems simpler nowadays to develop the theory of algebras first and then use it to give an account of group representations. This theory has been a powerful tool in the deeper study of groups, especially the modular theory (representations over a field of finite characteristic), which has played a key role in the classification of finite simple groups. The theory also has important applications to physics: quantum mechanics describes physical systems by means of states which are represented by vectors in Hilbert space (= infinite-dimensional complete unitary space). Any group which may act on the system, such as the rotation group or a permutation group of the constituent particles, acts by unitary transformations on this Hilbert space and any finite-dimensional subspace admitting the group leads to a representation of the group. If we know the irreducible representations of our group, this will often allow us to classify these spaces.

Of course an introductory chapter like the present one is not the place to develop modular representations, nor the applications to physics. The plan of the chapter is as follows. The first four sections give a concise account of the theory, based on the Wedderburn theorems of Ch. 5; they include the basic results on orthogonality and completeness and explain the role of characters. Over the complex numbers some simplifications can be made and they are described next. The rest of the chapter deals with representations and characters of the symmetric group, and describes induced representations, an important technique which is illustrated by the theorems of Frobenius and Burnside.

7.1 Basic definitions

Let G be any group (not necessarily finite). By a *representation* of G over a field k one understands a homomorphism

$$\rho: G \rightarrow \mathbf{GL}_d(k), \tag{1}$$

where $\mathbf{GL}_d(k)$ is the *general linear group* of degree d over k , i.e. the group of all

invertible $d \times d$ matrices over k . Thus we have a mapping $x \mapsto \rho(x)$ such that

$$\rho(xy) = \rho(x)\rho(y) \quad \text{for all } x, y \in G. \quad (2)$$

Since each matrix $\rho(x)$ is invertible, we have $\rho(1) = I$ and $\rho(x^{-1}) = \rho(x)^{-1}$. The integer d is called the *degree* of the representation.

For example, to find a representation of the cyclic group $\mathbf{C}_3 = \{1, t, t^2\}$ over \mathbf{R} of degree 2, we need to find $A \in \mathbf{GL}_2(\mathbf{R})$ such that $A^3 = I$. We may take $A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ and then ρ is defined by

$$\rho(t) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad \rho(t^2) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \quad \rho(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3)$$

Every group has the *trivial representation*, obtained by mapping every element of G to I . At the other extreme we have the *faithful* representations, defined as homomorphisms with trivial kernel; e.g. (3) is a faithful representation of \mathbf{C}_3 .

Two representations ρ, σ of a group G are said to be *equivalent*, if they have the same degree, d say, and there exists $P \in \mathbf{GL}_d(k)$ such that

$$\sigma(x) = P^{-1}\rho(x)P \quad \text{for all } x \in G. \quad (4)$$

It is clear that this is indeed an equivalence relation on the set of all representations of G . For example, if ω is a primitive cube root of 3, then

$$\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -\omega & -\omega^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -\omega & -\omega^2 \end{pmatrix} \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix};$$

therefore the representation ρ of \mathbf{C}_3 given by (3) is equivalent to σ , where σ is given by

$$\sigma(t) = \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad \sigma(t^2) = \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega \end{pmatrix}, \quad \sigma(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

If we interpret the matrices of a representation of G as linear transformations of a vector space we reach the notion of a G -module. A *G -module* is a vector space V over k such that each $x \in G$ defines a linear mapping $v \mapsto vx$ on V satisfying

$$v(xy) = (vx)y, \quad v1 = v, \quad \text{for all } v \in V, x, y \in G. \quad (5)$$

Given two G -modules U and V , a *homomorphism* or *G -homomorphism* from U to V is a k -linear mapping $f: U \rightarrow V$ such that

$$f(ux) = (fu)x, \quad \text{for all } u \in U, x \in G. \quad (6)$$

If a homomorphism from U to V is bijective, its inverse is easily seen to be a homomorphism from V to U ; this is called an *isomorphism* or *G -isomorphism*. We say that U and V are *isomorphic* and write $U \cong V$.

To establish the link between representations of G and G -modules, let us take a

finite-dimensional G -module V , with basis v_1, \dots, v_d over k . The action of G on V is completely described by the equations

$$v_i x = \sum_j \rho_{ij}(x) v_j \quad (x \in G), \quad (7)$$

where $\rho_{ij}(x) \in k$, and it is easily checked that the matrices $\rho(x) = (\rho_{ij}(x))$ form a representation of G ; we shall say that the G -module V with the basis v_1, \dots, v_d affords the representation ρ . Conversely, given a representation $\rho = (\rho_{ij})$ of G of degree d and a d -dimensional vector space V over k , we can turn V into a G -module by defining the action of $x \in G$ on a basis v_1, \dots, v_d by (7) and generally putting

$$(\sum_i \alpha_i v_i) x = \sum_i \alpha_i \rho_{ij}(x) v_j.$$

The verification that this provides a G -module is straightforward and may be left to the reader.

To see that the operations of passing between representations and modules are mutually inverse we need to examine the effect of a change of basis on the representation. Let V be a G -module affording the representation ρ relative to a basis v_1, \dots, v_d . Thus the equations (7) hold, which may be written concisely as

$$vx = \rho(x)v, \quad (8)$$

where $v = (v_1, \dots, v_d)^T$ stands for the column of basis vectors v_1, \dots, v_d . Suppose that $u = (u_1, \dots, u_d)^T$ is a second basis of V , affording the representation σ , so that

$$ux = \sigma(x)u. \quad (9)$$

If the matrix of transformation from u to v is denoted by P , we have

$$v = Pu, \quad u = P^{-1}v. \quad (10)$$

Hence

$$\sigma(x)u = ux = (P^{-1}v)x = P^{-1}(vx) = P^{-1}\rho(x)v = P^{-1}\rho(x)Pu.$$

It follows that

$$\sigma(x) = P^{-1}\rho(x)P; \quad (11)$$

thus ρ and σ are equivalent, and what we have shown is that different bases of a G -module afford equivalent representations. Moreover, since P may be any invertible matrix, we see that representations of G that are equivalent are afforded by the same G -module, for suitable bases. Further, if two G -modules afford the same representation, they must be isomorphic. For take the modules to be V, W with bases $v = (v_1, \dots, v_d)^T$, $w = (w_1, \dots, w_d)^T$ and let

$$vx = \rho(x)v, \quad wx = \rho(x)w.$$

Then the mapping $\sum_i \alpha_i v_i \mapsto \sum_i \alpha_i w_i$ is easily verified to be a G -isomorphism between V and W . It follows that isomorphic modules afford equivalent representations, for by changing the bases in one of the modules we can make the representations equal. Thus we have proved

PROPOSITION 1.1 *For any group G there is a natural bijection between equivalence classes of representations and isomorphism classes of G -modules.* ■

By this result we can use G -modules and representations of G interchangeably; we shall examine various concepts from both points of view, but first we must clarify the connexion with modules over a ring, discussed in Ch. 4. In order to do this we shall use the notion of a group algebra introduced in 5.1.

Let G be any group and kG its group algebra over k . We can form modules over kG , as for any ring, and it is clear that a kG -module is also a G -module. Conversely, a G -module V becomes a kG -module by the rule

$$u\left(\sum a_x x\right) = \sum a_x ux, \quad (u \in V),$$

where the summation is over $x \in G$. Given a G -module V , we can define a submodule as for modules over a ring, as a subspace V' of V admitting the G -action, i.e. such that $vx \in V'$ for $v \in V'$, $x \in G$. Alternatively we may regard V as kG -module and look for its kG -submodules; clearly they are just the G -submodules of V . Likewise the homomorphisms between G -modules are nothing other than the module homomorphisms between kG -modules.

Let us now examine the form taken by the representations corresponding to submodules. We consider a G -module V with a submodule V' . If v_1, \dots, v_d is a basis of V , adapted to V' , say v_1, \dots, v_t ($t \leq d$) is a basis of V' and the action is given by (8), then for $i \leq t$ we have $v_i x \in V'$ and so $\rho_{ij}(x) = 0$ for $i \leq t < j$. Thus ρ has the form

$$\rho(x) = \begin{pmatrix} \rho'(x) & 0 \\ \theta(x) & \rho''(x) \end{pmatrix}. \quad (12)$$

We note that ρ' is a representation afforded by V' , while ρ'' is afforded by the quotient module V/V' relative to the basis $\bar{v}_{t+1}, \dots, \bar{v}_d$, where \bar{u} denotes the residue class of u . Both ρ' and ρ'' are sometimes called *subrepresentations* of ρ .

We note that if the basis of V is chosen so that the last t members form a basis of V' (instead of the first t), then ρ takes the form

$$\rho(x) = \begin{pmatrix} \rho''(x) & \theta(x) \\ 0 & \rho'(x) \end{pmatrix}.$$

A representation is said to be *reducible* if it is equivalent to a representation of the form (12), where $0 < t < d$. Thus ρ is reducible iff the corresponding G -module V has a non-zero proper submodule, i.e. it is not simple. In the contrary case, when V is simple, the representation is called *irreducible*.

If ρ can be written in the form of a diagonal sum:

$$\rho(x) = \begin{pmatrix} \rho'(x) & 0 \\ 0 & \rho''(x) \end{pmatrix},$$

it is said to be *completely reduced*. Clearly this corresponds to V being directly decomposable. We observe that any finite-dimensional G -module V has a composition series:

$$V = V_0 \supset V_1 \supset V_2 \supset \cdots \supset V_r = 0,$$

such that V_{i-1}/V_i is simple. The corresponding representation can then be taken in the form

$$\rho(x) = \begin{pmatrix} \rho_1(x) & & & 0 \\ & \rho_2(x) & \dots & \\ * & & & \rho_r(x) \end{pmatrix}. \quad (13)$$

If ρ is *completely reducible*, i.e. we can find an equivalent representation of the form (13) in which $*$ = 0, this means that the corresponding G -module is a direct sum of simple modules, i.e. *semisimple*.

Just as for modules over a ring we can define *left* G -modules; they are vector spaces V with a G -action $v \mapsto xv$ such that

$$x(yv) = (xy)v, \quad 1v = v.$$

However, any such left G -module may be regarded as a right G -module by defining $v.x = x^{-1}v$ ($x \in G$). For we have $v.(xy) = (xy)^{-1}v = (y^{-1}x^{-1})v = y^{-1}(x^{-1}v) = y^{-1}(v.x) = (v.x).y = v.(xy)$.

In terms of the group algebra this may be expressed as follows: the group algebra kG has an antiautomorphism $*$, i.e. a linear mapping satisfying $(ab)^* = b^*a^*$, given by

$$(\sum a_x x)^* = \sum a_x x^{-1}. \quad (14)$$

Now any left kG -module V becomes a right kG -module on putting $v.a = a^*v$.

We remark that the mapping defined by (14) has the property $a^{**} = a$. An antiautomorphism whose square is the identity is called an *involution*; thus kG is an algebra with an involution.

Exercises

- (1) Let G be a group and consider the regular representation of kG (defined by right multiplication). Show that this representation always has the trivial representation $x \mapsto 1$ as a subrepresentation.
- (2) Let k be a field containing a primitive n th root of 1, say ω (hence of characteristic 0 or prime to n) and let C_n be the cyclic group of order n , with generator t . Show that $\rho_k: t^r \mapsto \omega^k$ is a representation of degree 1 of C_n for $k = 0, 1, \dots, n - 1$. Show that if ρ is any representation of C_n over k , then $\rho(t)$ can be transformed to diagonal form. Deduce that ρ can be written as a direct sum of copies of the ρ_k .
- (3) Show that if ρ is any representation of a group G , then $\check{\rho}$ defined as $\check{\rho}(x) = \rho(x^{-1})^T$ is

again a representation of G (it is called the *contragredient* of ρ). What are the conditions for $\check{\rho}$ to coincide with ρ ?

(4) Let G be a group and ρ, δ representations of G , where δ is of degree 1. Show that $\delta\rho$ is again a representation of G , and this is irreducible iff ρ is.

(5) Let ρ be an irreducible representation of a (finite) p -group G , acting on a vector space V over a field k of characteristic p . Show that V contains a vector $v \neq 0$ which is fixed under the action of G . Show that the matrices $\rho(x)$ ($x \in G$) have a common eigenvector and deduce that ρ must be the trivial representation. (Hint. Use the fact that $\rho(x) - 1$ is nilpotent; a matrix $\rho(x)$ with this property is called *unipotent*.)

(6) Let G be a p -group and k a field of characteristic p . Suppose that G acts transitively on a finite set S and let V be the k -space on S as basis, with the G -action defined by the permutations of S . Show that V has a unique maximal submodule; deduce that V is indecomposable, but not simple, unless it is one-dimensional

(7) Let ρ be an irreducible representation of a finite group G over a field of characteristic p . Show that if the degree of ρ is d , then any normal subgroup of index p^d lies in the kernel of ρ . Deduce that if ρ is faithful, then G has a non-trivial normal p -subgroup.

7.2 The averaging lemma and Maschke's theorem

For a closer study of representations we need to assume that our group is finite and we shall make this assumption from now on. The first important fact to note is that for a finite group every representation over a field of characteristic 0 is completely reducible.

THEOREM 2.1 (Maschke's theorem, 1899) *Let G be a finite group and k a field of characteristic 0 or not dividing $|G|$. Then every representation of G over k is completely reducible.*

Proof. Let ρ be a representation of G and suppose that ρ is reduced:

$$\rho(x) = \begin{pmatrix} \rho'(x) & 0 \\ \theta(x) & \rho''(x) \end{pmatrix}, \quad (1)$$

where ρ' , ρ'' are subrepresentations of degrees d' , d'' respectively. To establish complete reducibility it is enough to find a $d'' \times d'$ matrix μ such that

$$\begin{pmatrix} \rho'(x) & 0 \\ \theta(x) & \rho''(x) \end{pmatrix} \begin{pmatrix} I & 0 \\ \mu & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ \mu & I \end{pmatrix} \begin{pmatrix} \rho'(x) & 0 \\ 0 & \rho''(x) \end{pmatrix}.$$

When we multiply out, only the $(2, 1)$ -block gives anything new:

$$\theta(x) = \mu\rho'(x) - \rho''(x)\mu, \quad (2)$$

and we shall complete the proof by finding a matrix μ to satisfy this equation. By substituting from (1) in the relation $\rho(xy) = \rho(x)\rho(y)$ we obtain the following equation for $\theta(x)$:

$$\theta(xy) = \theta(x)\rho'(y) + \rho''(x)\theta(y). \quad (3)$$

Writing $|G| = m$ and noting that $m \neq 0$ in k , by hypothesis, we have

$$\begin{aligned} m\theta(x) &= \sum_y \theta(x)\rho'(y)\rho'(y^{-1}) \\ &= \sum_y [\theta(xy) - \rho''(x)\theta(y)]\rho'(y^{-1}). \end{aligned}$$

Put $z = xy$; then $y^{-1} = z^{-1}x$, and as y runs over G , so does z , for fixed x . Hence we can rewrite the last sum as

$$\sum_z \theta(z)\rho'(z^{-1}x) - \sum_y \rho''(x)\theta(y)\rho'(y^{-1}),$$

so

$$m\theta(x) = \sum_z \theta(z)\rho'(z^{-1})\rho'(x) - \sum_y \rho''(x)\theta(y)\rho'(y^{-1}),$$

and this has the form (2), if we abbreviate $m^{-1}\sum_y \theta(y)\rho'(y^{-1})$ as μ . ■

In view of its importance we shall give a second proof of this result, or rather, restate the same proof in module terms. The essential step is a lemma which is also useful elsewhere, but we shall first need to introduce some notation. If U, V are G -modules and α is a mapping from U to V , we shall write $\alpha: U \xrightarrow{k} V$, $\alpha: U \xrightarrow{G} V$ to indicate that α is k -linear or a G -homomorphism respectively. The space of all k -linear mappings from U to V is denoted by $\text{Hom}_k(U, V)$ and the subspace of G -homomorphisms by $\text{Hom}_G(U, V)$.

In the next lemma we shall (exceptionally) write mappings between right G -modules on the right, so for $\alpha: U \xrightarrow{k} V$ the condition for a G -homomorphism is that

$$(ux)\alpha = (u\alpha)x \quad \text{for all } u \in U, x \in G.$$

LEMMA 2.2 (Averaging lemma) *Let G be a finite group and k a field of characteristic 0 or prime to $|G|$. Given any two G -modules U, V and $\alpha: U \xrightarrow{k} V$, the mapping*

$$\alpha^*: u \mapsto |G|^{-1} \sum_x ((ux^{-1})\alpha)x \quad (4)$$

is a G -homomorphism from U to V . Moreover,

- (i) *if α is a G -homomorphism, then $\alpha^* = \alpha$,*
- (ii) *if $\alpha: U \xrightarrow{k} V$, $\beta: V \xrightarrow{G} W$, then $(\alpha\beta)^* = \alpha^*\beta$,*
- (iii) *if $\alpha: U \xrightarrow{G} V$, $\beta: V \xrightarrow{k} W$, then $(\alpha\beta)^* = \alpha\beta^*$.*

Proof. Let us fix $a \in G$ and write $y = xa, x = ya^{-1}$. Then as one of x, y runs over G , so does the other. Now for $\alpha: U \xrightarrow{k} V$, we have

$$|G|.ux^*a = \sum_x ux^{-1}\alpha xa = \sum_y uay^{-1}\alpha y = |G|.ua\alpha^*. \quad (5)$$

This shows α^* to be a G -homomorphism. If α is a G -homomorphism, each term in the sum in (5) is $uxa = u\alpha a$, so $\alpha^* = \alpha$ in this case and (i) follows. Now let $\beta: V \xrightarrow{G} W$; then

$$|G|.u(\alpha\beta)^* = \sum_x ux^{-1}\alpha\beta x = \sum_x ux^{-1}\alpha x\beta = |G|.u\alpha^*\beta.$$

Hence (ii) follows; (iii) is proved similarly. ■

We note that if neither α nor β is a G -homomorphism, there is nothing we can say. We can now prove the module form of Maschke's theorem, which states that every module extension splits, or equivalently, that the group algebra kG is semisimple.

THEOREM 2.3 (Maschke's theorem, form 2) *Let G be a finite group and k a field of characteristic 0 or prime to $|G|$. Then kG is semisimple.*

Proof. We shall show that every (finite-dimensional) G -module is semisimple, or equivalently, that every short exact sequence of G -modules

$$0 \longrightarrow V' \xrightarrow{\alpha} V \xrightarrow{\beta} V'' \longrightarrow 0 \quad (6)$$

splits. Such a sequence certainly splits as a sequence of k -spaces; for this just means that V' as k -subspace of V has a vector space complement. Thus we have a k -linear splitting map $\gamma: V \rightarrow V'$. We have $\alpha\gamma = 1_{V'}$; therefore $1 = 1^* = (\alpha\gamma)^* = \alpha\gamma^*$, and so γ^* is the desired G -homomorphism splitting the sequence (6). ■

Exercises

- (1) Let G be a finite group and V a finite-dimensional G -module over a field of characteristic prime to $|G|$. Show that if G acts trivially on every simple composition factor of V then the G -action on V is trivial.
- (2) For $G = \mathbf{C}_p$ and $k = \mathbf{F}_p$ define a two-dimensional space V with basis v_1, v_2 as G -module by $v_1t = v_1 + v_2, v_2t = v_2$, where t is a generator of \mathbf{C}_p . Verify that V is not semisimple; calculate the corresponding representation.
- (3) Show that the infinite cyclic group has, over a field of characteristic 0, a faithful two-dimensional representation which is not completely reducible.

(4) Let G be a finite group and k a field of characteristic dividing $|G|$. Show that the element $z = \sum_x x$ is central and nilpotent; deduce that kG not semisimple.

(5) Let k be a field of characteristic p and G a finite group. Show that for any element g of p -power order $g - 1$ is nilpotent in kG . Deduce that for a finite p -group G the radical of kG is the augmentation ideal. (*Hint.* Find a basis of nilpotent elements for the radical and use Th. 5.6.4).

7.3 Orthogonality and completeness

The representation theory of finite groups was developed by Frobenius in the 1880s and 1890s, using the determinant of the matrix of a general group element in the regular representation. I. Schur in his dissertation (1901) greatly simplified the theory using his lemma (in the form given below) and the averaging lemma 2.2. We begin by recalling Schur's lemma in a slightly more explicit form.

LEMMA 3.1 (Schur's lemma for algebraically closed fields) *Let k be an algebraically closed field, A a k -algebra and U, V any two simple A -modules, finite-dimensional over k . Then*

$$\text{Hom}_A(U, V) = \begin{cases} k & \text{if } U \cong V, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Proof. Let $\alpha: U \rightarrow V$ be a non-zero homomorphism. Since U, V are simple, we have $\ker \alpha = 0$, $\text{im } \alpha = V$, so α is an isomorphism, and this proves the second line. Now the first case follows as in Lemma 5.3.5. ■

We remark that the argument used in the above proof will also show that any division algebra A over an algebraically closed field k is just k itself. For A is simple as left A -module, finite-dimensional over k , and we can write $A \cong \text{End}_A(A)$ by Th. 5.1.3, so $A \cong k$ by Lemma 3.1.

Let G be a finite group; we shall define an inner product on the group algebra kG by the rule

$$(\sum a(x)x, \sum b(y)y) = |G|^{-1} \sum a(x^{-1})b(x). \quad (2)$$

It is clear that this product is bilinear; it is not symmetric, but satisfies the equation

$$(g, f) = (f^*, g^*), \quad (3)$$

where $*$ is the involution defined by (14) of 7.1. The product is regular, i.e. non-degenerate: if $(f, x) = 0$ for all $x \in G$, then $f(x^{-1}) = 0$ for all $x \in G$ and so $f = 0$. Of course from the point of view of the inner product (2) the multiplication on kG is immaterial, and kG may be thought of as the space of all functions on G .

Our next aim is to show that the different representation coefficients, regarded as functions on G , are orthogonal.

THEOREM 3.2 (Orthogonality relations for representations) *Let G be a finite group and k an algebraically closed field of characteristic 0 or prime to $|G|$. If ρ, σ are irreducible representations of degrees c, d respectively, then we have the relations*

$$|G|^{-1} \sum_{x \in G} \rho_{ij}(x^{-1}) \sigma_{pq}(x) = \begin{cases} 0 & \text{if } \rho \text{ and } \sigma \text{ are inequivalent,} \\ d^{-1} \delta_{jp} \delta_{iq} & \text{if } \rho = \sigma. \end{cases} \quad (4)$$

Thus different representation coefficients are orthogonal. We note that the alternatives on the right of (4) are not exhaustive: the representations ρ, σ may be equivalent but distinct. In that case (4) will not apply (but of course we can use (4) even then, after transforming one of ρ, σ into the other).

Proof. Take spaces U, V affording ρ, σ with bases $u_1, \dots, u_c, v_1, \dots, v_d$ and let $\alpha_{jp}: U \xrightarrow{k} V$ be the linear mapping defined by

$$u_i \alpha_{jp} = \delta_{ij} v_p.$$

Explicitly the (i, q) -entry of the matrix for α_{jp} is

$$(\alpha_{jp})_{iq} = \delta_{ij} \delta_{pq}; \quad (5)$$

by Lemma 2.2, α_{jp}^* is a G -homomorphism from U to V , and its matrix is given by applying $*$ to (5):

$$\begin{aligned} |G| \cdot (\alpha_{jp}^*)_{iq} &= \sum_{h,r,x} \rho_{ih}(x^{-1}) \delta_{hj} \delta_{pr} \sigma_{rq}(x) \\ &= \sum_x \rho_{ij}(x^{-1}) \sigma_{pq}(x). \end{aligned}$$

If ρ, σ are inequivalent, then $\alpha_{jp}^* = 0$ by Lemma 3.1, and this proves the first line of (4).

Next we take $\rho = \sigma$. By Lemma 3.1, $\alpha_{jp}^* = \lambda_{jp} \in k$, hence we have

$$\sum_x \rho_{ij}(x^{-1}) \rho_{pq}(x) = |G| \cdot \lambda_{jp} \delta_{qi}. \quad (6)$$

To find λ_{jp} we put $q = i$ and sum over i :

$$|G| \cdot d \cdot \lambda_{jp} = \sum_{i,x} \rho_{pi}(x) \rho_{ij}(x^{-1}) = \sum_x \rho_{pj}(1) = |G| \cdot \delta_{jp}.$$

By the hypothesis on k we can divide by $|G|$, hence $d \neq 0$ in k and $\lambda_{jp} = d^{-1} \delta_{jp}$. Inserting this value in (6) we obtain the second line of (4). ■

To illustrate the result, let us take the trivial representation for σ . Then $\sigma(x) = 1$

for all $x \in G$ and we find that every non-trivial irreducible representation ρ of G satisfies

$$\sum_x \rho_{ij}(x) = 0 \quad \text{for all } i, j. \quad (7)$$

More generally, if ρ is any representation, we can take it to be completely reduced, and we then see that (7) holds precisely when ρ does not contain the trivial representation.

In terms of the inner product (2) the relations (4) may be written

$$(\rho_{ij}, \sigma_{pq}) = \frac{1}{d} \delta_{jp} \delta_{iq} \quad \text{or } 0;$$

thus after normalizing them we obtain an orthonormal set of functions; in particular their number cannot exceed $\dim kG = |G|$. Hence we have

COROLLARY 3.3 *The coefficients of inequivalent irreducible representations of G are linearly independent; hence the number of such representations is finite, and if their degrees are d_1, \dots, d_t then*

$$\sum d_i^2 \leq |G|. \quad \blacksquare \quad (8)$$

Our next task is to show that equality holds in (8) if we take enough representations. This means that every k -valued function on G can be written as a linear combination of irreducible representation coefficients; this is expressed by saying that these coefficients form a *complete* system of functions on G .

To see this, let us go back to the group algebra kG . We have seen that this is semisimple, hence a direct product of full matrix rings over skew fields. But as we saw, the only skew field finite-dimensional over k is k itself, because k is algebraically closed. Thus kG is a direct product of full matrix rings over k :

$$kG \cong \prod_{i=1}^t \mathfrak{M}_{d_i}(k). \quad (9)$$

Here each factor provides an irreducible representation of G , of degree d_i , and these representations are inequivalent because the product is direct, so the coefficients corresponding to different factors are linearly independent. On counting dimensions in (9) we obtain the desired equality

$$\sum d_i^2 = |G|. \quad (10)$$

We can also take the regular representation of G , i.e. we take kG as G -module under right multiplication by G . Each irreducible representation ρ_i occurs d_i times, corresponding to the d_i rows of the corresponding matrix. Thus we again obtain the equation (10).

Let G be a group and U, V any G -modules. Then $U \otimes V$ may be defined as a G -

module by the equation

$$(u \otimes v)g = ug \otimes vg.$$

Since the right-hand side is bilinear in u and v , this defines an action and $U \otimes V$ is easily verified to be a G -module. If the representations afforded by U, V are ρ, σ relative to bases $u_1, \dots, u_m, v_1, \dots, v_n$ respectively, then

$$(u_i \otimes v_p)g = \sum \rho_{ij}(g)\sigma_{pq}(g)u_j \otimes v_q;$$

hence the representation of G afforded by $U \otimes V$ is the tensor product of the matrices: $\rho \otimes \sigma$. When G is finite, then by Maschke's theorem, each representation is a diagonal sum of irreducible ones, and if ρ_1, \dots, ρ_r are the inequivalent irreducible representations of G , it is enough to determine the products $\rho_i \otimes \rho_j$. We have

$$\rho_i \otimes \rho_j = \sum_k g_{ijk} \rho_k, \quad (11)$$

where the g_{ijk} are non-negative integers, indicating how often a given representation ρ_k occurs in the tensor product.

For each representation ρ of G we define its *kernel* as

$$K = \{x \in G \mid \rho(x) = I\}.$$

Thus K is a normal subgroup of G and G has a faithful irreducible representation iff some irreducible representation has trivial kernel. This need not be so, but in any case we have

THEOREM 3.4 *For any finite group G the intersection of the kernels of all the irreducible representations is trivial.*

For the regular representation of G is faithful; since it is a direct sum of irreducible representations, the conclusion follows. ■

Exercises

- (1) Find all irreducible representation of S_3 by reducing the regular representation.
- (2) Show that for any representation ρ of a group G the set $N = \{x \in G \mid \det \rho(x) = 1\}$ is a normal subgroup of G with cyclic quotient group.
- (3) Let G be a finite group, k an algebraically closed field of characteristic 0 and ρ, σ inequivalent irreducible representations of G of degrees c, d . Show that for any $c \times d$ matrix T , $\sum_x \rho(x^{-1})T\sigma(x) = 0$. Further show that for a $d \times d$ matrix T , $\sum_x \sigma(x^{-1})T\sigma(x) = d^{-1} \cdot |G| \cdot \text{Tr}(T)I$.
- (4) Show that if ρ_1, \dots, ρ_r are irreducible pairwise inequivalent representations of a group

and $\rho = \bigoplus c_i \rho_i$, then the centralizer of ρ has dimension $\sum c_i^2$. Use this fact to obtain another proof of (10).

(5) Let G be a finite group and let d be the degree of an irreducible representation of G over \mathbf{Z} (i.e. a homomorphism $G \rightarrow \mathbf{GL}_d(\mathbf{Z})$). Show that every prime dividing d must also divide $|G|$.

7.4 Characters

A one-dimensional representation is also called a *linear character* or simply a *character* if we are dealing with an abelian group. Such characters have already been discussed in 4.6, and the results proved there will be used freely. We recall that an irreducible representation of an abelian group over \mathbf{C} (an algebraically closed field) is necessarily one-dimensional, by Schur's lemma. For a non-abelian group there will always be irreducible representations of degrees greater than 1 (cf. Prop. 4.2 below) and the definition then runs as follows. Given any representation ρ of a group G over \mathbf{C} , its *character* is defined as

$$\chi(x) = \operatorname{tr} \rho(x), \quad x \in G, \quad (1)$$

where tr denotes the trace of the matrix $\rho(x)$; thus if $\rho(x) = (\rho_{ij}(x))$, then $\operatorname{tr} \rho(x) = \sum_i \rho_{ii}(x)$. When χ and ρ are related as in (1), ρ is said to *afford* the character χ . For example, any representation of degree 1 is its own character; in particular, the function $\chi_1(x) = 1$ for all $x \in G$ is the character afforded by the trivial representation, and is called the *trivial* or also the *principal* character.

Some obvious properties of characters are collected in

PROPOSITION 4.1 *The character of a representation is independent of the choice of basis, i.e. equivalent representations have the same character. Moreover, each character is a class function on G , i.e. it is constant on conjugacy classes:*

$$\chi(y^{-1}xy) = \chi(x), \quad x, y \in G. \quad (2)$$

The degree of χ is $\chi(1)$ and for any element x in G of order n , $\chi(x)$ is a sum of n th roots of 1.

If ρ_1, ρ_2 are any representations with the characters χ_1, χ_2 then the characters afforded by $\rho_1 \oplus \rho_2$ and $\rho_1 \otimes \rho_2$ are $\chi_1 + \chi_2$ and $\chi_1 \chi_2$ respectively.

Proof. Let χ be the character of ρ ; any equivalent representation has the form $T^{-1}\rho(x)T$ and since $\operatorname{tr}(BA) = \operatorname{tr}(AB)$ for any square matrices of the same size, we have

$$\operatorname{tr}(T^{-1}\rho(x)T) = \operatorname{tr} \rho(x),$$

so both ρ and $T^{-1}\rho T$ afford the same character. For the same reason, $\operatorname{tr} \rho(y^{-1}xy) = \operatorname{tr}(\rho(y)^{-1}\rho(x)\rho(y)) = \operatorname{tr} \rho(x)$, and (2) follows. $\chi(1)$ equals the degree

because $\text{char } \mathbf{C} = 0$, while $A = \rho(x)$ satisfies $A^n = I$ if $x^n = 1$. Thus A satisfies an equation with distinct roots and so can be transformed to diagonal form over \mathbf{C} ; its diagonal elements λ again satisfy $\lambda^n = 1$, so they are n th roots of 1, and $\chi(x)$ is the sum of these diagonal elements.

The final assertions follow because $\text{tr}(A \oplus B) = \text{tr } A + \text{tr } B$ and $\text{tr}(A \otimes B) = \text{tr } A \cdot \text{tr } B$. ■

The next result may be regarded as a generalization of Th. 4.6.1 on the duality of abelian groups.

PROPOSITION 4.2 *For any finite group G the number of linear characters is $(G:G')$, where G' is the derived group. Hence every non-abelian group has irreducible representations of degree greater than 1.*

Proof. Every homomorphism $\alpha: G \rightarrow \mathbf{C}^\times$ corresponds to a homomorphism from G/G' to \mathbf{C}^\times and conversely. But as we have seen in 4.6, the number of such homomorphisms is $|\widehat{G/G'}| = |G/G'| = (G:G')$. ■

In 7.3 we defined an inner product on kG ; we shall now see how in the case of $k = \mathbf{C}$ we can define a hermitian inner product on $\mathbf{C}G$. Let us put

$$(f, g) = |G|^{-1} \sum_x \overline{f(x)}g(x). \quad (3)$$

Since every character α is a sum of roots of 1, we have $\overline{\alpha(x)} = \alpha(x^{-1})$. Hence for characters the formula (3) can also be written

$$(\alpha, \beta) = |G|^{-1} \sum_x \alpha(x^{-1})\beta(x); \quad (4)$$

thus in this case it agrees with the inner product introduced in 7.3. From the orthogonality relations in Th. 3.2 we obtain the following orthogonality relations for irreducible characters, by putting $j = i$, $q = p$ in (4) of 7.3 and summing over i and p :

$$(\chi, \psi) = \begin{cases} 1 & \text{if } \chi = \psi, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Thus under the metric (4) the irreducible characters form an orthonormal system. Suppose that all the irreducible representations of G are ρ_1, \dots, ρ_r with characters χ_1, \dots, χ_r . Any representation of G is equivalent to a direct sum of irreducible ones, by complete reducibility (Th. 2.1):

$$\rho = v_1\rho_1 \oplus \dots \oplus v_r\rho_r.$$

Hence its character is $\chi = v_1\chi_1 + \dots + v_r\chi_r$ and here v_i is given by

$$v_i = (\chi, \chi_i),$$

using (5). This shows that any representation of G (over \mathbf{C}) is determined up to equivalence by its character. For example, the regular representation $\mu(x):a \mapsto ax$ has this form: $\mu = \bigoplus d_i \rho_i$ and so we again find

$$|G| = \mu(1) = \sum d_i^2.$$

The inner product (4) can also be expressed directly in terms of the modules affording the representation:

PROPOSITION 4.3 *Let G be a finite group and U, V any G -modules over an algebraically closed field k , affording representations with characters α, β respectively. Then*

$$(\alpha, \beta) = \dim_k (\text{Hom}_G(U, V)). \quad (6)$$

Proof. Suppose first that U, V are simple. Then by Lemma 3.1 the right-hand side of (6) is 1 or 0 according as U, V are or are not isomorphic, and this is just the value of the left, by (5). Now the general case follows because every G -module is a direct sum of simple G -modules. ■

The number on the right of (6) is also called the *intertwining number* of U and V .

Above we have found the character of the regular representation, and it is not difficult to obtain an explicit expression for it. Sometimes we shall want an expression for the character of the representation afforded by a given right ideal of the group algebra. Since the latter is semisimple, each right ideal is generated by an idempotent, and the next result expresses the character in terms of this idempotent.

PROPOSITION 4.4 *Let G be a finite group, $A = kG$ its group algebra and $I = eA$ a right ideal in A , with idempotent generator $e = \sum e(x)x$. Then the character afforded by I is*

$$\chi(g) = \sum_x e(xg^{-1}x^{-1});$$

more generally, we have for any $b \in A$,

$$\chi(b) = \sum_{x,v} e(xv^{-1}x^{-1})b(v).$$

Proof. Consider the operation $\rho(b):a \mapsto eab$, representing the projection on I followed by the right regular representation. We have

$$a \cdot \rho(b) = eab = \sum e(uv^{-1}w^{-1})a(w)b(v).u. \quad (7)$$

Now $x \cdot \rho(b) = \sum_y \rho_{xy}(b)y$, by expressing ρ in terms of the natural basis; hence on

putting $a = x$, $u = y$ in (7), we find

$$\rho_{xy}(b) = \sum e(yv^{-1}x^{-1})b(v).$$

Therefore the character afforded by I is

$$\chi(b) = \text{tr}(\rho(b)) = \sum_x \rho_{xx}(b) = \sum_{x,v} e(xv^{-1}x^{-1})b(v). \quad \blacksquare$$

So far we have regarded the characters as functions on G ; but as we saw in Prop. 4.1, they are really class functions and we may regard them as functions on the set of all conjugacy classes of G . We shall now interpret the orthogonality relations in this way and in the process find that the number of irreducible characters of G is equal to the number of conjugacy classes.

Our first remark is that $a(x)(x \in G)$ is a class function iff the element $a = \sum a(x)x$ lies in the centre of the group algebra. For we have

$$y^{-1}ay = \sum_y a(x)y^{-1}xy = \sum_z a(yzy^{-1})z \quad \text{for all } y \in G;$$

hence

$$y^{-1}ay = a \quad \text{for all } y \in G \Leftrightarrow a(yzy^{-1}) = a(z) \quad \text{for all } y, z \in G.$$

Thus an element $a = \sum a(x)x$ lies in the centre of kG iff $a(x)$ is constant on conjugacy classes. This just means that we can write $a = \sum c_\lambda e_\lambda$ where c_λ is the sum of all elements in a given conjugacy class C_λ . It follows that these class sums c_λ form a basis for the centre of kG . This proves the first part of our next result:

THEOREM 4.5 *Let G be a finite group and k an algebraically closed field of characteristic 0 or prime to $|G|$. An element $a = \sum a(x)x$ of kG lies in the centre if and only if $a(x)$ is a class function. Moreover, the class sums c_λ form a basis of the centre of kG and the irreducible characters over k form a basis for the class functions; thus if χ_1, \dots, χ_r are the different irreducible characters, then any class function α on G may be written in the form*

$$\alpha = \sum_i (\chi_i, \alpha) \chi_i. \quad (8)$$

Hence the number of irreducible characters equals the number of conjugacy classes of G .

To complete the proof we denote the number of irreducible characters by r and the number of conjugacy classes by s ; as we have seen, s is the dimension of the centre of kG . Now kG is a direct product of r full matrix rings over k . Each matrix ring has a one-dimensional centre (cf. 4.3) and the centre of the direct product is easily seen to be the direct product of the centres. Hence the centre of kG is r -dimensional over k , and it follows that $r = s$. The characters are independent, by the orthogonality relation (5), hence they form a basis, which is orthonormal, by (5), and we therefore have (8). \blacksquare

Let us consider the multiplication table for the basis c_1, \dots, c_s of the centre of kG . Any product of classes $C_\lambda C_\mu$ is a union of a number of classes, hence we have

$$c_\lambda c_\mu = \sum_v \gamma_{\lambda\mu v} c_v, \quad (9)$$

where the $\gamma_{\lambda\mu v}$ are non-negative integers. If ρ is any irreducible representation of G , then $\rho(c_\lambda) = \eta_\lambda I$ by Schur's lemma, where $\eta_\lambda \in k$. Let χ denote the character of ρ , d its degree and write $h_\lambda = |C_\lambda|$. Taking characters in the last equation, we find

$$h_\lambda \chi^{(\lambda)} = \eta_\lambda d. \quad (10)$$

If we apply ρ to (9) we obtain $\eta_\lambda \eta_\mu = \sum \gamma_{\lambda\mu v} \eta_v$, and it follows that η_μ is a root of the equation

$$\det(xI - \Gamma_\mu) = 0, \quad \text{where } \Gamma_\mu = (\gamma_{\lambda\mu v}). \quad (11)$$

This shows η_μ to be an algebraic integer. Further, (10) shows that for each irreducible character χ , $h_\mu \chi^{(\mu)}/d$ is a root of (11); since (11) is of degree r , its roots are the values $h_\mu \chi_i^{(\mu)}/d_i$ for the different irreducible characters of G .

As a consequence of this development we can show that the degrees of the irreducible representations divide the group order, if we are willing to use a result from Ch. 8.

PROPOSITION 4.6 *For any finite group G the degree of each irreducible representation over \mathbf{C} divides $|G|$.*

Proof. Let χ be an irreducible character and d its degree. As a sum of roots of 1, χ is an algebraic integer, and so is $h_\lambda \chi^{(\lambda)}/d$, as we saw above. By the orthogonality relation (5) we have

$$\sum \frac{h_\lambda \chi^{(\lambda)}}{d} \overline{\chi^{(\lambda)}} = \frac{|G|}{d},$$

and since sums and products of algebraic integers are integral (cf. 8.4), it follows that $|G|/d$ is integral, as we had to show. ■

We note that the relations (5) can be written $|G|^{-1} \sum_i \overline{\chi_i^{(\lambda)}} \chi_j^{(\lambda)} h_\lambda = \delta_{ij}$, where $h_\lambda = |C_\lambda|$. This tells us that the $r \times r$ matrix $([h_\lambda/|G|]^{1/2} \chi_i^{(\lambda)})$ is unitary; hence so is its conjugate transpose and we have

$$|G|^{-1} \sum h_\lambda^{1/2} h_\mu^{1/2} \overline{\chi_i^{(\lambda)}} \chi_i^{(\mu)} = \delta_{\lambda\mu}.$$

When $\lambda \neq \mu$, we can omit h_λ , and so we obtain the second orthogonality relation for characters:

PROPOSITION 4.7 *For any finite group G , if $\chi_i^{(\lambda)}$ is the value of the i th irreducible character on the class C_λ and $|C_\lambda| = h_\lambda$, then*

$$\sum_i \chi_i^{(\lambda)} \chi_i^{(\mu)} = \begin{cases} |G|/h_\lambda & \text{if } \lambda = \mu, \\ 0 & \text{if } \lambda \neq \mu. \end{cases} \quad \blacksquare$$

The character of a representation may also be used to describe its kernel:

PROPOSITION 4.8 *Let G be a finite group and ρ a representation of G over \mathbf{C} with character χ . Then the kernel of ρ is given by*

$$K_\rho = \{x \in G \mid \chi(x) = \chi(1)\}.$$

Proof. Denote the degree of ρ by d , so that $\chi(1) = d$. If $x \in G$ has order n , then x is represented by a matrix $\rho(x)$, whose eigenvalues are n th roots of 1. Thus we have

$$\chi(x) = \omega_1 + \cdots + \omega_d, \quad \text{where } \omega_i^n = 1.$$

Since $\chi(x) = \chi(1) = d$, it follows that

$$d = |\omega_1 + \cdots + \omega_d| \leq |\omega_1| + \cdots + |\omega_d| = d.$$

Here equality must hold, which is possible only if $\omega_1 = \cdots = \omega_d$ and since $\sum \omega_i = d$, each ω_i must be 1. Since $\rho(x)$ satisfies an equation with distinct roots, it can be diagonalized (by Th. 5, 11.2, p. 343 of Vol. 1). Hence $\rho(x) = I$ and so x is in the kernel of ρ . The converse is clear. ■

Combining this result with Th. 3.4, we obtain

COROLLARY 4.9 *Let G be a finite group. Given $x \in G$, if $\chi(x) = \chi(1)$ for all irreducible characters χ of G , then $x = 1$. ■*

As a final result on characters we give a formula for characters of permutation modules. By a *permutation module* for a group G we understand a G -module V with a basis B on which G acts by permutations. The character afforded by V , $\chi(x)$, is just the number of points of B fixed by $x \in G$. Suppose that G is transitive on B and let H be the stabilizer of a point $p \in B$. Then each point of B is fixed by $|H|$ elements, for its stabilizer is conjugate to H . Recalling that $|B| = (G:H)$ (cf. p. 54 of Vol. 1), we therefore have

$$\sum_x \chi(x) = |B| \cdot |H| = |G|.$$

For general permutation modules we can apply the result to each orbit and obtain

PROPOSITION 4.10 *Let G be a finite permutation group acting on a set B and*

denote the character afforded by the corresponding permutation module by χ . Then

$$\sum_x \chi(x) = n \cdot |G|, \quad (12)$$

where n is the number of orbits. ■

An elaboration of this argument allows us to evaluate (χ, χ) .

THEOREM 4.11 *Let G be a transitive permutation group, denote by H the stabilizer of a point and let r be the number of double cosets in the decomposition $G = \bigcup Hs_iH$. If χ is the character afforded by the corresponding permutation module, then*

$$(\chi, \chi) = |G|^{-1} \sum \chi(x)^2 = r. \quad (13)$$

Proof. Let G act on B and consider the action of H on B . We may replace B by the coset decomposition with respect to H ; $G = \bigcup Hx_\lambda$. Here each orbit of H corresponds to a double coset Hs_iH . Now take $a \in G$, say $a = hs_ih'$, where $h, h' \in H$. If $\chi(a) = t$, then a leaves t cosets Hx_λ fixed, i.e. $a \in x_\lambda^{-1}Hx_\lambda$ for t values of λ . Now the action of $x_\lambda^{-1}Hx_\lambda$ on B yields $r \cdot |H|$ for the sum of its characters, by Prop. 4.10. There are $(G:H)$ such conjugates, so taking all these characters, we obtain $r \cdot |H| \cdot (G:H) = r \cdot |G|$. The character sum includes each value $\chi(a)$, and if $\chi(a) = t$, this term occurs t times, so in all we obtain $\sum \chi(a)^2$, i.e. (13). ■

Examples. We end this section with some examples of representations and characters; throughout, k is algebraically closed of characteristic 0.

1. Let A be a finite abelian group. Then kA is a commutative semisimple algebra, hence a direct product of copies of k , and all its irreducible representations are of degree 1. We take a basis a_1, \dots, a_m of A (in multiplicative notation), where a_i has order n_i , and denote by ε_i any primitive n_i th root of 1; then for any integers v_1, \dots, v_m , the mapping

$$a_1^{\alpha_1} \dots a_m^{\alpha_m} \mapsto \varepsilon_1^{\alpha_1 v_1} \dots \varepsilon_m^{\alpha_m v_m}$$

is a representation. We get distinct characters for n_i different values of v_i , and the $n_1 \dots n_m$ characters so obtained are all different and constitute all the irreducible characters of A (cf. Th. 4.6.1).

2. Consider D_m , the dihedral group of order $2m$, with generators a, b and defining relations $a^m = 1$, $b^2 = 1$, $b^{-1}ab = a^{-1}$. Every element can be uniquely expressed in the form $a^\alpha b^\beta$, where $0 \leq \alpha < m$, $0 \leq \beta < 2$. It is easily verified that the conjugacy classes are for odd m : $\{a^r, a^{-r}\}$, $r = 1, \dots, \frac{1}{2}(m-1)$, $\{1\}$, $\{a^\alpha b\}$; and for even m : $\{a^r, a^{-r}\}$, $r = 1, \dots, \frac{1}{2}m-1$, $\{1\}$, $\{a^{m/2}\}$, $\{a^{2\alpha}b\}$, $\{a^{2\alpha+1}b\}$. Further it may be checked that the index of the derived group in D_m is 2 when m is odd and 4 when m is even.

To find the representations of D_m , we have a homomorphism $D_m \rightarrow C_2$

obtained by mapping $a \mapsto 1$, which gives rise to two representations, the trivial representation and $a \mapsto 1, b \mapsto -1$. Further representations are obtained by taking a primitive m th root of 1, say ω , and writing

$$a \mapsto \begin{pmatrix} \omega^i & 0 \\ 0 & \omega^{-i} \end{pmatrix}, \quad b \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (i = 1, \dots, [m/2]). \quad (14)$$

When m is odd, we thus obtain $\frac{1}{2}(m-1)$ irreducible representations of degree 2 and two of degree 1, and this is a complete set, because the total number is $\frac{1}{2}(m-1) + 2$, which is the number of conjugacy classes. We also note the degree equation

$$\frac{1}{2}(m-1).2^2 + 2.1^2 = 2m.$$

When m is even, (14) becomes reducible for $i = m/2$ and we obtain two more representations of degree 1, $a \mapsto -1, b \mapsto \pm 1$. We now have $\frac{1}{2}m + 3$ classes and the degree equation becomes

$$(\frac{1}{2}m-1).2^2 + 4.1^2 = 2m.$$

3. Character tables for the symmetric groups S_3, S_4 . In the tables below the rows indicate the different characters, while the columns indicate the conjugacy classes, headed by a typical element and the order of each class. We recall (3.5 in Vol. 1) that the class of each permutation is determined by its cycle structure; hence the number of conjugacy classes of S_n is the number of partitions of n into positive integers. Moreover, the derived group is A_n and its index in S_n is 2; hence there are just two linear characters, the trivial character and the sign character.

	1	3	2	1	6	8	6	3
	1	(12)	(123)	1	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	χ_1	1	1	1	1
χ_2	1	-1	1	χ_2	1	-1	1	-1
χ_3	2	0	-1	χ_3	3	1	0	-1
				χ_4	3	-1	0	1
				χ_5	2	0	-1	0
								2

In each table the first row is the trivial character and the second row the sign character. The degrees in the first column are obtained by solving the degree equation $\sum d_i^2 = n!$. Each character χ gives rise to a ‘conjugate’ character $\chi\chi_2$, corresponding to the tensor product $\rho \otimes \rho_2$ of the representations. Thus χ_3 and χ_4 in the second table are conjugate and χ_3 in the first table and χ_5 in the second table are self-conjugate, hence they vanish on odd classes. Now the remaining values are found by orthogonality, using Prop. 4.7.

We note that the characters for S_3, S_4 are rational. This is a general feature of symmetric groups; in fact, as we shall see in 7.6, the irreducible representations of S_n can be expressed over \mathbb{Q} .

Exercises

- (1) Verify the calculations in the above examples, and make a character table for the quaternion group of order 8.
- (2) Show that any two elements x, y of a group G are conjugate iff $\chi(x) = \chi(y)$ for all irreducible characters χ .
- (3) Show that any character which is zero on all elements $\neq 1$ of G is a multiple of the regular representation.
- (4) Show that if $S = \{x_1, \dots, x_n\}$ is a transitive G -set, then the vector space spanned by the $x_i - x_j$ is a G -module affording a representation which does not contain the trivial representation.
- (5) Let χ be a character of a finite group G over a field of characteristic 0. Show that (χ, χ) is a positive integer.
- (6) Let ρ be an irreducible representation of degree d of G , with character χ . Show that the simple factor of the group algebra corresponding to ρ has unit element e given by $|G| \cdot e = d \cdot \sum \chi(x^{-1})x$.
- (7) If g_{ijk} is defined as in (11) in 7.3, show that $g_{ijk} = (\chi_i \chi_j, \chi_k)$. Use Prop. 4.7 to evaluate the sum of all the g_{ijk}^2 and deduce the formula $\sum_{ijk} g_{ijk}^2 = |G| \sum_\lambda h_\lambda^{-1}$. Show that the value of this sum for S_3 is 11 and for S_4 is 43, and verify the formula in these cases.
- (8) Let G be a finite group with irreducible characters χ_1, \dots, χ_r , and U a G -module with basis u_1, \dots, u_m . Show that the character afforded by the G -module $U \otimes U \otimes \dots \otimes U$ (n factors) is $\sum n_i \chi_i$, where n_i is the coefficient of t^n in $|G|^{-1} \sum_x \bar{\chi}_i(x) \det(I - t\rho(x))^{-1}$, ρ being the representation afforded by U . Deduce that the number of invariants of degree n in the u 's is the coefficient of t^n in $|G|^{-1} \sum \det(I - t\rho(x))^{-1}$ (Molien 1897).

7.5 Complex representations

When we restrict our representations to be over the complex numbers, several simplifications can be made. In the first place we can then confine attention to unitary representations; secondly, by examining the different types of irreducible representations we obtain an estimate for the sum of their degrees in terms of the number of elements of order 2 in the group.

Throughout this section we only consider complex representations, thus we take \mathbf{C} as our ground field. We recall that a square matrix P over \mathbf{C} is said to be *unitary* if $PP^\text{H} = I$, where $P^\text{H} = \bar{P}^\text{T}$ is the transpose of the complex conjugate of P . The $d \times d$ unitary matrices over \mathbf{C} form a group, the *unitary group*, written $\mathbf{U}_d(\mathbf{C})$. By a *unitary representation* of G we understand a homomorphism $G \rightarrow \mathbf{U}_d(\mathbf{C})$. In the special case where the representation is real, we have an *orthogonal representation* because $\mathbf{U}_d(\mathbf{C}) \cap \mathbf{GL}_d(\mathbf{R}) = \mathbf{O}_d(\mathbf{R})$, the orthogonal group.

Any G -module U affording a unitary representation of G has a hermitian metric defined on it which is invariant under G . Thus there is a positive definite hermitian form (u, v) on U :

$$(\lambda u + \lambda' u', v) = \lambda(u, v) + \lambda'(u', v), \quad (v, u) = \overline{(u, v)}, \quad (u, u) > 0 \text{ for } u \neq 0, \quad (1)$$

which in addition satisfies

$$(ux, vx) = (u, v) \quad \text{for all } u, v \in U, x \in G. \quad (2)$$

For, given a unitary representation ρ , relative to the basis v_1, \dots, v_d of U , let us define a hermitian form by writing $(v_i, v_j) = \delta_{ij}$. If the vectors u, v have coordinate rows α, β , then $(u, v) = \alpha\beta^H$ and $(ux, vx) = \alpha\rho(x)(\beta\rho(x))^H = \alpha\rho(x)\rho(x)^H\beta^H = \alpha\beta^H$, because $\rho(x)\rho(x)^H = I$ by unitarity. Conversely, if we have a positive definite hermitian form satisfying (2), then transformation by x preserves the metric and so must be unitary.

Any unitary representation is completely reducible. To verify this assertion we take the corresponding module U . If W is any submodule, then its orthogonal complement $W^\perp = \{u \in U | (u, w) = 0 \text{ for all } w \in W\}$ is again a G -submodule and is complementary to W . In terms of representations we can also verify this fact by noting that a reduced matrix $\rho(x)$ must be fully reduced, because its transpose is $\overline{\rho(x^{-1})}$.

The importance of unitary representations is underlined by the following result, which incidentally provides another proof of Maschke's theorem.

PROPOSITION 5.1 *Every complex representation of a finite group G is equivalent to a unitary representation; every real representation is equivalent to an orthogonal representation.*

Proof. Let ρ be the representation of G and U a G -module affording ρ . We have to find a positive definite hermitian form on U which is invariant under U . Take any positive definite hermitian form $h(u, v)$ on U and define

$$(u, v) = \sum_{x \in G} h(ux, vx).$$

For any $a \in G$ we have

$$(ua, va) = \sum_x h(uax, vax) = \sum_y h(uy, vy) = (u, v).$$

Thus (u, v) is invariant under G , and it is clearly positive definite hermitian, as a sum of such forms. On choosing an orthonormal basis, we obtain the desired unitary representation; starting from a real representation, we obtain an orthogonal representation of G . ■

In dealing with unitary representations, we usually restrict equivalence to be by unitary matrices. Thus two unitary representations ρ, σ are *unitarily equivalent* if

$\sigma(x) = P\rho(x)P^{-1}$ for a unitary matrix P . For irreducible representations this is automatic, for if ρ, σ are equivalent, we have $\sigma(x)S = S\rho(x)$ for an invertible matrix S . Taking hermitian conjugates, we have $S^H\sigma(x)^H = \rho(x)^HS^H$, hence

$$\rho(x)S^HS = \rho(x)S^H\sigma(x)^H\sigma(x)S = \rho(x)\rho(x)^HS^HS\rho(x) = S^HS\rho(x).$$

Since ρ is irreducible, we have $S^HS = \lambda I$ by Schur's lemma, and here $\lambda > 0$, because S^HS is positive definite. Writing $\lambda = \mu^2$ ($\mu > 0$) and $T = \mu^{-1}S$, we obtain a unitary matrix T such that $\sigma(x) = T\rho(x)T^{-1}$.

The irreducible complex representations may be classified as follows. Let ρ be an irreducible complex representation (not necessarily unitary). If ρ is equivalent to a real representation, it is said to be of the *first kind*. If ρ is not of the first kind, but is equivalent to its conjugate $\bar{\rho}$, it is said to be of the *second kind*; in the remaining case ρ is of the *third kind*. Our next result shows how to distinguish the first two of these cases. In the proof we shall need the elementary fact that a symmetric unitary matrix can always be written as the square of a symmetric unitary matrix. We recall the proof.

Let P be symmetric and unitary; as a unitary matrix P is similar to a diagonal matrix, say $S^{-1}PS = D$ for a unitary matrix S , and

$$SDS^{-1} = P = P^T = (S^T)^{-1}DST,$$

i.e. $S^TSD = DST$. Now D is diagonal and again unitary, so its diagonal elements have absolute value 1 and we can find a unitary diagonal matrix E such that $E^2 = D$ and

$$S^TSE = ESTS. \quad (3)$$

Put $Q = SES^{-1}$; then Q is unitary, because E is, and $Q^2 = SE^2S^{-1} = P$. Moreover, $Q^T = (S^T)^{-1}EST = SES^{-1}$ by (3), hence $Q^T = Q$, so Q is also symmetric.

PROPOSITION 5.2 *Let ρ be a complex irreducible representation of G such that*

$$\overline{\rho(x)} = P^{-1}\rho(x)P \quad \text{for some } P \in \mathbf{U}_d(\mathbb{C}). \quad (4)$$

Then either $P^T = P$ and ρ is of the first kind, or $P^T = -P$ and ρ is of the second kind.

Proof. Taking complex conjugates in (4) we have $\rho(x) = P^T\overline{\rho(x)}(P^T)^{-1}$, because $P^{-1} = P^H = \overline{P^T}$; therefore

$$P^T P^{-1} \rho(x) = P^T \overline{\rho(x)} P^{-1} = \rho(x) P^T P^{-1},$$

hence $P^T P^{-1} = \lambda I$, so $P^T = \lambda P$. Transposing, we find $P = \lambda P^T = \lambda^2 P$, hence $\lambda^2 = 1$ and so $\lambda = \pm 1$. This shows that either $P^T = P$ or $P^T = -P$.

Now it is clear from (4) that P cannot be of the third kind, so it is enough to show that ρ is of the first kind iff $P^T = P$. If ρ is of the first kind, then there is a

matrix L such that $L^{-1}\rho(x)L$ is real for all $x \in G$, hence by (4),

$$L^{-1}\rho(x)L = \bar{L}^{-1}\overline{\rho(x)}\bar{L} = \bar{L}^{-1}P^{-1}\rho(x)P\bar{L}.$$

It follows that $P\bar{L}L^{-1}$ commutes with $\rho(x)$ and so $P\bar{L}L^{-1} = \alpha I$, i.e. $P = \alpha L\bar{L}^{-1}$. Now if $P^T = -P$, then $P^{-1} = P^H = -\bar{P}$ and so $I = PP^{-1} = -\alpha L\bar{L}^{-1}\cdot\alpha\bar{L}L^{-1} = -\alpha\bar{\alpha}I$, which is a contradiction. Therefore $P^T = P$ in this case.

Conversely, assume that $P^T = P$. Then P is symmetric unitary and by the above remark, $P = Q^2$, where Q is symmetric unitary. Hence $\bar{Q} = Q^H = Q^{-1}$, and $\overline{Q^{-1}\rho(x)Q} = QP^{-1}\rho(x)PQ^{-1} = Q^{-1}\rho(x)Q$. Therefore ρ is equivalent to a real representation, and so is of the first kind. ■

If in Prop. 5.2, ρ is of the second kind and its degree is denoted by d , we have $P^T = -P$, hence $\det P = (-1)^d \det P$, and it follows that $(-1)^d = 1$. Hence d must be even and we have

COROLLARY 5.3 *Any complex irreducible representation of the second kind has even degree.* ■

For any character χ of G let us define its *indicator* as

$$v(\chi) = |G|^{-1} \sum_{x \in G} \chi(x^2). \quad (5)$$

The three kinds of irreducible representation may be described in terms of the indicator:

THEOREM 5.4 *Let χ be a complex irreducible character of a finite group G and $v(\chi)$ its indicator as in (5). Then χ is of the first kind if $v(\chi) = 1$, of the second kind if $v(\chi) = -1$ and of the third kind if $v(\chi) = 0$.*

Proof. Let ρ be a representation affording χ and denote its degree by d . We may take ρ to be unitary, and then have

$$\begin{aligned} |G| \cdot v(\chi) &= \sum_{i,x} \rho_{ii}(x^2) = \sum_{ij} \sum_x \rho_{ij}(x)\rho_{ji}(x) \\ &= \sum_{ij} \sum_x \rho_{ij}(x)\overline{\rho_{ij}(x^{-1})}. \end{aligned}$$

If ρ is of the third kind, ρ and $\bar{\rho}$ are inequivalent and then $v(\chi) = 0$ by the orthogonality relations (Th. 3.2). If ρ is of the first kind, we may assume that $\bar{\rho} = \rho$; in that case

$$|G| \cdot v(\chi) = \sum \sum \rho_{ij}(x)\rho_{ij}(x^{-1}) = \sum \delta_{ji}\delta_{ij}|G|/d = |G|,$$

and hence $v(\chi) = 1$. Finally, if ρ is of the second kind, then by Prop. 5.2 there is a unitary matrix P such that $P^T = -P$ and $P^{-1}\rho(x)P = \overline{\rho(x)}$; hence on writing

$P = (p_{ij})$, $P^{-1} = (q_{ij})$ we have, again by the orthogonality relations,

$$\begin{aligned} v(\chi) &= |G|^{-1} \sum \sum \rho_{ij}(x) q_{ir} \rho_{rs}(x^{-1}) p_{sj} = d^{-1} \sum q_{ir} p_{sj} \delta_{jr} \delta_{is} \\ &= d^{-1} \sum q_{ir} p_{ir} = d^{-1} \operatorname{tr}(P^{-1} P^T) = \frac{\operatorname{tr}(-I)}{d} = -1. \end{aligned}$$

Thus $v(\chi) = -1$, as we wished to show. ■

Finally we show how the indicator is related to the solutions of the equation $x^2 = a$ in G .

PROPOSITION 5.5 *Let G be a finite group. Given $a \in G$, let $t(a)$ be the number of solutions of the equation $x^2 = a$ in G . If χ_1, \dots, χ_r are the different complex irreducible characters of G , then*

$$t(a) = \sum_{i=1}^r v(\chi_i) \chi_i(a). \quad (6)$$

Proof. It is clear that $t(a)$ is a class function on G , so it can be written in the form $t(a) = \sum c_i \chi_i(a)$, and it only remains to show that $c_i = v(\chi_i)$. The sets $T(a) = \{x \in G \mid x^2 = a\}$ form a partition of G , and we have by Th. 4.5, using (8) to determine the coefficient c_i :

$$\begin{aligned} |G| \cdot c_i &= \sum t(a) \overline{\chi_i(a)} = \sum_{a \in G} \sum_{x \in T(a)} \overline{\chi_i(x^2)} \\ &= \sum_{x \in G} \overline{\chi_i(x^2)} = |G| \cdot \overline{v(\chi_i)}, \end{aligned}$$

by (5). Since the indicator is always real, it follows that $c_i = v(\chi_i)$, as we had to show. ■

Let us apply this result for $a = 1$. In this case $\chi_i(1) = d_i$ is the degree of χ_i . Bearing in mind that $v(\chi) \leq 1$, with equality precisely when χ is of the first kind, by Th. 5.4, we obtain

COROLLARY 5.6 *If t is the number of elements of order 2 in G , and the degrees of its complex irreducible representations are d_1, \dots, d_r , then*

$$t + 1 = \sum v(\chi_i) d_i \leq \sum d_i, \quad (7)$$

with equality if and only if all the irreducible characters are of the first kind. ■

As an example consider D_m , the dihedral group of order $2m$, where m is odd. We saw that there are $(m-1)/2$ characters of degree 2 and two linear characters, and there are m elements of order 2, viz. $a^b b$ in the notation of 7.4. The inequality (7) in this case reads

$$m + 1 \leq \frac{1}{2}(m-1) \cdot 2 + 2 \cdot 1.$$

Since equality holds here, all the representations must be of the first kind, and going through the proof of Prop. 5.2, we find that the representation given at the end of 7.4 is equivalent to the real representation

$$a \mapsto \frac{1}{2} \begin{pmatrix} \lambda + \lambda^{-1} & \lambda - \lambda^{-1} \\ \lambda^{-1} - \lambda & \lambda + \lambda^{-1} \end{pmatrix}, \quad b \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \lambda = \omega^j,$$

$j = 1, \dots, (m-1)/2.$

As a second example consider the degrees of the irreducible representations of S_5 . We shall see in 7.6 that all irreducible representations of the symmetric group can be taken to be rational (even integral). The elements of order 2 in S_5 are transpositions or products of two transpositions. There are $5 \cdot 4 / 2 = 10$ transpositions $(i \ j)$ and $10 \cdot 3 / 2 = 15$ products of the form $(i \ j)(k \ l)$; hence $t = 25$, $t+1 = 26$. To find the number of conjugacy classes we count the partitions of 5: (5) , $(4, 1)$, $(3, 2)$, $(3, 1^2)$, $(2, 1^3)$, $(2^2, 1)$, (1^5) . So we have to look for seven positive integers, divisors of $5! = 120$, whose sum is 26 and whose squares have the sum 120, by (10) of 7.3. Prop. 4.2 tells us that just two of the numbers are 1, because $(S_5 : A_5) = 2$, and this leaves d_1, \dots, d_5 such that

$$\sum_1^5 d_i = 24, \quad \sum_1^5 d_i^2 = 118.$$

Further, we can rule out the low dimensions 2, 3 because an integral representation can be reduced mod 2, and we cannot have a homomorphism of S_5 into $\mathbf{GL}_2(\mathbf{F}_2)$ (order $(2^2 - 1)(2^2 - 2) = 6$) or into $\mathbf{GL}_3(\mathbf{F}_2)$ (order $(2^3 - 1)(2^3 - 2)(2^3 - 2^2) = 192$) which maps A_5 non-trivially. This leaves as the only possibility for the degrees; 1, 1, 4, 4, 5, 5, 6.

Exercises

- (1) Use the information on S_5 to make a character table for it.
- (2) Find the matrix of transformation which reduces the representation (14) of 7.4 for \mathbf{D}_m to the form given above.
- (3) Let V be a simple G -module with real character. Show that there is just one invariant bilinear form b on V , up to scalar multiples. Show further that b is either symmetric or antisymmetric, and that the latter can happen only when $\dim V$ is even.
- (4) Show that in a group G of odd order no element $\neq 1$ is conjugate to its inverse. Deduce that for any $y \neq 1$, $\sum_i \chi_i(y)^2 = \sum_i \chi_i(y) \overline{\chi_i(y^{-1})} = 0$. Hence show that G has an irreducible character of the third kind.
- (5) Show that every non-trivial irreducible character of a group of odd order is of the third kind. (*Hint.* Use the methods of Ex. (4) and the fact that $\chi(1)$ is an odd integer.)

7.6 Representations of the symmetric group

In principle all the irreducible representations of a finite group can be obtained by taking the regular representation and reducing it completely. This is a lengthy undertaking, but for certain types of groups there are more direct methods. We shall describe such a method for symmetric groups; it is due mainly to Frobenius and A. Young, with some simplifications by J. von Neumann.

We recall that the symmetric group of degree n , S_n , is the group of all permutations of $1, 2, \dots, n$. It has order $n!$ and each permutation can be written as a product of disjoint cycles, e.g.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 5 & 6 & 8 & 1 & 9 & 2 & 4 & 3 \end{pmatrix} = (2 \ 5 \ 1 \ 7)(3 \ 6 \ 9)(4 \ 8).$$

The cycles commute, so we can arrange them by decreasing length. If the lengths are $\alpha_1, \dots, \alpha_h$, we may thus suppose that

$$\alpha_1 + \alpha_2 + \dots + \alpha_h = n, \quad (1)$$

and

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_h. \quad (2)$$

If λ_i of the α_i are 1, λ_2 are 2, etc., we can also write $1^{\lambda_1} 2^{\lambda_2} \dots r^{\lambda_r}$ (where r is the largest α_i) for the set of α 's. This is called the *cycle structure* of the permutation. Two permutations have the same cycle structure iff they are conjugate: If

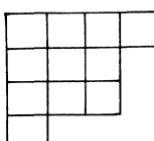
$$\begin{aligned} f &= (a_1 \cdots a_{\alpha_1})(a_{\alpha_1+1} \cdots a_{\alpha_1+\alpha_2}) \cdots (\cdots a_n), \\ g &= (b_1 \cdots b_{\alpha_1})(b_{\alpha_1+1} \cdots b_{\alpha_1+\alpha_2}) \cdots (\cdots b_n), \end{aligned}$$

then $g = p^{-1}fp$, where $p: a_i \mapsto b_i$. Hence two permutations with the same cycle structure are conjugate; the converse is clear.

It follows that the number of conjugacy classes of S_n equals the number of sequences $(\alpha_1, \dots, \alpha_h)$ of positive integers satisfying (1) and (2); by Th. 4.5 this is also the number of inequivalent irreducible representations of S_n . To get a complete system of irreducible representations of S_n we need only construct for each sequence $(\alpha_1, \dots, \alpha_h)$ an irreducible representation, such that representations corresponding to different sequences are inequivalent. This will be our aim in what follows.

We shall write $\alpha \vdash n$ to indicate that $\alpha = (\alpha_1, \dots, \alpha_h)$ is a descending partition on n , as in (1), (2). To each such partition there corresponds a diagram of n squares arranged in h rows of α_i squares in the i th row. For example

$(4, 3^2, 1)$



This is called a *Young diagram* and is again denoted by α . Since $\alpha \vdash n$, we can write the numbers 1 to n in these squares (in some order). The result is called a *Young tableau*. For example

$(3, 2, 1)$

2	6	5
1	3	
4		

We see that for each diagram there are $n!$ distinct tableaux. If T_α is a Young tableau and $g \in S_n$, then $T_\alpha g$ denotes the tableau obtained from T_α by applying g .

We can let each Young tableau represent a permutation by regarding the rows as cycles. In this way the tableau just illustrated represents the permutation $(2 \ 6 \ 5)(1 \ 3)$. If T_α represents g_0 in this way, then $T_\alpha g$ represents $g^{-1}g_0g$, as is easily verified.

We now fix a tableau T_α and define two subgroups of S_n as follows: P_α is the set of all permutations leaving each symbol in its row, briefly the set of *row permutations* of T_α ; and Q_α is the set of all permutations leaving each symbol in its column, the set of *column permutations* of T_α . Here P_α and Q_α depend of course on the tableau T_α and not merely on the diagram α . For example, for the above tableau P_α is generated by $(2 \ 6)$, $(2 \ 6 \ 5)$, $(1 \ 3)$, while Q_α is generated by $(2 \ 1)$, $(2 \ 1 \ 4)$, $(3 \ 6)$. If we apply g to T_α and use the remark made earlier, we obtain

LEMMA 6.1 *Let T_α be a Young tableau with groups P_α, Q_α and let $g \in S_n$. Then the groups for $T_\alpha g$ are $g^{-1}P_\alpha g$ and $g^{-1}Q_\alpha g$. ■*

Let A be the group algebra of S_n (over the rational numbers, say) and write $\varepsilon(g)$ or ε_g for the sign of the permutation g . Writing p, q for the typical elements of P_α, Q_α respectively, we define two elements of A , the sum of the row permutations and the alternating sum of the column permutations:

$$f_\alpha = \sum_p p, \quad g_\alpha = \sum_q \varepsilon_q q. \quad (3)$$

LEMMA 6.2 *Let T_α be a Young tableau and f_α, g_α as in (3), and write $|P_\alpha| = r_\alpha$, $|Q_\alpha| = s_\alpha$. Then*

$$pf_\alpha = f_\alpha p = f_\alpha \quad \text{for } p \in P_\alpha, \quad (4)$$

$$qg_\alpha = g_\alpha q = \varepsilon_q g_\alpha \quad \text{for } q \in Q_\alpha, \quad (5)$$

$$f_\alpha^2 = r_\alpha f_\alpha, \quad g_\alpha^2 = s_\alpha g_\alpha. \quad (6)$$

Proof. We have $pf_\alpha = \sum_p pp' = \sum p' = f_\alpha$, $\varepsilon_q qg_\alpha = \sum \varepsilon_q \varepsilon_q q q' = g_\alpha$, which establishes half of (4), (5); the other half follows similarly. Now $f_\alpha^2 = \sum_p pf_\alpha = \sum f_\alpha = r_\alpha f_\alpha$, $g_\alpha^2 = \sum \varepsilon_q qg_\alpha = \sum g_\alpha = s_\alpha g_\alpha$. ■

Next comes a basic combinatorial lemma on which everything else depends. We order the partitions lexicographically, by writing $\alpha > \beta$ if the first non-zero difference $\alpha_i - \beta_i$ is positive (where α or β is completed by 0's if necessary). This provides a total ordering of the set of all partitions.

LEMMA 6.3 *Let $\alpha, \beta \vdash n$ and take any tableaux T_α, T'_β . If $\alpha \geq \beta$ and no two numbers in the same row of T_α occur in the same column of T'_β , then (i) $\alpha = \beta$ and (ii) $T'_\beta = T_\alpha qp$ for some $p \in P_\alpha, q \in Q_\alpha$.*

Some care is needed here, for we have abused notation by writing P_α instead of the more accurate $P(T_\alpha)$. In fact we shall take P_α, Q_α to be the groups associated with T_α and write P'_β, Q'_β for the groups of T'_β .

Proof. Since $\alpha \geq \beta$, we have $\alpha_1 \geq \beta_1$. The first row of T_α has α_1 numbers, which must be in different columns of T'_β , so $\beta_1 \geq \alpha_1$, i.e. $\beta_1 = \alpha_1$. Now for a certain column permutation $q'_1 \in Q'_\beta$ we can bring these numbers into the top row, though possibly in a different order from that in T_α .

Leaving out the top row in $T'_\beta q'_1, T_\alpha$ we can repeat the argument, showing that $\beta_2 = \alpha_2$ and finding q'_2 such that $T'_\beta q'_1 q'_2$ has the same numbers as T_α in the second row as well as in the top row. After h steps (if α has h parts) we get $q' = q'_1 q'_2 \cdots q'_h$ such that $T'_\beta q'$ differs from T_α only by a row permutation: $T'_\beta q' = T_\alpha p$, where $p \in P_\alpha, q' \in Q'_\beta$. Now $T_\alpha = T'_\beta q' p^{-1}$, hence

$$Q_\alpha = (q' p^{-1})^{-1} Q'_\beta q' p^{-1} = p Q'_\beta p^{-1};$$

it follows that $q = p q'^{-1} p^{-1} \in Q_\alpha$. Therefore $qp = p q'^{-1}$ and we have $T'_\beta = T_\alpha qp$, as claimed, with $p \in P_\alpha, q \in Q_\alpha$. ■

COROLLARY 6.4 *Given $\alpha, \beta \vdash n$ and any tableaux T_α, T'_β , if $\alpha > \beta$, then*

$$f_\alpha x^{-1} g_\beta x = 0 \quad \text{for all } x \in S_n, \tag{7}$$

$$f_\alpha A g_\beta = 0. \tag{8}$$

Proof. We begin by showing that

$$f_\alpha g_\beta = 0 \quad \text{for } \alpha > \beta. \tag{9}$$

By Lemma 6.3 there must be two numbers i, k which lie in the same row of T_α and in the same column of T'_β . Write $t = (i \ k)$; then $f_\alpha t = f_\alpha, t g_\beta = -g_\beta$, hence $f_\alpha g_\beta = f_\alpha t^2 g_\beta = -f_\alpha g_\beta$ and (9) follows. Now $x^{-1} g_\beta x$ corresponds to the tableau $T_\beta x$ and (7) follows by applying (9) with $g'_\beta = x^{-1} g_\beta x$. Replacing x^{-1} by y and multiplying by y on the right, we have $f_\alpha y g_\beta = 0$, and now (8) follows by summing. ■

Given a Young tableau T_α , let us put

$$h_\alpha = f_\alpha g_\alpha = \sum_{p,q} \varepsilon_q p q. \quad (10)$$

Clearly $h_\alpha \neq 0$, because the coefficient of 1 is 1:

$$h_\alpha(1) = 1. \quad (11)$$

We may consider h_α as an operator, symmetrizing the rows and antisymmetrizing the columns; it is called the *Young symmetrizer* associated with the tableau T_α .

PROPOSITION 6.5 *Let h_α be the Young symmetrizer associated with a tableau T_α . Then the relation*

$$p a \varepsilon_q q = a \quad \text{for all } p \in P_\alpha, q \in Q_\alpha \quad (12)$$

holds for $a = h_\alpha$, and any $a \in A$ satisfying (12) must be of the form $a = \lambda h_\alpha$, where λ is a scalar. Moreover,

$$h_\alpha b h_\beta = 0 \quad \text{for } \alpha > \beta \text{ and any } b \in A, \quad (13)$$

$$h_\alpha b h_\alpha = \mu h_\alpha \quad \text{for any } b \in A. \quad (14)$$

Proof. By (4), $p f_\alpha g_\alpha = f_\alpha g_\alpha$, while (5) yields $f_\alpha g_\alpha q = \varepsilon_q f_\alpha g_\alpha$, hence (12) holds for $a = h_\alpha$. Now let $a = \sum a(x)x$ satisfy (12). Then

$$\sum_x \varepsilon_q a(x) p x q = \sum a(x) x \quad \text{for all } p \in P_\alpha, q \in Q_\alpha. \quad (15)$$

Comparing coefficients for $x = pq$, we find

$$\varepsilon_q a(1) = a(pq); \quad (16)$$

we claim that $a(x) = 0$ when x is not of the form pq . Consider T_α and $T'_\alpha = T_\alpha x^{-1}$; by Lemma 6.3 there are two numbers j, k in the same row in T_α and in the same column in T'_α (because T'_α is not of the form $T_\alpha qp$). Put $t = (j - k)$; then we have $t \in P_\alpha, t \in Q'_\alpha$, where Q'_α corresponds to T'_α ; therefore $t \in x Q_\alpha x^{-1}$ or also $x^{-1} t x \in Q_\alpha$. In (15) let us take $p = t, q = x^{-1} t x$; comparing the coefficient of x we find

$$(-1)a(x) = a(x),$$

hence $a(x) = 0$. Together with (16) this shows that $a = a(1) \cdot \sum p q \varepsilon_q = a(1) \cdot h_\alpha$ as claimed.

Now when $\alpha > \beta$, then by Cor. 6.4, $h_\alpha b h_\beta = f_\alpha g_\alpha b f_\beta g_\beta \in f_\alpha A g_\beta = 0$. This proves (13); and (14) follows because $h_\alpha b h_\alpha$ satisfies (12). ■

We now obtain the irreducible representations of S_n by expressing the group algebra as a direct sum of minimal right ideals.

THEOREM 6.6 *With each Young diagram α let us associate one definite Young*

tableau T_α and construct the corresponding Young symmetrizer h_α as element of the group algebra $A = \mathbf{QS}_n$:

$$h_\alpha = \sum \varepsilon_q pq, \quad (p \in P_\alpha, q \in Q_\alpha).$$

Then the $I^\alpha = h_\alpha A$ are simple submodules for the right regular representation of S_n ; they are pairwise non-isomorphic and afford a complete system of irreducible representations for S_n .

Proof. Let us show that $I^\alpha = h_\alpha A$ is a minimal right ideal of A . Given a right ideal m satisfying $m \subseteq I^\alpha$, we have $mh_\alpha \subseteq I^\alpha h_\alpha \subseteq \mathbf{Q}h_\alpha$. We distinguish two cases. (i) $mh_\alpha = \mathbf{Q}h_\alpha$. Then $I^\alpha = h_\alpha A = mh_\alpha A \subseteq m$, hence $I^\alpha = m$. (ii) $mh_\alpha = 0$. Then $m^2 = mI^\alpha = mh_\alpha A = 0$, so $m^2 = 0$ and therefore $m = 0$. It follows that I^α is minimal and the representation induced by the regular representation is irreducible.

We next show that I^α, I^β are not isomorphic for $\alpha \neq \beta$. Suppose that $\alpha > \beta$ say. By Cor. 6.4, $h_\alpha A h_\beta = 0$, hence $I^\alpha h_\beta = 0$, but $I^\alpha h_\alpha \neq 0$, because $h_\alpha \neq 0$. Now the number of distinct diagrams is the number of partitions of n , which is just the number of conjugacy classes of S_n , as we have seen; hence by Th. 4.5, we have a complete set of irreducible representations. ■

It still remains to calculate the irreducible characters. By Prop. 6.5 we have

$$h_\alpha^2 = \mu_\alpha h_\alpha, \tag{17}$$

therefore $e_\alpha = \mu_\alpha^{-1} h_\alpha$ is an idempotent and the right ideal of A generated by e_α is a minimal right ideal. To find μ_α consider the operator of left multiplication by h_α on I^α . By (17) this is

$$\lambda(h_\alpha) = \mu_\alpha \cdot 1. \tag{18}$$

On the other hand, for $a \in I^\alpha$, $a\lambda(h_\alpha) = h_\alpha a = \sum h_\alpha(uv^{-1})a(v) \cdot u$, hence the matrix of $\lambda(h_\alpha)$ (in the natural basis) is given by

$$\lambda_{xy}(h_\alpha) = h_\alpha(yx^{-1}),$$

and so $\text{tr}(\lambda(h_\alpha)) = \sum h_\alpha(xx^{-1}) = \sum h_\alpha(1) = \sum 1 = n!$, by (11). A comparison with (18) shows that $n! = \text{tr}(\lambda(h_\alpha)) = d_\alpha \mu_\alpha$, where d_α is the degree of the representation. Thus we obtain

$$\mu_\alpha = n!/d_\alpha.$$

Recalling Prop. 4.4, we find that the corresponding character is given by

$$\chi_\alpha(x) = \frac{d_\alpha}{n!} \sum h_\alpha(yx^{-1}y^{-1}). \tag{19}$$

From this formula it is possible to calculate χ_α explicitly in terms of the partition α of n ; we shall give the result here without proof and refer to Weyl (1939, Ch. VII) for details.

Let $x \in S_n$ have the cycle structure $\beta = (1^{\beta_1} 2^{\beta_2} \dots n^{\beta_n})$; since the value of a character χ at x depends only on β , we shall write it as $\chi(\beta)$. Let x_1, \dots, x_h be any variables and denote the Vandermonde determinant formed from them by writing down the terms of the main diagonal:

$$|x_1^{h-1} \quad x_2^{h-2} \quad \cdots \quad x_h^0|.$$

This is an alternating function of the x 's, briefly denoted by Δ in what follows. More generally, for any positive integers $\alpha_1, \dots, \alpha_h$ the function

$$|x_1^{\alpha_1 + h - 1} \quad x_2^{\alpha_2 + h - 2} \quad \cdots \quad x_h^{\alpha_h}|$$

formed in the same way, is an alternating function of the x 's, zero unless all the α 's are in descending order, and it may be written as $S(\alpha_1, \dots, \alpha_h)\Delta$, where S as quotient of two alternating functions is symmetric in the x 's. Such a function S is called a *bialternant* or *S-function*.

Given $\alpha = (\alpha_1, \dots, \alpha_h) \vdash n$, let us write $r_1 = \alpha_1 + (h - 1), r_2 = \alpha_2 + (h - 2), \dots, r_h = \alpha_h$. We define the power sums in the x 's as $s_i = \sum x_v^i$ and for a given $\beta = (\beta_1, \dots, \beta_t)$ satisfying $\sum i\beta_i = n$ put $\sigma(\beta) = s_1^{\beta_1} s_2^{\beta_2} \cdots s_t^{\beta_t}$. With these notations we have the following relation for the characters of S_n corresponding to partitions into at most h parts:

$$\sigma(\beta). |x_1^{h-1} \quad x_2^{h-2} \quad \cdots \quad x_h^0| = \sum_{\alpha} \chi_{\alpha}(\beta) |x_1^{r_1} \quad x_2^{r_2} \quad \cdots \quad x_h^{r_h}|. \quad (20)$$

Thus $\chi_{\alpha}(\beta)$ is the coefficient of $x_1^{r_1} x_2^{r_2} \cdots x_h^{r_h}$ in the expansion of the left-hand side of (20). On dividing by Δ we may write (20) as

$$\sigma(\beta) = \sum_{\alpha} \chi_{\alpha}(\beta) S(\alpha_1, \dots, \alpha_h); \quad (21)$$

this shows that $\chi_{\alpha}(\beta)$ is the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_h^{\alpha_h}$ in the expansion of the left-hand side of (21). Further, the degree of the irreducible character χ_{α} is given by

$$d_{\alpha} = n! \prod_{j < i} (r_j - r_i) / r_1! r_2! \cdots r_h!.$$

Exercises

- (1) A group is called *ambivalent* if each element is conjugate to its inverse. Such a group, if non-trivial, must have even order. Show that the symmetric group S_n of any degree n is ambivalent. (It can be shown that the alternating group A_n is ambivalent precisely when $n = 1, 2, 5, 6, 10, 14$.)
- (2) Compute the characters of S_4, S_5 by the methods of this section.
- (3) For each partition α of n define the *conjugate* partition as α' , where α'_i is the number of parts of α that are $\geq i$. Describe the relation between the Young diagrams of α and α' , and show that $\alpha'' = \alpha$.
- (4) Show that the character corresponding to the conjugate partition is given by $\chi_{\alpha'} = \varepsilon \chi_{\alpha}$, where ε is the sign character.

- (5) Show that for $n \geq 4$, S_n has at least two irreducible representations of degree $n - 1$, and show that for $n = 6$ it has four.
- (6) Show that any irreducible representation of S_n of degree greater than 1 is faithful except the representation corresponding to 2^2 for $n = 4$.

7.7 Induced representations

There are various relations between the representations of a group and those of its subgroups which are often needed. In the first place, if G is a group and H a subgroup, then each G -module V is an H -module by restriction of the action to H . This H -module is denoted by $\text{res}_H^G V$ or $\text{res } V$ or also V_H ; if the representation afforded by V is ρ , then the corresponding representation of V_H is written $\text{res}_H^G \rho$ or ρ_H .

We next describe a process of passing from a representation of a subgroup of G to one of G . Let H be a subgroup of finite index r in G and consider a right H -module U . From it we can form a right G -module

$$\text{ind}_H^G U = U^G = U \otimes_H kG, \quad (1)$$

called the *induced* G -module. To find the representation afforded by U^G we note that the subspace $U \otimes H = \{u \otimes 1 \mid u \in U\}$ of U^G is an H -module in a natural way: for any $h \in H$ we have $(u \otimes 1)h = uh \otimes 1$. More generally we can define the subspace $U \otimes Ha = \{u \otimes a \mid u \in U\}$ of (1) as an H^a -module, where $H^a = a^{-1}Ha$, by the rule

$$(u \otimes a)x = uaxa^{-1} \otimes a \quad \text{for } x \in H^a.$$

This is well-defined because $axa^{-1} \in H$ precisely when $x \in H^a$. When $U \otimes Ha$ is considered as right H^a -module in this way we shall denote it by U^a .

Now take a coset representation of G :

$$G = Ht_1 \cup \dots \cup Ht_r. \quad (2)$$

With its help we can write (1) as

$$U^G = U \otimes Ht_1 \oplus \dots \oplus U \otimes Ht_r = U^{t_1} \oplus \dots \oplus U^{t_r}. \quad (3)$$

Each U^{t_λ} is an H^{t_λ} -module and under the action of G these terms are permuted among themselves.

Given $x \in G$, we can for each $\lambda = 1, \dots, r$ find a unique $h \in H$ and μ , $1 \leq \mu \leq r$, such that $t_\lambda x = ht_\mu$. Then we have

$$(u \otimes t_\lambda)x = uh \otimes t_\mu;$$

this shows how the action of x permutes the terms in (3), as well as acting on each. To find the representation afforded by U^G , let us take a basis u_1, \dots, u_n of U and let $\rho(h) = (\rho_{ij}(h))$ ($h \in H$) be the corresponding representation of H . Then the elements $u_i \otimes t_\lambda$ form a basis of U^G and we have, for any $x \in G$,

$$(u_i \otimes t_\lambda)x = u_i \otimes t_\lambda x = u_i \otimes ht_\mu,$$

where μ is chosen so that $t_\lambda x t_\mu^{-1} = h \in H$. Hence we find

$$(u_i \otimes t_\lambda)x = \sum_j \rho_{ij}(h)u_j \otimes t_\mu.$$

Thus from a representation $(\rho(h))$ of degree d we obtain the induced representation $(\rho(t_\lambda x t_\mu^{-1}))$ of degree rd , where $r = (G:H)$. We note that this representation consists of $r d \times d$ blocks, one in each row and one in each column:

$$\rho^G = \begin{pmatrix} P_{11} & \cdots & P_{1r} \\ \vdots & & \vdots \\ P_{r1} & \cdots & P_{rr} \end{pmatrix},$$

where

$$P_{\lambda\mu} = \begin{cases} \rho(t_\lambda x t_\mu^{-1}) & \text{if } t_\lambda x t_\mu^{-1} \in H, \\ 0 & \text{otherwise.} \end{cases}$$

This is called the *induced* representation and is denoted by $\text{ind}_H^G \rho$ or ρ^G . Hence if the character of ρ is α , then ρ^G has the character $\text{ind}_H^G \alpha = \alpha^G$, given by

$$\begin{aligned} \alpha^G(x) &= \text{tr}(\rho^G(x)) \\ &= \sum_{\lambda} \text{tr} P_{\lambda\lambda} \\ &= \sum \text{tr}(\rho(t_\lambda x t_\lambda^{-1})), \end{aligned}$$

where the sum is over all λ such that $t_\lambda x t_\lambda^{-1} \in H$. Let us define $\dot{\alpha}$ on G by

$$\dot{\alpha}(x) = \begin{cases} \alpha(x) & \text{if } x \in H \\ 0 & \text{otherwise.} \end{cases}$$

Bearing in mind that $\alpha(h) = \text{tr}(\rho(h))$, we can rewrite the equation for $\alpha^G(x)$ as

$$\alpha^G(x) = \sum_{\lambda} \dot{\alpha}(t_\lambda x t_\lambda^{-1}). \quad (4)$$

Since α is a class function on H , we have $\alpha(ht_\lambda x t_\lambda^{-1} h^{-1}) = \alpha(t_\lambda x t_\lambda^{-1})$ for any $h \in H$. It follows that

$$\begin{aligned} |H| \cdot \alpha^G(x) &= \sum_{h \in H} \sum_{\lambda} \dot{\alpha}(ht_\lambda x t_\lambda^{-1} h^{-1}) \\ &= \sum_{g \in G} \dot{\alpha}(gxg^{-1}). \end{aligned}$$

If we define α^g by $\alpha^g(x) = \dot{\alpha}(gxg^{-1})$, we can express this as a formula

$$\alpha^G = |H|^{-1} \sum_{g \in G} \alpha^g. \quad (5)$$

We remark that for any class function α on H this formula defines a class function α^G on G . To illustrate (5), consider the case $H = 1$. There is only one character, namely the trivial character 1. We see from (1) that the induced representation is just the regular representation of G . More generally, let us take an arbitrary subgroup H of G but take the trivial character 1_H on H . Then we obtain as

induced character

$$\text{ind}_H^G(1_H(x)) = n_x,$$

where for any $x \in G$, n_x is the number of elements $t_\lambda x t_\lambda^{-1}$ ($\lambda = 1, \dots, r$) that lie in H .

It is clear from (1) that ind_H^G is a covariant functor; thus

$$\text{ind}_H^G(U \oplus U') \cong \text{ind}_H^G U \oplus \text{ind}_H^G U'. \quad (6)$$

Further, if $K \subseteq H \subseteq G$, then $U \otimes_K kG \cong U \otimes_K kH \otimes_H kG$; hence

$$\text{ind}_H^G(\text{ind}_K^H U) \cong \text{ind}_K^G U. \quad (7)$$

It follows that for any character α of K (indeed, for any class function),

$$(\alpha^H)^G = \alpha^G. \quad (8)$$

The following important relation between induction and restriction was proved by Frobenius in 1898:

THEOREM 7.1 (Frobenius reciprocity theorem) *Let G be a finite group and H a subgroup.*

(i) *Given a G -module V and an H -module U , we have the natural isomorphism:*

$$\text{Hom}_H(U, \text{res}_H^G V) \cong \text{Hom}_G(\text{ind}_H^G U, V). \quad (9)$$

(ii) *If α, β are class functions on H, G respectively and $\alpha^G = \text{ind}_H^G \alpha$ is defined as in (5), then*

$$(\alpha, \text{res}_H^G \beta)_H = (\text{ind}_H^G \alpha, \beta)_G, \quad (10)$$

where the subscripts denote the groups on which the scalar products are taken.

Proof. By adjoint associativity (cf. 4.7), we have

$$\begin{aligned} \text{Hom}_G(\text{ind}_H^G U, V) &\cong \text{Hom}_G(U \otimes_H kG, V) \\ &\cong \text{Hom}_G(kG, \text{Hom}_H(U, \text{res}_H^G V)) \\ &\cong \text{Hom}_H(U, \text{res}_H^G V). \end{aligned}$$

This establishes (9). For (10) we have

$$\begin{aligned} (\alpha^G, \beta)_G &= |G|^{-1} \sum_{x \in G} \alpha^G(x) \overline{\beta(x)} = |G|^{-1} |H|^{-1} \sum_{x, g} \dot{\alpha}(x^g) \overline{\beta(x)} \\ &= |G|^{-1} |H|^{-1} \sum \dot{\alpha}(x^g) \overline{\beta(x^g)} \\ &= |H|^{-1} \sum \dot{\alpha}(x^g) \overline{\beta(x^g)} \\ &= |H|^{-1} \sum_{x \in H} \alpha(x) \overline{\beta(x)} \\ &= (\alpha, \text{res}_H^G \beta)_H. \end{aligned}$$

When α, β are characters, (10) is of course an immediate consequence of (9). Since every class function is a linear combination of characters, this provides another proof of (10). ■

We next derive a formula due to Mackey in 1951, which describes the effect of inducing and then restricting a representation. Before stating it let us recall that for any group G with subgroups H, K we have a double coset decomposition

$$G = \bigcup HaK \quad (11)$$

into disjoint sets. It is obtained by letting the direct product $H \times K$ act on G by the rule $x \mapsto h^{-1}xk$ for $x \in G, (h, k) \in H \times K$. Each orbit has the form HaK , for some $a \in G$, and (11) is the partition of G into its orbits. The element $(h, k) \in H \times K$ leaves $a \in G$ fixed iff $h^{-1}ak = a$, i.e. $k = a^{-1}ha$; hence the stabilizer of a is $H^a \cap K$, and so we obtain the following formula for the size of a double coset:

$$|HaK| = |H| \cdot |K| / |H^a \cap K|.$$

THEOREM 7.2 (Mackey) *Let G be a finite group, H, K subgroups of G and (11) a double coset decomposition of G for H, K . If V is any H -module and $H_i = H^{a_i} \cap K$, then*

$$\text{res}_K \text{ind}_H^G(V) = \bigoplus_i \text{ind}_{H_i}^K \text{res}_{H_i}(V^{a_i}). \quad (12)$$

Hence if α is the character of the representation afforded by V and α^g is again defined by $\alpha^g(x) = \alpha(x^{g^{-1}})$, for $x, x^g \in H$, then

$$\text{res}_K \text{ind}_H^G(\alpha) = \sum_i \text{ind}_{H_i}^K \text{res}_{H_i}(\alpha^{a_i}). \quad (13)$$

Proof. Take a coset decomposition $G = \bigcup Ht$ of G for H ; then $\text{ind}_H^G V = \bigoplus V^t$, and for fixed i , the set $\{Ht | t \in Ha_i K\}$ is an orbit for the action of K on the coset space G/H . Hence we obtain

$$\text{ind}_H^G V = \bigoplus W_i,$$

where $W_i = \bigoplus \{V^t | t \in Ha_i K\}$ is a K -module. Fix $a = a_i$ and put $W = W_i$. For any $y, z \in K$ we have

$$Hay = Haz \Leftrightarrow yz^{-1} \in H^a \cap K = D, \text{ say.}$$

Hence if $K = \bigcup Dy$ is a coset decomposition of K over D , then $HaK = \bigcup Hay$, and so $H^a K = \bigcup H^a y$. Now

$$W = \bigoplus (V \otimes ay) = \bigoplus V^a y;$$

here $V \otimes a = V^a$ is an H^a -module, hence also a D -module, and we have

$$W = ((V^a)_D)^K = \text{ind}_D^K(V^a)_D,$$

by the definition of induced module. Now (12) follows by summing over i and restricting to K , and (13) follows by taking characters in (12). ■

Let us apply this result in the special case $K = H$.

PROPOSITION 7.3 *Let G be a finite group, H a subgroup with double coset decomposition $G = \bigcup Ha_iH$, and α a character of H . Then the induced character $\text{ind}_H^G\alpha$ is irreducible if and only if α is irreducible and $(\alpha^{a_i}, \text{res}_{H_i}(\alpha))_{H_i} = 0$ for all $a_i \notin H$, where $H_i = H^{a_i} \cap H$.*

Proof. Write $\beta = \text{ind}_H^G\alpha$; by Frobenius reciprocity (10) we have

$$(\beta, \beta)_G = (\alpha, \text{res}_H\beta)_H, \quad (14)$$

and by Mackey's formula (13),

$$\text{res}_H\beta = \sum_i \text{ind}_{H_i}^H \text{res}_{H_i}(\alpha^{a_i}).$$

Applying Frobenius reciprocity once again, we have

$$(\alpha, \text{ind}_{H_i}^H \text{res}_{H_i}(\alpha^{a_i}))_H = (\text{res}_{H_i}(\alpha), \alpha^{a_i})_{H_i} = d_i \text{ say.}$$

Substituting into (14), we obtain

$$(\beta, \beta)_G = \sum_i d_i.$$

If the double coset $H = HH$ is represented by a_1 , then $d_1 = (\alpha, \alpha) \geq 1$. For β to be irreducible, we must have $(\beta, \beta)_G = 1$, i.e. $d_1 = 1$ and $d_i = 0$ for $i > 1$, and these are just the conditions stated. ■

In particular, when H is normal in G , then $H_i = H$, $\text{res}_{H_i}\alpha = \alpha$ and we find

COROLLARY 7.4 *Let G be a finite group and H a normal subgroup. For any character α of H , $\text{ind}_H^G\alpha$ is irreducible if and only if α is irreducible and different from α^a for all $a \notin H$.* ■

Exercises

- (1) Show that if $N \triangleleft G$ and α is a class function on N , then $\text{ind}_N^G\alpha$ vanishes outside N .
- (2) Let α be a character of a subgroup H of G and define α^G as in (5). Show directly that $(\alpha^G, \chi)_G$ is a non-negative integer for every irreducible character χ of G , and deduce that α^G is a character of G .
- (3) Given a group G with subgroup H and characters α, β on H, G respectively, show that $\text{ind}_H^G(\alpha, \text{res}_H\beta) = (\text{ind}_H^G\alpha)\beta$.

- (4) Given a group G and subgroups H, K , show that if α, β are characters on H, K respectively, then

$$(\text{ind}_H^G \alpha, \text{ind}_K^G \beta)_G = \sum (\alpha^x, \beta^y)_{H^x \cap K^y},$$

where $x^{-1}y$ runs over a set of double coset representatives.

- (5) (A. H. Clifford) Let G be a finite group and H a normal subgroup. Show that if V is a simple G -module and U a simple H -submodule of V (*qua* H -module), then Ux is a simple H -module for each $x \in G$. Deduce that V is semisimple as H -module. Show also that for each simple H -type α the α -socle of V has the same length.

7.8 Applications: the theorems of Burnside and Frobenius

In this section we shall apply the earlier work to prove some results of pure group theory. In the first place we have the $p^\alpha q^\beta$ -theorem of Burnside, which states that any finite group whose order is divisible by only two distinct primes is soluble. This was proved by Burnside in 1904 using character theory; for odd primes a purely group theoretic proof was found by Thompson about 1960, but this is far more complicated.

Let G be a finite group; we denote its conjugacy classes again by $\{C_\lambda\}$ and denote the sum of the elements in C_λ (in the group algebra kG) by c_λ and their number by h_λ . We recall from (10) of 7.4 that for any irreducible character χ of degree d ,

$$h_\lambda \chi^{(\lambda)} = \eta_\lambda d, \quad (1)$$

where η_λ is given by $\rho(c_\lambda) = \eta_\lambda I$, ρ being the corresponding representation. Further we saw in 7.4 that both $\chi^{(\lambda)}$ and η_λ are algebraic integers.

LEMMA 8.1 *Let G be a finite group, ρ a complex irreducible representation of G of degree d with character χ and let C_λ be a conjugacy class of G with h_λ elements. If h_λ is prime to d , then either χ vanishes on C_λ or $\rho(x) = c$, 1 for all $x \in C_\lambda$, for some $c \in \mathbb{C}$.*

Proof. We have $\chi^{(\lambda)} = \omega_1 + \dots + \omega_d$, where each ω_i is a root of 1; hence $|\chi^{(\lambda)}| \leq d$, with strict inequality unless all the ω_i are equal. In the latter case the elements of C_λ are represented by scalars; otherwise we have $|\chi^{(\lambda)}/d| < 1$, and the same holds for all conjugates of $\chi^{(\lambda)}$ over \mathbb{Q} . Since d is prime to h_λ , there exist integers a, b such that $ah_\lambda + bd = 1$, and so

$$\begin{aligned} \chi^{(\lambda)} &= (ah_\lambda + bd)\chi^{(\lambda)} = ad\eta_\lambda + bd\chi^{(\lambda)}, \quad \text{by (1),} \\ &= d(a\eta_\lambda + b\chi^{(\lambda)}). \end{aligned}$$

Now both η_λ and $\chi^{(\lambda)}$ are algebraic integers, hence so is $\chi^{(\lambda)}/d$, and the same holds for its conjugates. Therefore the norm $N(\chi^{(\lambda)}/d)$ is an integer; as we have seen, it is

less than 1 in absolute value, so $N(\chi^{(\lambda)}/d) = 0$, and it follows that $\chi^{(\lambda)} = 0$, as claimed. ■

This lemma has the following important consequence. We recall that for any element x of a finite group G the number of conjugates of x in G is the index $(G:C_x)$ where C_x is the centralizer of x in G (cf. Vol. 1, p. 55). In what follows we understand by a *simple* group a simple non-abelian group.

THEOREM 8.2 (Burnside) *If a finite group G has a conjugacy class $C_\lambda \neq \{1\}$ consisting of p^m elements, where p is a prime, then G cannot be simple.*

Proof. If $m = 0$, then the unique element in C_λ lies in the centre of G , so G has a non-trivial centre and cannot be simple. We may therefore assume that $m \geq 1$. Let d_1, \dots, d_r be the degrees of the irreducible representations of G ; we have seen in (10) in 7.3 that

$$\sum d_i^2 = |G|, \quad (2)$$

and here the right-hand side is divisible by p , because $p^m = |C_\lambda|$ is the index of a centralizer. On the left of (2) we have $d_1 = 1$, so two cases can arise. Either there are more linear representations of G ; since their number is $(G:G')$, we have $G' \subset G$ in this case by Prop. 4.2, and so G is not simple. Or there are representations ρ_i of G of degree $d_i > 1$ and prime to p . Choose such a representation ρ_i and take $x \in C_\lambda$; by Lemma 8.1, either $\chi_i^{(\lambda)} = 0$ or $\rho_i(x)$ is a scalar. In the latter case either ρ_i is not faithful or x lies in the centre of G ; each time it follows that G is not simple. This only leaves the alternative $\chi_i^{(\lambda)} = 0$. Thus for any character χ_j we have either $\chi_j^{(\lambda)} = 0$ or $\chi_j^{(1)} \equiv 0 \pmod{p}$, except for $j = 1$, when $\chi_1^{(1)} = \chi_1^{(\lambda)} = 1$. But by orthogonality we have $\sum_j \chi_j^{(1)} \chi_j^{(\lambda)} = 0$, hence $1 \equiv 0 \pmod{p}$, which is a contradiction. ■

Now it is an easy matter to deduce the solubility condition.

THEOREM 8.3 (Burnside) *Every group of order $p^\alpha q^\beta$, where p, q are primes, is soluble.*

Proof. Let Q be a Sylow q -subgroup; its order is q^β . If $\beta = 0$, the result follows because G is then a p -group and so has a conjugacy class with p^m elements, hence by Th. 8.2, G is not simple (in any case this is a well-known fact, proved in 9.8, p. 291 of Vol. 1). Otherwise take $x \neq 1$ in the centre of Q (which we know is non-trivial, cf. Th. 3 of 9.8, Vol. 1). The centralizer C of x in G contains Q , hence $(G:C) = p^m$, where $m \leq \alpha$. It follows that x has p^m conjugates. By Th. 8.2, G cannot be simple, and now an induction on $|G|$ shows G to be soluble. ■

Secondly we prove a result of Frobenius in 1901 on the existence of complements in a group. If G is any group and H, K are subgroups such that

$H \cap K = 1$, $HK = G$, then each of H, K is called a *complement* of the other; this term is used particularly when one of H, K is normal in G . The main result is preceded by a lemma which is used in the proof.

LEMMA 8.4 *Let G be a finite group and H a non-trivial subgroup such that $H^x \cap H = 1$ for $x \notin H$. If α is a class function on H such that $\alpha(1) = 0$ and $\alpha^G = \text{ind}_H^G \alpha$, then $(\alpha^G)_H = \alpha$ and for any class function β on H , $(\alpha^G, \beta^G)_G = (\alpha, \beta)_H$.*

Proof. We have $\alpha^G(1) = (G:H)\alpha(1) = 0$, by hypothesis. Let us fix $h \in H$, such that $h \neq 1$. Then by definition, we have

$$\alpha^G(h) = |H|^{-1} \sum_{x \in G} \alpha(h^x). \quad (3)$$

Consider a term in this sum; if $\alpha(h^x) \neq 0$, then $h^x \in H^x \cap H$, so $x \in H$ and $\alpha(h^x) = \alpha(h)$. Thus all the non-zero terms in the sum on the right of (3) are $\alpha(h)$, and there is a term for each $x \in H$. Hence

$$\alpha^G(h) = |H|^{-1} \sum_{x \in H} \alpha(h^x) = \alpha(h),$$

and this proves the first assertion. For any other class function β on H , Frobenius reciprocity gives $(\alpha^G, \beta^G)_G = ((\alpha^G)_H, \beta)_H = (\alpha, \beta)_H$, by what has been proved. ■

The actual result of Frobenius can be stated in two equivalent forms, in terms of permutation groups or abstractly.

THEOREM 8.5 (Frobenius) *Let G be a finite transitive permutation group such that no element $\neq 1$ has more than one fixed point. Then the elements without fixed point, together with 1, form a normal subgroup N of G .*

The normal subgroup N is called the *Frobenius kernel*. We remark that if H is the stabilizer of a point, then $NH = G$ by transitivity (see the proof below), while $N \cap H = 1$ from the hypothesis of the theorem. Thus N is a complement of H , and the theorem may also be stated in the following form:

THEOREM 8.6 *Let G be a finite group with a subgroup H such that*

$$H^x \cap H = 1 \quad \text{for all } x \notin H, \quad (4)$$

i.e. H intersects each of its conjugates in 1 and is its own normalizer. Then H has a normal complement in G .

A subgroup H with the property (4) is called a *Frobenius subgroup*. So Th. 8.6 states that every Frobenius subgroup has a normal complement.

Proof. Let us first show the equivalence of Ths. 8.5 and 8.6. Under the hypotheses of Th. 8.5 let H be the stabilizer of a point. If $H \triangleleft G$, then all points

have the same stabilizer. In that case $H = G$ or 1 and the conclusion follows with $N = 1$ or G respectively, so we may exclude this case. Further, the conditions of Th. 8.5 mean that any $x \in G \setminus H$ moves the point fixed by H to another point, not fixed by any element in H^+ (where $H^+ = H \setminus \{1\}$), so that $H^x \cap H = 1$, while the elements moving all points make up precisely the set $G \setminus \bigcup H^x$. Conversely, given the hypotheses of Th. 8.6, we see on taking the coset representation on G/H that no element $\neq 1$ fixes more than one point and the elements moving all points comprise the complement of the union of all the conjugates of H in G . If they, together with 1, form a subgroup (and this is what we have to prove), then it must be a complement of H in G .

Let us denote this set, viz. $(G \setminus (\bigcup H^x) \cup \{1\})$, by N . It is clear that N is a normal set (i.e. a union of conjugacy classes), and writing $|G| = n$, $|H| = m$, $n = mr$, we have $|N| = n - r(m - 1) = r$, so if N is a subgroup, it will be a complement because $r = (G:H)$ and clearly $H \cap N = 1$. So all we need show is that N is a subgroup. We shall obtain it as the kernel of a certain representation.

Let α be any class function on H and put $\bar{\alpha} = \alpha - \alpha(1) \cdot 1_H$, where 1_H is the trivial character on H . Then $\bar{\alpha}(1) = 0$, and $\bar{\alpha}^G$ is a class function on G . Put

$$\alpha^* = \bar{\alpha}^G + \alpha(1) \cdot 1_G. \quad (5)$$

Then

$$\begin{aligned} (\alpha^*, \alpha^*)_G &= (\bar{\alpha}^G, \bar{\alpha}^G)_G + 2\alpha(1)(\bar{\alpha}^G, 1_G)_G + \alpha(1)^2 \\ &= (\bar{\alpha}, \bar{\alpha})_H + 2\alpha(1)(\bar{\alpha}, 1_H)_H + \alpha(1)^2, \end{aligned}$$

by Lemma 8.4. Hence

$$(\alpha^*, \alpha^*)_G = (\bar{\alpha} + \alpha(1)1_H, \bar{\alpha} + \alpha(1)1_H)_H = (\alpha, \alpha)_H.$$

If α is an irreducible character, then α^* is either a character or a difference of characters of G , i.e. a *virtual* character, and moreover $(\alpha^*, \alpha^*)_G = (\alpha, \alpha)_H = 1$, so either α^* or $-\alpha^*$ is an irreducible character of G ; in fact it must be α^* , because $\alpha^*(1) = \bar{\alpha}^G(1) + \alpha(1) > 0$. Further we have by Lemma 8.4,

$$\text{res}_H \alpha^* = \text{res}_H \bar{\alpha} + \alpha(1)1_H = \bar{\alpha} + \alpha(1)1_H = \alpha.$$

Now consider $K = \bigcap \ker \alpha^*$, where the intersection is taken over all irreducible characters α on H and \ker indicates the kernel of the corresponding representation. Then $K \triangleleft G$ and for any irreducible character α on H and $x \in H \cap K$ we have

$$\alpha(x) = \alpha^*(x) = \alpha^*(1);$$

hence $H \cap K = 1$ by Cor. 4.9. Since K is normal, we also have $H^x \cap K = 1$ for all $x \in G$, and so $K \subseteq N$. Here equality holds, for if $x \in N^+$, then $x \notin H^y$ for all $y \in G$, hence $\bar{\alpha}^G(x) = 0$ for all irreducible characters α of H . So in that case $\alpha^*(x) = \alpha^*(1)$ by (5), and so $x \in K$ by Prop. 4.8. This shows that $N = K$ and it proves N to be a subgroup. ■

So far no proof of this result without representation theory is known, although special cases (e.g. for soluble groups) have been proved in this way. Burnside has

shown that the Sylow p -subgroups of the Frobenius subgroup are either cyclic or, in case $p = 2$, generalized quaternion groups, and Thompson has shown that the Frobenius kernel is nilpotent, cf. Huppert (1967).

Exercises

- (1) Let the finite group G be a split extension of N by a subgroup H . Show that H is a Frobenius subgroup iff H acts freely on N^+ , i.e. for any $x \in N^+, h \in H^+, x^h \neq x$.
- (2) Let G be a finite group with a Frobenius subgroup H . Show that under the action of H on the coset space G/H there is one orbit of length 1, while all the others have length $|H|$. Deduce that $|H|$ divides $(G:H) - 1$ and that the Frobenius kernel N has order prime to its index.
- (3) Show that if a conjugacy class C of a non-trivial group G has p^m elements, where p is prime, then the set $C^{-1}C$ generates a proper normal subgroup of G .
- (4) Let G be a Frobenius group, acting on a set of order p^m , where p is a prime. Give a direct proof of the existence of a Frobenius kernel.
- (5) A subgroup K of a finite group G is called a *Hall subgroup* if its order is prime to its index. Show that every normal Hall subgroup of G is characteristic in G .
- (6) Verify that in a dihedral group of order $2m$, where m is odd, the cyclic subgroup of order m is a Frobenius kernel.

Further exercises on Chapter 7

- (1) Let G be a soluble group of order divisible by a prime p . If $N = G_{i-1}/G_i$ is a chief factor of order p^r of G (cf. Th. 5 of **9.6**, p. 276 of Vol. 1), show that the action of G on N induced by inner automorphisms is a representation of degree r over the field \mathbf{F}_p of p elements.
- (2) Show that an irreducible representation of a p -group over a field of characteristic p is necessarily trivial. Find a (reducible) representation of degree 2 of \mathbf{C}_p over \mathbf{F}_p which is faithful.
- (3) Show that any representation of a finite group G over \mathbf{Q} is equivalent to a representation over \mathbf{Z} . (*Hint.* Take a basis (u_i) of the representation module, let N be a common denominator for all the representation coefficients and consider the subgroup A of $\sum \mathbf{Z} u_i$ generated by all the expressions $N u_i x$ ($x \in G$). Show that A is again a G -module and find a \mathbf{Z} -basis for it.)
- (4) Show that in an irreducible representation of a finite group G over an algebraically closed field k , each element of the centre of G is represented by a scalar matrix. Show that this holds even if k is not algebraically closed but contains a primitive $|G|$ th root of 1.

- (5) Show that if G is a group with a faithful irreducible representation over a field of characteristic 0, then the centre of G must be cyclic. (*Hint.* Use Ex. (4) to prove first the special case where the ground field contains a primitive $|G|$ th root of 1.)
- (6) Show that for a permutation representation of G with character χ the number of orbits is $(\chi, 1_G)$, where 1_G is the trivial character.
- (7) Show that if a character χ of a faithful representation assumes r distinct values on G , then the Vandermonde determinant $(\chi^{(\lambda)i})$ is non-zero. Deduce that every irreducible character is a constituent of at least one of $1_G, \chi, \chi^2, \dots, \chi^{r-1}$.
- (8) Show that a doubly transitive permutation representation is the sum of the trivial representation and one other irreducible representation. (*Hint.* The stabilizer must be transitive. Now use Th. 4.11 and the orthogonality relations.)
- (9) For each representation $\rho(x) = (\rho_{ij}(x))$ of degree d define d^2 elements of the group algebra as $\rho_{ij} = \sum \rho_{ij}(x)x$. Show that the orthogonality relations may be stated as follows. For any irreducible representations ρ, σ that are inequivalent, $\rho_{ij}\sigma_{pq} = 0$, while $\rho_{ij}\rho_{pq} = (|G|/d)\delta_{jp}\rho_{iq}$.
- (10) (J. A. Green) Show that if V is an irreducible G -module over a finite field F , then $E = \text{End}_G(V)$ is a finite field. Further if $|E| = q$, then $V \otimes V \otimes \cdots \otimes V$ (n factors) has exactly $(q^n - 1)/(q - 1)$ irreducible submodules.
- (11) Let G be a finite group and $X = (\chi_i^{(\lambda)})$ its character table. Show that any automorphism of G induces a permutation of the rows of X , and a permutation of the columns, where the effect of α on X is defined by $\chi_i^\alpha(x) = \chi_i(x^\alpha)$. If $P(\alpha), Q(\alpha)$ are the permutation matrices describing the effects of α on the rows and columns of X respectively, then $X^\alpha = P(\alpha)X = XQ(\alpha)$. Deduce that $P(\alpha), Q(\alpha)$ are conjugate and hence have the same trace. Hence show that for any group A of automorphisms acting on G the number of orbits of the set of irreducible characters is the same as the number of orbits of the conjugacy classes of G under the action of A .
- (12) Let G be a Frobenius group with kernel K and complement H . Show that for any non-trivial irreducible character α of K , $\text{ind}_K^G \alpha$ is an irreducible character of G . (*Hint.* Consider the group of automorphisms of K induced by H ; show that $c \in H^+$ fixes no conjugacy class $\neq \{1\}$ of K , and use Ex. (11) to show that for any non-trivial character χ of K , $\chi^c \neq \chi$. Now apply reciprocity to show that $\text{ind}_K^G \alpha$ is irreducible.)
- (13) Let G, K, H be as in Ex. (12). Show that any irreducible character χ of G is either trivial on K or induced up from an irreducible character ψ of K . Deduce that for such χ, ψ and any $c \in H$, $\text{res}_{H^+}^G \chi(c) = \delta_{c1} \psi(1) \cdot |H|$.
- (14) Show that the affine group $A = \text{Aff}_1(\mathbf{F}_q)$ of all transformations $x \mapsto ax + b$ ($a, b \in \mathbf{F}_q, a \neq 0$) is a Frobenius group with the translation group as kernel and the stabilizer of 0 as complement.

8

Valuation theory

Valuation theory may be described as the study of divisibility (in commutative rings) in its purest form, but that is only one aspect. The general formulation leads to the introduction of topological concepts like completion, which provides a powerful tool. It also emphasizes the parallel with the absolute value on the real and complex numbers. After the initial definitions we shall prove the essential uniqueness of the absolute value on \mathbf{R} and \mathbf{C} (in 8.2), and go on to describe the p -adic numbers, before looking at simple cases of the extension problem.

8.1 Divisibility and valuations

Let R be an integral domain with field of fractions K ; our object will be to study the divisibility in R . We note that the divisibility relation can be defined for K as well by writing for any $a, b \in K^\times$, $a|b$ (*a divides b*) whenever $b = ac$ for some $c \in R$. This relation is reflexive and transitive, thus it is a preordering of K^\times . Moreover, this preordering is preserved by multiplication:

$$a|b \Rightarrow ad|bd \quad \text{for all } a, b, d \in K^\times,$$

and so K^\times may be regarded as a preordered group. It is not generally ordered, because $a|b, b|a$ need not imply $a = b$; but if we define

$$a \sim b \quad \text{whenever} \quad a|b \quad \text{and} \quad b|a,$$

then we obtain an equivalence on K^\times . Its classes are just the classes of associated elements in K^\times and these classes form a partially ordered group Γ . If $v(a)$ denotes the class of $a \in K^\times$, then $a \mapsto v(a)$ is a homomorphism from K^\times to Γ . We shall write the operation in Γ as addition and put $v(a) \leq v(b)$ or $v(b) \geq v(a)$ to indicate that $a|b$. Then the mapping v satisfies the conditions:

$$\mathbf{V.1} \quad v(a), v(b) \geq \gamma \Rightarrow v(a+b) \geq \gamma, \quad \text{for all } \gamma \in \Gamma,$$

$$\mathbf{V.2} \quad v(ab) = v(a) + v(b),$$

$$\mathbf{V.3} \quad a \in R \Leftrightarrow v(a) \geq 0.$$

It is convenient to have v defined on the whole of K , including 0. To achieve this we shall introduce a symbol ∞ satisfying $\infty + \gamma = \infty$, while $\gamma < \infty$ for all

$\gamma \in \Gamma$. If we now put $v(0) = \infty$, then V. 1–3 hold for all $a, b \in K$ and $\gamma \in \Gamma \cup \{\infty\}$. Of course this is just a formal device, designed to ensure that our formulae hold without exception; in defining v we can ignore $v(0)$ since its value is prescribed.

Given a field K , it is clear that any function v from K^\times to a partially ordered group Γ , satisfying V. 1, 2, determines the subset $R = \{x \in K \mid v(x) \geq 0\}$, which is easily seen to be a subring, and V. 3 holds for this R . So here we have another way of describing R , which stresses the divisibility in R . In general this offers no advantage over the usual description, but it suggests looking at rings for which Γ has a particularly simple form.

An important simplification is obtained by assuming Γ to be totally ordered. Then V. 1 can be replaced by

$$\mathbf{V.1'} \quad v(a + b) \geq \min \{v(a), v(b)\}.$$

In this case v is called a *valuation* on K and K is *valuated*. Thus a valuation on K is a mapping $v: K^\times \rightarrow \Gamma$ to a totally ordered group Γ , together with $v(0) = \infty$, satisfying V. 1' and V. 2. The ring R defined by V. 3 is then called the *valuation ring*, or the ring of *valuation integers*. It is easy to characterize valuation rings directly:

PROPOSITION 1.1 *Let K be a field. Then a subring V of K is a valuation ring in K if and only if for any $x \in K^\times$, either $x \in V$ or $x^{-1} \in V$.*

Proof. Let v be a valuation on K and $V = \{x \in K \mid v(x) \geq 0\}$ the ring of integers. Then for any $x \in K^\times$, $v(x) + v(x^{-1}) = v(xx^{-1}) = v(1) = 0$; hence either $v(x) \geq 0$ or $v(x^{-1}) \geq 0$, so V satisfies the given condition. Conversely, let V be a subring of K satisfying this condition and denote by Γ the group of classes of associated elements, defined as above. Then for any value γ in Γ , $\gamma = v(x)$ for some $x \in K^\times$. Now either $x \in V$ or $x^{-1} \in V$ and accordingly $\gamma \geq 0$ or $-\gamma \geq 0$; this means that Γ is totally ordered. ■

Frequently the value group Γ will be still further restricted. Thus if $\Gamma = \mathbf{R}$, we speak of a *real-valued* valuation. When $\Gamma = \mathbf{Z}$, the simplest non-trivial case, we speak of a *principal valuation*; this is sometimes called a *discrete rank 1 valuation*.

Every field has the *trivial* valuation, given by

$$v(x) = \begin{cases} 0 & \text{if } x \neq 0, \\ \infty & \text{if } x = 0. \end{cases}$$

Examples. Here are other examples, important in what follows:

1. The p -adic valuation on \mathbf{Q} . Every rational number can, up to sign, be written uniquely as a product of powers of different primes. Fix a prime p ; for any $a \in \mathbf{Q}^\times$ let p^v (where $v \in \mathbf{Z}$) be the exact power of p occurring in a and define

$v(a) = v$. This is easily seen to be a valuation, called the p -adic valuation. For example, taking $p = 3$, we have $v(111) = 1$, $v(10/9) = -2$. We remark that this valuation is completely determined by its values on \mathbf{Z} . This is a general property: If K is the field of fractions of an integral domain A , then any valuation defined on A has a unique extension to K (cf. Ex. (2)).

2. On $k(x)$, the field of rational functions in x over a field k , a valuation may be defined by writing any element φ of $k(x)^\times$ in reduced form: $\varphi = f/g$, $f, g \in k[x]$, $(f, g) = 1$, and setting $v(\varphi) = \deg g - \deg f$. The integers in this valuation are the rational functions remaining finite at ∞ . Other valuations are obtained by singling out an irreducible polynomial p over k and defining the value of φ as the exponent of p in a complete factorization of φ . In particular, for $k = \mathbf{C}$ (or any algebraically closed field) any irreducible polynomial has the form $x - \alpha$ (up to a constant factor) and the integers for the corresponding valuation are the rational functions that remain finite at $x = \alpha$. Thus the valuations we have found correspond to the different places on the Riemann sphere.

In what follows we shall mainly be concerned with principal valuations; although some of the results will hold for the general case, we shall not always mention this fact explicitly.

If v is a principal valuation on a field K , then the image $v(K^\times)$ is a subgroup of \mathbf{Z} . It is either 0, when v is trivial, or it is of the form $v(K^\times) = r\mathbf{Z}$ for some $r > 0$, and hence is isomorphic to \mathbf{Z} . In that case $(1/r)v(x)$ is again a valuation, this time with value group exactly \mathbf{Z} . We call the valuation *normalized* if the value group is the whole of \mathbf{Z} and express the preceding statement by saying: every non-trivial valuation can be normalized. Nevertheless, it is not always expedient to normalize our valuations; e.g. a normalized valuation on K may be non-trivial on a subfield F and yet not normalized on F .

We now turn to consequences of the definition. As in every homomorphism of groups, the neutral element is preserved, i.e. $v(1) = 0$. It follows that

$$v(-1) = 0 \quad \text{and} \quad v(-x) = v(x) \quad \text{for all } x, \quad (1)$$

for $v(-1) + v(-1) = v((-1)^2) = v(1) = 0$, hence $v(-1) = 0$, and now $v(-x) = v(-1) + v(x) = v(x)$.

The following relation will be frequently needed.

LEMMA 1.2 *For any x, y in a field with a valuation v ,*

$$v(x - y) \geq \min \{v(x), v(y)\}, \quad (2)$$

and here the inequality is strict unless $v(x) = v(y)$.

Proof. From V.1' we have $v(x - y) \geq \min \{v(x), v(-y)\} = \min \{v(x), v(y)\}$, by (1), and this proves (2). If $v(x) \neq v(y)$, say $v(x) > v(y)$, then $v(y) \geq \min \{v(x), v(x - y)\}$ and $v(x - y) \geq v(y)$ by (2), hence $v(x - y) = v(y)$. Similarly if $v(y) > v(x)$. ■

By an easy induction we find that $v(\sum_1^n a_i) \geq \min \{v(a_1), \dots, v(a_n)\}$, and hence we obtain, as in Lemma 1.2,

$$\text{If } \sum_1^n a_i = 0, \text{ then } v(a_i) = v(a_j) \text{ for some } i \neq j. \quad (3)$$

If we think of $v(x)$ as indicating the degree of divisibility of x by a certain prime, property (2) is obvious. Later (in 8.2) we shall meet a less obvious interpretation.

We have already seen that associated with every valuation v there is a ring

$$V = \{x \in K \mid v(x) \geq 0\},$$

the ring of *valuation integers*. The set of all non-units in this ring is

$$\mathfrak{p} = \{x \in K \mid v(x) > 0\};$$

this is easily verified to be an ideal, by V.1', 2, and it is maximal, since a proper ideal cannot contain any units. Moreover, it is the unique maximal ideal of V . The quotient ring V/\mathfrak{p} is easily seen to be a field \bar{K} , called the *residue class field* of the valuation v .

If v is the trivial valuation, then $V = K$, $\mathfrak{p} = 0$ and the residue class field is just K itself. Leaving this case aside, we may characterize the valuation rings of principal valuations as follows:

PROPOSITION 1.3 *Let K be a field with a valuation v and valuation ring V . Then v is a principal valuation if and only if V is a principal ideal domain, and in that case K contains an element p such that*

$$K^\times = \langle p \rangle \times U,$$

where $\langle p \rangle$ is the (multiplicative) cyclic group on p and U is the group of units of V . Thus every element a of K^\times has the form $a = p^n u$, where $n \in \mathbb{Z}$ and $u \in U$, and if v is normalized, then $v(a) = n$; $a \in V$ if and only if $n \geq 0$.

The element p is called a *uniformizer* or also a *prime element* of v ; it is determined up to a unit factor.

Proof. Let v be a principal valuation and take an element p of least positive value. By taking v to be normalized we may assume that $v(p) = 1$; hence for any $a \in K^\times$, if $v(a) = n \in \mathbb{Z}$, then $v(ap^{-n}) = v(a) - nv(p) = 0$, so $ap^{-n} = u \in U$, and we can therefore write $a = p^n u$, where $n = v(a)$, $u \in U$. It is clear that the representation is unique once p has been chosen. Moreover, V is principal; all its ideals have the form 0 or (p^n) , $n \geq 0$.

Conversely, assume that V is principal; then its maximal ideal can be written

in the form (p) . We claim that $\bigcap(p^n) = 0$. For if not, suppose that $\bigcap(p^n) = (q)$; then $q = a_n p^n$ for all $n \geq 0$, hence $qp^{-1} = a_n p^{n-1}$ for all $n \geq 1$, so $qp^{-1} \in \bigcap(p^n) = (q)$, say $qp^{-1} = qb$, $q(1 - pb) = 0$. But $q \neq 0$, hence $pb = 1$, a contradiction, and this shows that $\bigcap(p^n) = 0$. Now take $a \in K^\times$; by what has been proved, $a \notin (p^{n+1})$ for some n . Choose the least such n ; then $a \in (p^n)$, so $a = p^n u$, $u \in V$, and here u is a unit, for if $u \in (p)$, we would have $a \in (p^{n+1})$. Thus we have expressed a in the form $a = p^n u$ ($u \in U$), and this form is unique, for if $p^n u = p^m v$, where $u, v \in U$ and $n \geq m$, say, then $p^{n-m} = vu^{-1} \in U$, hence $n = m$, $v = u$. It follows that $v(a) = nv(p)$ for $a \in K^\times$, and v is normalized iff $v(p) = 1$. ■

A valuation ring which is also a principal ideal domain but not a field is called a *principal valuation ring*. This term (due to Mumford) seems preferable to the usual term for such rings: ‘discrete valuation rings of rank 1’.

Two valuations v_1, v_2 on a field K are said to be *equivalent* if there is an order-isomorphism θ between their value groups such that $v_2(x) = v_1(x)^\theta$ for all $x \in K$. It is clear from this definition that equivalent valuations have the same valuation ring:

THEOREM 1.4 *On any field K there is a bijection between the equivalence classes of valuations on K and valuation rings in K . In this correspondence principal valuation rings correspond to principal valuations.* ■

We have already briefly mentioned some examples of valuations. Generally, in any UFD R with field of fractions K , we can write each element of K^\times in the form

$$a = up_1^{\alpha_1} \cdots p_r^{\alpha_r} \quad (\alpha_i \in \mathbf{Z}),$$

where p_1, \dots, p_r are pairwise non-associated atoms of R . We single out an atom p and write $a = p^v a'$, where a' is prime to p (i.e. the exponent of p in the factorization of a' is 0). If we now put $v(a) = v$, this provides a valuation on K associated with the atom p , called again the *p-adic valuation* on K . Here the valuation ring V is the set of all elements $p^n a'$, with $n \geq 0$, i.e. the set of elements with denominator prime to p , while the maximal ideal in V is the set of all elements with numerator divisible by p .

In the particular case $R = \mathbf{Z}$ we get a *p-adic valuation* for each prime number p (introduced by K. Hensel in 1908). The elements of V are then called *p-adic integers*; e.g. $5/24$ is a 7-adic integer, but not a 3-adic integer. No other valuations exist on \mathbf{Q} , by

PROPOSITION 1.5 *The only non-trivial valuations on \mathbf{Q} are the p-adic valuations.*

Proof. By Th. 1.4 it is enough to determine all valuation rings on \mathbf{Q} . Let V be a valuation ring on \mathbf{Q} , with maximal ideal \mathfrak{p} say. Since $1 \in V$, it follows that $\mathbf{Z} \subseteq V$. Now $\mathbf{Z} \cap \mathfrak{p}$ is a prime ideal in \mathbf{Z} ; either $\mathbf{Z} \cap \mathfrak{p} = 0$, then every non-zero element of \mathbf{Z}

is a unit in V , so $V = \mathbf{Q}$ and the valuation is trivial. Or $\mathbf{Z} \cap \mathfrak{p} = p\mathbf{Z}$ for some prime number p ; then every $n \in \mathbf{Z}$ is either divisible by p or prime to p , and hence is a unit in V . This means that we have the p -adic valuation on \mathbf{Q} . ■

If we examine what makes this proof work, we find that it depends essentially on the fact that \mathbf{Z} is a principal ideal domain; e.g. the proof also applies to the polynomial ring $k[x]$ over a field k . Here the valuations of the field of fractions $k(x)$ which are trivial on k are the p -adic valuations for the different irreducible polynomials p in $k[x]$, and one extra valuation is associated with the degree. For let V be a valuation ring in $k(x)$ containing k and assume first that $x \in V$; then $k[x] \subseteq V$, and the same argument as in the proof of Prop. 1.5 shows that the valuation is associated with some irreducible polynomial, because every maximal ideal of $k[x]$ has such a polynomial as generator. If $x \notin V$, then $y = x^{-1} \in V$ and the same conclusion holds with x replaced by y . Moreover, $y \in \mathfrak{p}$, so if we are given $f = a_0 + a_1x + \dots + a_nx^n$ ($a_n \neq 0$), then $f = (a_0y^n + \dots + a_n)y^{-n}$, and so $v(f) = -n$. Thus we get an extra valuation $v(f) = -\deg f$. In particular, when k is algebraically closed, then the valuations correspond to the elements of k , together with a point ‘at infinity’, as we have seen in the case $k = \mathbf{C}$. The residue class field is k at each point, while the valuation indicates the *order* of the function at the given place: if $f = (x - \alpha)^n u$ or $f = x^{-n} u$, n indicates the order to which f vanishes at α or at ∞ .

Exercises

- (1) Show that a finite field has only the trivial valuation.
- (2) Let v be a mapping from a commutative ring R to a totally ordered group Γ (with $v(0) = \infty$) such that V.1' and V.2 hold. Show that R is an integral domain and if K is its field of fractions, then there is exactly one extension of v to a valuation on K , defined by $v(a/b) = v(a) - v(b)$.
- (3) Show that the relation $a|b$ in an integral domain R is an ordering iff 1 is the only unit in R . Give examples of such rings.
- (4) Show that a principal ideal domain with a single maximal ideal is either a principal valuation ring or a field.
- (5) Show that every automorphism of the additive group of \mathbf{R} which preserves the natural order is of the form $x \mapsto \lambda x$, where $\lambda > 0$. Deduce that two real-valued valuations v_1, v_2 are equivalent iff $v_2(x) = \lambda v_1(x)$ for all $x \in K$ and a fixed $\lambda > 0$.

8.2 Absolute values

It is possible to interpret a valuation as a distance function. This is a fruitful approach, which enables us to introduce metric notions such as completion. We begin with a quite general definition.

Let R be a commutative ring. An *absolute value* on R is a real-valued function $x \mapsto |x|$ on R such that

- A.1 $|x + y| \leq |x| + |y|$ (triangle inequality),
- A.2 $|xy| = |x| \cdot |y|$,
- A.3 $|x| \geq 0$ with equality iff $x = 0$.

Any non-trivial ring R with an absolute value, briefly a *valued* ring, is an integral domain. For $1 \neq 0$ by definition, and $x, y \neq 0$, then $|xy| = |x| \cdot |y| \neq 0$, hence $xy \neq 0$. Conversely, any integral domain has an absolute value: we put

$$|x| = \begin{cases} 1 & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases} \quad (1)$$

This is called the *trivial* absolute value, all others are *non-trivial*. A field with the trivial absolute value is said to be discretely valued, or *discrete*. If K is a field with a real-valued valuation v , then we can define an absolute value on K by putting

$$|x| = 2^{-v(x)}.$$

A.2,3 are immediate, while A.1 holds in the stronger form

- A.1' $|x + y| \leq \max \{|x|, |y|\}$ (ultrametric inequality).

As in Lemma 1,2 we derive from A.1' the consequence

$$|x - y| \leq \max \{|x|, |y|\}, \quad \text{with equality unless } |x| = |y|. \quad (2)$$

Geometrically this may be expressed by saying that 'all triangles are isosceles'.

An absolute value is called *non-archimedean* if it satisfies A.1', *archimedean* otherwise. If $| \cdot |$ is non-archimedean, then by setting $v(x) = -\log_2 |x|$ we obtain a valuation; hence the non-archimedean absolute values just correspond to the real-valued valuations, with the trivial absolute value corresponding to the trivial valuation. An example of an archimedean absolute value is the usual absolute value on \mathbf{Q} , or more generally, the absolute value defined on any archimedean ordered field, as in 6.6.

The following criterion for a non-archimedean absolute value is often useful. In the proof we shall need an elementary limit: for any real positive α ,

$$\lim_{n \rightarrow \infty} (1 + n\alpha)^{1/n} = 1. \quad (3)$$

To prove (3) we note that $(1 + n\alpha)^{1/n} \geq 1$ for all n ; further, for any $\delta > 0$ we have $(1 + \delta)^n \geq 1 + n\alpha$ for all sufficiently large n (by the binomial theorem), and so $(1 + n\alpha)^{1/n} \leq 1 + \delta$. Since δ was arbitrary, (3) follows.

PROPOSITION 2.1 *For any absolute value $| \cdot |$ on a field K the following*

conditions are equivalent:

- (a) $| \cdot |$ is non-archimedean,
- (b) $|n \cdot 1| \leq 1$ for all $n \in \mathbf{Z}$,
- (c) $|n \cdot 1|$ is bounded for all $n \in \mathbf{N}$.

Proof. (a) \Rightarrow (b). If $| \cdot |$ is non-archimedean, then $|n \cdot 1| \leq \max\{|1|, \dots, |1|\} = 1$, so (b) holds. (b) \Rightarrow (c) is clear, and to prove (c) \Rightarrow (a), suppose that $|n \cdot 1| \leq M$ for all $n \in \mathbf{N}$; then for any $x, y \in K$ we have

$$|x + y|^n = |(x + y)^n| = \left| \sum \binom{n}{i} x^i y^{n-i} \right| \leq \sum M |x|^i |y|^{n-i}.$$

Hence

$$|x + y|^n \leq (n+1)M \cdot \max\{|x|^n, |y|^n\}.$$

Taking n th roots, we find

$$|x + y| \leq (1+n)^{1/n} M^{1/n} \max\{|x|, |y|\};$$

letting n tend to ∞ and remembering (3), we obtain the ultrametric inequality, hence $| \cdot |$ must be non-archimedean. ■

Since (c) clearly holds in finite characteristic, we deduce

COROLLARY 2.2 *A field with an archimedean absolute value must be of characteristic 0.* ■

If R is a ring with an absolute value $| \cdot |$, we can define a metric on R by putting $d(x, y) = |x - y|$. This makes R into a metric space; the ring operations are continuous by A.1,2, so that we have a topological ring. We shall indicate how to form the completion of such a ring, rather in the fashion in which \mathbf{R} was formed from \mathbf{Q} in 6.6. The same method applies, because the construction in 6.6 depended not on the ordering of \mathbf{Q} but only on the ordering of the absolute values. We shall sketch the method, using the fact that \mathbf{R} is already known to be complete.

As in 6.6, we shall say that a sequence $\{c_v\}$ of elements of R converges to $c \in R$ if $|c_v - c| \rightarrow 0$ as $v \rightarrow \infty$. By a *Cauchy sequence* we understand a sequence $\{c_v\}$ such that $|c_\mu - c_v| \rightarrow 0$ as $\mu, v \rightarrow \infty$. As before we see that every convergent sequence is a Cauchy sequence; if the converse holds, R is said to be *complete* with respect to the absolute value. To form the completion we take the set C of all Cauchy sequences over R and verify that this is a ring under componentwise addition and multiplication, with the constant sequences ($a_v = a$ for all v) as a subring isomorphic to R . The sequences convergent to 0, the *null sequences*, again form an ideal \mathfrak{n} in C ; here we use the fact that for every Cauchy sequence $\{c_v\}$ the sequence $|c_v|$ is bounded. Let us embed R in C by identifying $a \in R$ with the constant

sequence $\{a, a, \dots\}$. Then $R \cap \mathfrak{n} = 0$, so if we set $\tilde{R} = C/\mathfrak{n}$, we have a mapping $R \rightarrow C \rightarrow C/\mathfrak{n} = \tilde{R}$, which is an embedding, because the kernel is $R \cap \mathfrak{n} = 0$, by the second isomorphism theorem.

We extend the absolute value to \tilde{R} by putting $|\{c_v\}| = \lim |c_v|$; by the completeness of \mathbf{R} this defines an absolute value on \tilde{R} , which extends the given absolute value on R . Moreover, R is dense in \tilde{R} , i.e. every element of \tilde{R} is a limit of elements of R . Suppose further that R is a field; then so is \tilde{R} . For if $c \in \tilde{R}$, say $c = \lim c_v$, and $c \neq 0$, then as in 6.6 we can show that c_v^{-1} is bounded for all $v > v_0$ and $c_\mu^{-1} - c_v^{-1} = c_\mu^{-1}(c_v - c_\mu)c_v^{-1} \rightarrow 0$ as $\mu, v \rightarrow \infty$, and so $\{c_v^{-1}\}_{v > v_0}$ is again a Cauchy sequence, convergent to the element c^{-1} of \tilde{R} . It only remains to show that \tilde{R} is complete and is determined up to isomorphism by R : this follows as before. Summing up, we have

THEOREM 2.3 *Let R be a commutative absolute valued ring. Then there exists a complete absolute valued ring \tilde{R} and an embedding $R \rightarrow \tilde{R}$ preserving absolute values, such that the image is dense in \tilde{R} , and \tilde{R} is uniquely determined by R , up to (metric) isomorphism. If R is a field, then so is \tilde{R} . ■*

The ring \tilde{R} is again called the *completion* of R with respect to the absolute value. For example, consider \mathbf{Q} with the p -adic valuation v_p , for a given prime p . The completion is called the *field of p -adic numbers* and is denoted by \mathbf{Q}_p . The elements x of \mathbf{Q}_p such that $v_p(x) \geq 0$ form a subring \mathbf{Z}_p , the *ring of p -adic integers*. The elements of \mathbf{Q}_p may be written in the form of a series

$$a = \sum_{i=-k}^{\infty} a_i p^i \quad (0 \leq a_i < p). \quad (4)$$

If $a_i = 0$ for $i < 0$, we have an element of \mathbf{Z}_p . The finite series form a subring of \mathbf{Z}_p , isomorphic to \mathbf{Z} ; here (4) reduces to the expression of an ordinary integer in the base of p .

We note that if the absolute value on a field K corresponds to a principal valuation, then the value groups for K and its completion \tilde{K} are the same; for $v(K)$ is a discrete subgroup of \mathbf{R} , hence closed, and $v(a) = \lim v(a_v)$ whenever $a_v \rightarrow a \in \tilde{K}$. The residue class field likewise is unchanged by completion: let V, V' be the valuation rings and $\mathfrak{p}, \mathfrak{p}'$ their maximal ideals in K, \tilde{K} respectively. Then $V \subseteq V'$, $\mathfrak{p} \subseteq \mathfrak{p}'$; since $V \cap \mathfrak{p}' \supseteq \mathfrak{p}$ and $V \cap \mathfrak{p}' \neq V$, we have $V \cap \mathfrak{p}' = \mathfrak{p}$ by the maximality of the latter, and $V + \mathfrak{p}' = V'$, because any $c \in V'$ can be written $c = c_0 + c_1$, where $c_0 \in V, c_1 \in \mathfrak{p}'$. Hence $V'/\mathfrak{p}' = (V + \mathfrak{p}')/\mathfrak{p}' \cong V/(V \cap \mathfrak{p}') = V/\mathfrak{p}$.

We have seen that an absolute value defines a topology. Two absolute values on a field K are said to be *equivalent* if they induce the same topology on K . Any absolute value on a field inducing the discrete topology must be trivial: if $| \cdot |$ is not trivial, then for some $a \in K$, $|a| \neq 0, 1$, hence $a \neq 0$ and $x = a$ or $x = a^{-1}$ satisfies $|x| < 1$; it follows that $x^n \rightarrow 0$. The relation between equivalent absolute values can be described quite explicitly as follows:

PROPOSITION 2.4 *Let $|\cdot|_1, |\cdot|_2$ be any two non-trivial absolute values on a field K . Then the following conditions are equivalent:*

- (a) $|\cdot|_1$ is equivalent to $|\cdot|_2$,
- (b) $|x|_1 < 1 \Rightarrow |x|_2 < 1$ for all $x \in K$,
- (c) $|x|_1 = |x|_2^\gamma$ for all $x \in K$, and some real constant γ .

The third condition shows that the notion defined here agrees with the definition of equivalence of valuations given in 8.1. We also note that not every value of γ in (c) will give an absolute value. Given an absolute value $|\cdot|$, the function $|x|^\gamma$ is again an absolute value for $0 < \gamma \leq 1$, and in certain cases (e.g. if $|\cdot|$ is non-archimedean) γ may be taken > 1 , but not always.

Proof. (a) \Rightarrow (b). If (a) holds and $|x|_1 < 1$, then $x^n \rightarrow 0$ in the $|\cdot|_1$ -metric, hence also in the $|\cdot|_2$ -metric, and so $|x|_2 < 1$.

(b) \Rightarrow (c). We first show that (b) is in fact symmetric, i.e.

$$|x|_2 < 1 \Leftrightarrow |x|_1 < 1. \quad (5)$$

If this were not so, then for some $a \in K$, $|a|_2 < 1$ and $|a|_1 \geq 1$, or on writing $b = a^{-1}$, $|b|_1 \leq 1$ and $|b|_2 > 1$. Choose $c \in K$ such that $0 < |c|_1 < 1$; then by (b), $0 < |c|_2 < 1$, hence $|b^n c|_1 < 1$ for all $n \geq 0$, so $|b^n c|_2 < 1$, i.e. $|b|_2^n < |c|_2^{-1}$ for all n , which is possible only if $|b|_2 \leq 1$. This contradiction shows that (5) holds.

Let us write $f_i(x) = -\log|x|_i$ ($i = 1, 2$); then (5) takes the form

$$f_1(x) > 0 \Leftrightarrow f_2(x) > 0, \quad (6)$$

and we must show that $f_2(x) = cf_1(x)$, bearing in mind that $f_i(xy) = f_i(x) + f_i(y)$, by the definition and A.2. Take $a \in K^\times$ such that $f_1(a) > 0$; then for any $x \in K$ and integers m, n we have

$$\begin{aligned} f_1(x) > (m/n)f_1(a) &\Leftrightarrow f_1(x^n a^{-m}) > 0 \\ &\Leftrightarrow f_2(x^n a^{-m}) > 0 \\ &\Leftrightarrow f_2(x) > (m/n)f_2(a). \end{aligned}$$

Thus for all rational r we have

$$f_1(x) > rf_1(a) \Leftrightarrow f_2(x) > rf_2(a).$$

Letting r tend to $f_1(x)/f_1(a)$ from below, we find that

$$\frac{f_2(x)}{f_2(a)} \geq \frac{f_1(x)}{f_1(a)};$$

by symmetry we have equality here, thus

$$\frac{f_2(x)}{f_1(x)} = \frac{f_2(a)}{f_1(a)} = c,$$

for all $x \in K$, which shows that (c) holds. Finally, to prove that (c) \Rightarrow (a), we clearly have for any $\gamma > 0$, $|a_v|_1 \rightarrow 0 \Leftrightarrow |a_v|_1^\gamma \rightarrow 0$. ■

Just as this result shows that equivalent absolute values are very much alike, so the next shows that inequivalent ones are very different.

THEOREM 2.5 (Approximation theorem, Artin–Whaples 1945) *Let $|\cdot|_1, \dots, |\cdot|_n$ be non-trivial absolute values on a field K which are pairwise inequivalent. Then for any $a_1, \dots, a_n \in K$ and any $\varepsilon > 0$, there exists $\alpha \in K$ such that*

$$|\alpha - a_i|_i < \varepsilon \quad \text{for } i = 1, \dots, n.$$

Proof. First we observe that for any absolute value $|\cdot|$, as $r \rightarrow \infty$,

$$\lim\left(\frac{a^r}{1 + a^r}\right) = \begin{cases} 0 & \text{if } |a| < 1, \\ 1 & \text{if } |a| > 1. \end{cases} \quad (7)$$

Our next objective is to find $c \in K$ such that

$$|c|_1 > 1, \quad |c|_i < 1 \quad \text{for } i = 2, \dots, n. \quad (8)$$

Since $|\cdot|_1$ and $|\cdot|_2$ are inequivalent, we can by Prop. 2.4 find $a \in K$ such that $|a|_1 > 1 \geq |a|_2$ and $b \in K$ such that $|b|_1 \leq 1 < |b|_2$. Now $c = ab^{-1}$ satisfies $|c|_1 > 1 > |c|_2$. This is (8) in case $n = 2$; we may therefore take $n > 2$ and use induction on n . By hypothesis there exists $a \in K$ such that $|a|_1 > 1 > |a|_i$ ($i = 3, \dots, n$) and there exists $b \in K$ such that $|b|_1 > 1 > |b|_2$. Either $|a|_2 \leq 1$; then we put $c_r = a^r b$; now $|c_r|_1 > 1 > |c_r|_2$ for all r , and if r is large enough, then $|c_r|_i < 1$ for $i = 3, \dots, n$ also. Or $|a|_2 > 1$; then we put $c_r = a^r b / (1 + a^r)$. By (7), as $r \rightarrow \infty$, $|c_r|_1 \rightarrow |b|_1 > 1$, $|c_r|_2 \rightarrow |b|_2 < 1$, $|c_r|_i \rightarrow 0$ ($i = 3, \dots, n$), hence (8) is satisfied by $c = c_r$ when r is large enough.

Take c satisfying (8); from (7) it follows that the sequence $c^r / (1 + c^r)$ tends to 1 at $|\cdot|_1$ and to 0 at $|\cdot|_i$ for $i = 2, \dots, n$. Thus, given $\delta > 0$, and $1 \leq i \leq n$, we can find $u_i \in K$ such that $|u_i - 1|_i < \delta$, $|u_i|_j < \delta$ for $j \neq i$. We take such u_i for $i = 1, \dots, n$ and put $\alpha = \sum a_i u_i$. Then

$$|\alpha - a_i|_i \leq |a_i(u_i - 1)|_i + \sum_{j \neq i} |a_j u_j|_i < \delta n M,$$

where $M = \max_{ij} \{|a_j|_i\}$. So by choosing $\delta < \varepsilon/nM$ we obtain the required element α . ■

COROLLARY 2.6 *If $|\cdot|_1, \dots, |\cdot|_r$ are inequivalent non-trivial absolute values, then there is no relation*

$$|x|_1^{r_1} \cdots |x|_n^{r_n} = 1 \quad \text{for all } x \in K,$$

except the trivial one, where $r_1 = \cdots = r_n = 0$.

For if $r_1 \neq 0$ say, choose $x \in K$ so that $|x|_1$ is small and $|x - 1|_i$ is small for $i = 2, \dots, n$. Then $|x|_i^{r_i}$ is near 1 for $i > 1$ and the given relation cannot hold. ■

By contrast, for infinitely many absolute values there is a product formula of this form (see Ex. (6)).

Later we shall study methods of extending absolute values to extension fields, and it is convenient to treat the existence and uniqueness separately. The uniqueness can be proved under quite general conditions, namely for any finite-dimensional space over a complete valued field. We digress briefly to introduce the necessary definitions.

Let K be a field with an absolute value $|\cdot|$. A *normed vector space* over K is a vector space V over K with a function $x \mapsto \|x\|$ taking values in \mathbf{R} and satisfying the following conditions:

- N.1 $\|x + y\| \leq \|x\| + \|y\| \quad \text{for all } x, y \in V,$
- N.2 $\|\alpha x\| = |\alpha| \cdot \|x\| \quad \text{for } x \in V, \alpha \in K,$
- N.3 $\|x\| \geq 0, \quad \text{with equality if and only if } x = 0.$

Clearly any field extension, with an absolute value extending that of K , is a normed vector space. Moreover, any finite-dimensional vector space over K has at least one norm, the *cubical norm*, defined as follows. Pick a basis e_1, \dots, e_n in V and for $x = \sum \alpha_i e_i$ define $\|x\| = \max \{|\alpha_1|, \dots, |\alpha_n|\}$. It is straightforward to verify that this is a norm. We can define convergence, Cauchy sequences and completeness in V as in the case of fields, using the norm in place of the absolute value. Then we have

PROPOSITION 2.7 *Let V be a normed space of finite dimension over a complete absolute valued field K . Then V is complete and its topology is induced by any cubical norm on V ; thus the topology on V is uniquely determined.*

Proof. We first show that V is complete in the cubical norm. Let e_1, \dots, e_n be a basis of V and let $x^{(v)} = \sum \alpha_i^{(v)} e_i$ be a Cauchy sequence: $\|x^{(\mu)} - x^{(v)}\| \rightarrow 0$, hence $|\alpha_i^{(\mu)} - \alpha_i^{(v)}| \rightarrow 0$ for $i = 1, \dots, n$. By the completeness of K , $\lim \alpha_i^{(v)} = \alpha_i$ exists in K . Put $x = \sum \alpha_i e_i$; then $|\alpha_i - \alpha_i^{(v)}| \rightarrow 0$ as $v \rightarrow \infty$ ($i = 1, \dots, n$), hence $\|x - x^{(v)}\| \rightarrow 0$, i.e. $x^{(v)} \rightarrow x$, and so V is indeed complete in the cubical norm defined by the e_i .

Now let $N(x)$ be any norm on V ; we must show that this defines the same topology as the cubical norm $\|\cdot\|$. We use induction on $\dim V$, which is finite by hypothesis. For $\dim V = 0$ there is nothing to prove, so let $\dim V > 0$. In the first place, if $x = \sum \alpha_i e_i$, then

$$N(x) \leq \sum N(\alpha_i e_i) = \sum |\alpha_i| \cdot N(e_i) \leq \|x\| \cdot (\sum N(e_i)).$$

Thus $N(x) \leq c \|x\|$ for a fixed c , hence convergence in the cubical topology entails convergence in the N -topology (i.e. the cubical topology is finer than the N -

topology). Conversely, let $\{x^{(v)}\}$ be a sequence such that $N(x^{(v)}) \rightarrow 0$ and suppose that $\|x^{(v)}\| \not\rightarrow 0$. Write $x^{(v)} = \sum \alpha_i^{(v)} e_i$, so that $\|x^{(v)}\| = \max_i \{|\alpha_i^{(v)}|\}$. If $|\alpha_i^{(v)}| \rightarrow 0$, for each i , we would have $\|x^{(v)}\| \rightarrow 0$. This is not so, hence for some i , say $i = 1$, $|\alpha_i^{(v)}| \not\rightarrow 0$; going over to a subsequence, we may assume that $|\alpha_1^{(v)}| \geq \varepsilon$ for some $\varepsilon > 0$ and all v . We write $y^{(v)} = (\alpha_1^{(v)})^{-1} x^{(v)} = \sum \beta_i^{(v)} e_i$, where $\beta_1^{(v)} = 1$ by construction. Then

$$N(y^{(v)}) = |\alpha_1^{(v)}|^{-1} N(x^{(v)}) \leq \varepsilon^{-1} N(x^{(v)}),$$

hence again $N(y^{(v)}) \rightarrow 0$, and so $\sum \beta_i^{(v)} e_i \rightarrow -e_1$ in the N -topology. But e_2, \dots, e_n span an $(n-1)$ -dimensional subspace; by induction this has a unique topology and is complete, hence closed. Thus e_1 belongs to the subspace spanned by e_2, \dots, e_n , a contradiction. Therefore $\|x^{(v)}\| \rightarrow 0$ and the result follows. ■

Any discrete space is necessarily complete; hence we have

COROLLARY 2.8 *A finite-dimensional normed space over a discrete field is itself discrete.* ■

As a further consequence we have the uniqueness property of extensions.

THEOREM 2.9 *Let K be a complete valued field and L a finite algebraic extension field of K . Then the absolute value on K has at most one extension to L , and L is complete.*

Proof. Let $|\cdot|$ be the absolute value on K and let $|\cdot|_1, |\cdot|_2$ be two absolute values on L extending $|\cdot|$. By Prop. 2.7 both induce the same topology on L , and L is complete in this topology. If $|\cdot|$ is trivial, then K is discrete, hence so is L and $|\cdot|_1, |\cdot|_2$ are both trivial. Otherwise we have $|x|_1 = |x|_2^\gamma$ for all $x \in L$ and some γ , by Prop. 2.4, and there exists $a \in K$ such that $|a| \neq 0, 1$. Taking $x = a$ we see that $\gamma = 1$, hence $|\cdot|_1 = |\cdot|_2$. ■

Any absolute value on a finite field must be trivial, since the topology is discrete. More directly, in a field of q elements, every $x \neq 0$ satisfies $x^{q-1} = 1$, hence $|x|^{q-1} = 1$ and so $|x| = 1$. More generally, this clearly holds for any algebraic extension of a finite field.

We end this section by determining all complete valued fields with archimedean absolute value, following Ostrowski (with simplifications by E. Artin). It turns out that besides the two well-known examples of \mathbf{R} and \mathbf{C} there are no others.

THEOREM 2.10 (Ostrowski's first theorem) *Any non-trivial absolute value on \mathbf{Q} is equivalent either to the usual absolute value or to a p -adic valuation.*

Proof. Let f be any absolute value on \mathbf{Q} . If f is non-archimedean, then the

corresponding valuation must be p -adic for some prime p , as we have seen in Prop. 1.5. So we may assume that f is archimedean. We observe that for any $n \in \mathbb{N}$, $f(n) \leq f(1) + \dots + f(1) = n$, hence

$$f(n) \leq |n| \quad \text{for all } n \in \mathbb{Z}. \quad (9)$$

Given any integers $m, n > 1$, we express m in the base of n :

$$m = a_0 + a_1 n + \dots + a_v n^v, \quad \text{where } 0 \leq a_i < n, a_v \neq 0. \quad (10)$$

In particular, $m \geq n^v$ and so

$$v \leq \frac{\log m}{\log n}. \quad (11)$$

By (10), $f(m) \leq f(a_0) + f(a_1)f(n) + \dots + f(a_v)f(n)^v$, but $f(a_i) \leq a_i < n$, therefore

$$f(m) \leq n[1 + f(n) + \dots + f(n)^v].$$

According as $f(n) \leq 1$ or > 1 , we replace all terms in the brackets by the first or last term and so obtain

$$f(m) \leq n(1 + v) \max \{1, f(n)^v\}.$$

By (11) we can rewrite this as

$$f(m) \leq n \left(1 + \frac{\log m}{\log n} \right) \cdot \max \{1, f(n)^{\log m / \log n}\}.$$

In this formula replace m by m^r and then take r th roots:

$$f(m) \leq n^{1/r} \left(1 + \frac{r \cdot \log m}{\log n} \right)^{1/r} \cdot \max \{1, f(n)^{\log m / \log n}\}.$$

Letting $r \rightarrow \infty$ and remembering (3), we obtain

$$f(m) \leq \max \{1, f(n)^{\log m / \log n}\}. \quad (12)$$

Since f is archimedean, we know that $f(n_0) > 1$ for some n_0 , by Prop. 2.1. Taking $m = n_0$, we see that $f(n) > 1$ for all $n > 1$, so we can rewrite (12) as

$$f(m) \leq f(n)^{\log m / \log n},$$

i.e.

$$\frac{\log f(m)}{\log m} \leq \frac{\log f(n)}{\log n}.$$

By symmetry we have equality, say both sides equal γ . Then $\log f(m) = \gamma \cdot \log m$ for

all $m \geq 1$, i.e. $f(m) = m^\gamma$. It follows that $f(-m) = f(m) = m^\gamma$ and $f(m/n)^\gamma = (m/n)^\gamma$, hence f is equivalent to the usual absolute value on \mathbf{Q} , as claimed. ■

It remains to determine the archimedean absolute values on arbitrary fields. As a complete proof from first principles is rather long (and the result is not needed elsewhere in the book), we shall merely indicate how this follows from standard results in analysis.

THEOREM 2.11 (Ostrowski's second theorem) *Let K be a field with an archimedean absolute value for which K is complete. Then $K \cong \mathbf{R}$ or $K \cong \mathbf{C}$, and the absolute value is equivalent to the usual absolute value.*

Proof. Let f be the given absolute value on K . By Cor. 2.2, K has characteristic 0 and so contains \mathbf{Q} as a subfield. Moreover, since f is archimedean, $f(n) > 1$ for some n , so f is non-trivial on \mathbf{Q} and by Th. 2.10 there exists $\alpha \in \mathbf{R}$ such that

$$|x| = f(x)^\alpha \quad (13)$$

for all $x \in \mathbf{Q}$. We shall use (13) to define $| \cdot |$ over the whole of K ; this extends the usual absolute value of \mathbf{Q} and, moreover, it satisfies the triangle inequality on K . For we have

$$f(x + y) \leq f(x) + f(y) \leq 2 \max \{f(x), f(y)\},$$

hence by (13).

$$|x + y| \leq 2^\alpha \cdot \max \{|x|, |y|\}.$$

By repeated application we have for any $x_i \in K$,

$$|x_1 + \cdots + x_{2r}| \leq (2^r)^\alpha \cdot \max \{|x_1|, \dots, |x_{2r}|\}. \quad (14)$$

Given $x_1, \dots, x_n \in K$, we can choose r so that $2^{r-1} \leq n < 2^r$. If we apply (14) with this value of r , we get

$$|x_1 + \cdots + x_n| \leq (2n)^\alpha \cdot \max \{|x_1|, \dots, |x_n|\}.$$

In particular,

$$\begin{aligned} |a + b|^n &= \left| \sum \binom{n}{i} a^i b^{n-i} \right| \\ &\leq [2(n+1)]^\alpha \cdot \max \left\{ |a|^n, \binom{n}{1} |a|^{n-1} |b|, \dots, |b|^n \right\} \\ &\leq [2(n+1)]^\alpha \cdot (|a| + |b|)^n. \end{aligned}$$

Taking n th roots, we have

$$|a + b| \leq [2(n+1)]^{\alpha/n} (|a| + |b|),$$

and letting $n \rightarrow \infty$, we obtain (by (3)),

$$|a + b| \leq |a| + |b| \quad \text{for all } a, b \in K.$$

Thus we have an extension of $|\cdot|$ to K satisfying the triangle inequality as well as being multiplicative. Since K is complete, it must contain \mathbf{R} . If $x^2 + 1 = 0$ has a root in K , we adjoin this root to \mathbf{R} and find that $\mathbf{C} \subseteq K$. If not, we adjoin a root i of $x^2 + 1 = 0$ to K and obtain $\mathbf{C} \subseteq K(i)$, and we have to show that $K = \mathbf{C}$ or $K(i) = \mathbf{C}$ respectively. In either case we have a complete normed space over \mathbf{C} , where in the second case the norm is given by

$$\|x + iy\| = (|x|^2 + |y|^2)^{1/2}.$$

It is easily seen that this is indeed a norm on $K(i)$. By Prop. 2.7 $K(i)$ is then a complete space and it only remains to show that a field K which is a complete normed space over \mathbf{C} is \mathbf{C} itself (Gelfand–Mazur theorem). We shall outline the proof, referring e.g. to Rudin (1966) for details.

Suppose that $c \in K$, $c \notin \mathbf{C}$; then $c - \alpha \neq 0$ for all $\alpha \in \mathbf{C}$ and so $(c - \alpha)^{-1}$ is an element of K for each $\alpha \in \mathbf{C}$. We consider linear functionals on K , i.e. \mathbf{C} -linear mappings from K to \mathbf{C} . Such a functional f is bounded if $|f(x)| \leq \gamma \|x\|$ for some constant γ depending on f . When K is finite-dimensional over \mathbf{C} , it is algebraic over \mathbf{C} and hence $K = \mathbf{C}$, because \mathbf{C} is algebraically closed. In the infinite-dimensional case f is bounded iff it is continuous, or equivalently if the hyperplane $\ker f$ is closed. Given $b \in K^\times$, we can find f such that $f(b) \neq 0$, by constructing a closed hyperplane not containing b , and this can be done by Zorn's lemma (this is the assertion of the Hahn–Banach theorem).

Now return to $(c - \alpha)^{-1}$, and consider $f((c - \alpha)^{-1})$, for a bounded linear functional f . This is an analytic function of α for all $\alpha \in \mathbf{C}$ and it is bounded as $\alpha \rightarrow \infty$, hence by Liouville's theorem it is a constant, so for any $\alpha, \beta \in \mathbf{C}$,

$$f((c - \alpha)^{-1} - (c - \beta)^{-1}) = f((c - \alpha)^{-1}) - f((c - \beta)^{-1}) = 0.$$

Since this holds for all f , we must have $(c - \alpha)^{-1} - (c - \beta)^{-1} = 0$, which is a contradiction when $\alpha \neq \beta$. ■

Exercises

- (1) Show that an absolute value $N(x)$ on a field satisfies the triangle inequality whenever $N(x) \leq 1 \Rightarrow N(1 + x) \leq 2$, and it is ultrametric if $N(x) \leq 1 \Rightarrow N(x + 1) \leq 1$.
- (2) In a non-archimedean metric show that two balls either are disjoint or one is contained in the other.
- (3) Let $|\cdot|$ be a non-archimedean absolute value. Show that if an infinite series $\sum a_n$ is convergent and $|a_n| < |a_1|$ for all $n > 1$, then $|\sum a_n| = |a_1|$ (this is called the *principle of domination*).

- (4) On a field with a non-archimedean absolute value, an infinite series is convergent iff the n th term tends to 0.
- (5) Show that the trivial valuation of a field has no non-trivial extension to an algebraic extension field.
- (6) Show that the absolute values on \mathbf{Q} can be normed so that $\prod_i |a|_i = 1$ for all $a \neq 0$. Do the same for $k(x)$, where k is a finite field.
- (7) Show that every non-trivial subgroup of \mathbf{R} is either of the form $\alpha\mathbf{Z}$, where $\alpha > 0$, or is dense in \mathbf{R} . Deduce that every discrete \mathbf{R} -valued valuation is principal.
- (8) Let F be a complete field for an archimedean absolute value. Show directly (without using Ostrowski's theorem) that $1 + a$ is a square whenever $4|a| < |4|$.

8.3 The p -adic numbers

The definition of the p -adic field \mathbf{Q}_p as a completion of \mathbf{Q} in 8.2 suggests that the methods of analysis may be applicable. This is indeed the case, as was first observed by Hensel, and in this section we look at ways of solving equations over \mathbf{Q}_p . Throughout this section, p will be a fixed prime number, arbitrary except when otherwise specified. The p -adic valuation on \mathbf{Q}_p will be denoted by v or v_p .

Every natural number a has a p -adic expansion, i.e. we can express it in the base p :

$$a = a_0 + a_1 p + a_2 p^2 + \cdots + a_r p^r \quad (0 \leq a_i < p). \quad (1)$$

In terms of the valuation v_p we can think of a as being obtained by successive approximations: $a_0, a_0 + a_1 p, \dots, a$. Similarly, the general element of \mathbf{Z}_p can be written as an infinite series

$$b = b_0 + b_1 p + b_2 p^2 + \cdots; \quad (2)$$

it is the limit of the sequence of integers formed by its partial sums $b_0, b_0 + b_1 p, \dots$. For the elements of \mathbf{Q}_p we also have to allow a finite number of negative powers of p :

$$c = c_{-k} p^{-k} + c_{1-k} p^{1-k} + \cdots + c_{-1} p^{-1} + c_0 + c_1 p + \cdots. \quad (3)$$

For example, when $p = 7$, we have

$$-1 = 6 + 6.7 + 6.7^2 + \cdots. \quad (4)$$

Rational numbers are in \mathbf{Z}_7 as long as their denominator is prime to 7. Thus to find the 7-adic expansion of $1/2$ we have $1/2 = (-6+7)/2 = -3 + \frac{1}{2} \cdot 7 = -3 - 3.7 - 3.7^2 - \cdots$, hence

$$\frac{1}{2} = \frac{8-7}{2} = 4 - \frac{1}{2} \cdot 7 = 4 + 3.7 + 3.7^2 + \cdots. \quad (5)$$

With the help of this expansion we can solve $x^2 = 2$ in \mathbf{Z}_7 , using the binomial theorem. We have $x^2 = 2$ precisely when $(2x)^2 = 8 = 1 + 7$, hence $2x = (1 + 7)^{1/2}$, and

$$x = \frac{1}{2}(1 + 7)^{1/2} = \sum \frac{1}{2} \binom{1/2}{n} 7^n.$$

The latter is an element of \mathbf{Z}_7 , as we see by using (5). We shall soon meet a systematic way of solving such equations.

Clearly \mathbf{Z}_p is a principal valuation ring, the ideal of non-units being generated by p . If we adjoin an inverse for p , we obtain $\mathbf{Q}_p = \mathbf{Z}_p[p^{-1}]$. Let us write U for the group of units of \mathbf{Z}_p ; then every $x \in \mathbf{Q}_p^\times$ can be uniquely written in the form

$$x = p^v u, \quad \text{where } v = v(x) \text{ and } u \in U,$$

and of course $x \in \mathbf{Z}_p$ iff $v \geq 0$.

If in (2) we ignore all terms after p^{n-1} , we obtain an integer mod p^n ; thus we have a natural homomorphism $\varepsilon_n: \mathbf{Z}_p \rightarrow \mathbf{Z}/p^n$ which consists in mapping b , given by (2), to the residue class (mod p^n) of $b_0 + b_1 p + \dots + b_{n-1} p^{n-1}$. Clearly ε_n is surjective, with kernel $p^n \mathbf{Z}_p$, so we have an exact sequence

$$0 \rightarrow \mathbf{Z}_p \xrightarrow{p^n} \mathbf{Z}_p \xrightarrow{\varepsilon_n} \mathbf{Z}/p^n \rightarrow 0, \quad (6)$$

where p^n indicates multiplication by p^n . In particular this shows that

$$\mathbf{Z}_p/p^n \mathbf{Z}_p \cong \mathbf{Z}/p^n. \quad (7)$$

Let us consider polynomials with coefficients in \mathbf{Q}_p . On multiplying by a suitable power of p we may assume the coefficients to lie in \mathbf{Z}_p . Moreover, we can always arrange for the polynomial to be *primitive*; in the present case this means that at least one of the coefficients is a unit. If we apply ε_n to the coefficients of f , we obtain a polynomial with coefficients in \mathbf{Z}/p^n which we shall denote by $\varepsilon_n f$. Suppose that the equation $\varepsilon_n f = 0$ has a solution in \mathbf{Z}/p^n for all n ; we wish to show that $f = 0$ has a solution in \mathbf{Z}_p .

Let

$$a_n = a_{n0} + a_{n1}p + \dots + a_{n(n-1)}p^{n-1} \quad (0 \leq a_{ij} < p)$$

be a root of $\varepsilon_n f = 0$. Then we have an infinite sequence of integers

$$\begin{aligned} a_1 &= a_{10}, \\ a_2 &= a_{20} + a_{21}p, \\ a_3 &= a_{30} + a_{31}p + a_{32}p^2, \\ &\dots \end{aligned} \quad (8)$$

Consider the numbers in the first column on the right: a_{10}, a_{20}, \dots ; they can assume only finitely many values $0, 1, \dots, p-1$, so at least one of these values

must occur infinitely often. By choosing an appropriate subsequence of the equations (8) we may therefore assume that $a_{10} = a_{20} = \dots = b_0$, say. Next consider the sequence of coefficients of p in the equations that remain: a_{21}, a_{31}, \dots ; again we can pass to an infinite subsequence in which all the coefficients of p are the same, say b_1 . Continuing in this way, we obtain a p -adic integer

$$b = b_0 + b_1 p + \dots,$$

with the property that for each n_0 there is an $n > n_0$ such that $\varepsilon_n b = a_n$ and hence $\varepsilon_n f(b) = 0$. Thus $f(b) \equiv 0 \pmod{p^n}$ for arbitrarily large n , and it follows that $f(b) = 0$. This establishes

PROPOSITION 3.1 *Let $f \in \mathbf{Z}_p[x]$, where p is any prime number. If the equation $\varepsilon_n f = 0$ has a root in \mathbf{Z}/p^n for all n , then $f = 0$ has a root in \mathbf{Z}_p . More generally, this applies for polynomials in several variables.* ■

For the benefit of readers who have met inverse limits we remark that instead of forming \mathbf{Z}_p as a completion we can also construct it as an inverse limit of the rings \mathbf{Z}/p^n . Each of the latter, being finite, is compact and the diagonal argument in the proof of Prop. 3.1 can also be used to deduce the compactness of \mathbf{Z}_p from this fact.

In practice it is not necessary to solve each of the equations $\varepsilon_n f = 0$; it suffices to solve one or two and then improve the approximation by a method analogous to the Newton–Fourier rule (Vol. 1, **6.8**). We state the essential step separately in the next lemma, where f' as usual denotes the derivative of f .

LEMMA 3.2 *Given $f \in \mathbf{Z}_p[x]$, suppose that $\alpha \in \mathbf{Z}_p$ satisfies $f(\alpha) \equiv 0 \pmod{p^n}$ and $v(f'(\alpha)) = r$, where $0 \leq r < n$. Then $\beta = \alpha - f(\alpha)/f'(\alpha)$ is in \mathbf{Z}_p ,*

$$f(\beta) \equiv 0 \pmod{p^{n+1}}, \quad v(f'(\beta)) = r,$$

and

$$\beta \equiv \alpha \pmod{p^{n-r}}.$$

Proof. We write $\beta = \alpha + \gamma p^{n-r}$ and try to determine $\gamma \in \mathbf{Z}_p$. To this end consider $f(\alpha + h)$; if we regard this as a polynomial in h and expand it in powers of h , we obtain

$$f(\alpha + h) = f(\alpha) + h f'(\alpha) + h^2 g(h), \quad \text{where } g \in \mathbf{Z}_p[x]. \quad (9)$$

The form of the coefficient of h follows by Taylor's theorem, but we note that a straightforward application of Taylor's theorem to derive (9) would not make it clear that g has integral coefficients.

If we now put $h = \gamma p^{n-r}$ in (9), we get

$$f(\beta) = f(\alpha) + p^{n-r} \gamma f'(\alpha) + p^{2n-2r} c,$$

for some $c \in \mathbf{Z}_p$. By hypothesis, $f(\alpha) = ap^n$, $f'(\alpha) = up^r$, where $a \in \mathbf{Z}_p$ and $u \in U$, the unit group of \mathbf{Z}_p . Hence $f(\beta) = [a + \gamma u + p^{n-2r} c]p^n$; we can solve $a + \gamma u \equiv 0 \pmod{p^{n-r}}$ to obtain $\gamma \in \mathbf{Z}_p$.

p) for γ , because u is a unit, and with this value for γ we have $f(\beta) \equiv 0 \pmod{p^{n+1}}$ and $\beta \equiv \alpha \pmod{p^{n-r}}$. The last congruence shows that $f'(\beta) \equiv f'(\alpha) \pmod{p^{n-r}}$, hence $f'(\beta) \equiv p^r u \pmod{p^{n-r}}$, and since $n-r > r$, we see that $v(f'(\beta)) = r$. It is easily seen that $\gamma = -a/u$, and this gives the stated value for β . ■

Explicitly, if α is an approximate zero of f , then $\alpha - f(\alpha)/f'(\alpha)$ is a better approximation.

By iteration we obtain the general result:

THEOREM 3.3 *Let $f \in \mathbf{Z}_p[x_1, \dots, x_m]$, $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbf{Z}_p^m$ be such that*

$$f(\alpha) \equiv 0 \pmod{p^n}.$$

Suppose that for some j, r ($1 \leq j \leq m, 0 \leq 2r < n$), $v(D_j f(\alpha)) = r$, where $D_j f$ is the partial derivative with respect to x_j . Then there is a zero $\beta = (\beta_1, \dots, \beta_m)$ of f in \mathbf{Z}_p^m and $\beta_i \equiv \alpha_i \pmod{p^{n-r}}$.

Proof. Assume first that $m = 1$. We apply the lemma to $\alpha^{(0)} = \alpha$ and obtain $\alpha^{(1)} \in \mathbf{Z}_p$ such that $\alpha^{(1)} \equiv \alpha^{(0)} \pmod{p^{n-r}}$ and $f(\alpha^{(1)}) \equiv 0 \pmod{p^{n+1}}$, $v(f'(\alpha^{(1)})) = r$. By induction on k we obtain a sequence $\{\alpha^{(k)}\}$ such that

$$\alpha^{(k+1)} \equiv \alpha^{(k)} \pmod{p^{n+k-r}}, \quad (10)$$

and further $f(\alpha^{(k)}) \equiv 0 \pmod{p^{n+k}}$, $v(f'(\alpha^{(k)})) = r$. The congruence (10) shows that $\{\alpha^{(k)}\}$ is a Cauchy sequence; its limit β satisfies $f(\beta) = 0$, $\beta \equiv \alpha \pmod{p^{n-r}}$, and this is what we had to prove.

Now the general case is easily deduced by writing $f_1(x) = f(\alpha_1, \dots, \alpha_{j-1}, x, \alpha_{j+1}, \dots, \alpha_m)$ and applying the result just proved to f_1 . ■

When $n = 1, r = 0$, we obtain the usual form of Newton's rule:

COROLLARY 3.4 *Let $f \in \mathbf{Z}_p[x]$ and let \bar{f} be the polynomial obtained by reducing the coefficients mod p , thus $\bar{f} = \varepsilon_1 f$. Then every simple zero of \bar{f} in \mathbf{Z}/p can be lifted to a zero in \mathbf{Z}_p .* ■

For multiple zeros this breaks down, e.g. $x^2 + 1 = 0$ has a root $(\text{mod } 2)$, but no root in \mathbf{Z}_2 . A look at Th. 3.3 suggests that we reduce mod 4 (not mod 2), and indeed, $x^2 + 1 \equiv 0 \pmod{4}$ has no solutions. In the case of a double zero of f , Th. 3.3 tells us that we need to start from a root α of $f(x) \equiv 0 \pmod{p^3}$ when $p \neq 2$, and for $p = 2$ find a root of $f(x) \equiv 0 \pmod{4}$.

We now examine the structure of \mathbf{Q}_p^\times more closely. We know already that $\mathbf{Q}_p^\times \cong \mathbf{Z} \times U$, by Prop. 1.3, and it remains to describe the group of units U . Each element of U is of the form

$$u = a_0 + a_1 p + a_2 p^2 + \dots, \quad 1 \leq a_0 < p, 0 \leq a_i < p \ (i > 0).$$

Consider the natural homomorphism $U \rightarrow \mathbf{Z}/p^n$ obtained by ignoring powers of p

higher than p^{n-1} . The kernel is the set $U_n = 1 + p^n \mathbf{Z}_p$ and we have the chain

$$U = U_0 \supset U_1 \supset U_2 \supset \dots, \quad \bigcap U_n = 1. \quad (11)$$

U_1 is called the group of 1-units ('Einseinheiten'); they are the p -adic units that are congruent to 1 (mod p). We note that the first factor group in the series (11) is isomorphic to the multiplicative group of \mathbf{F}_p while the remaining factors are isomorphic to the additive group of \mathbf{F}_p :

$$U_0/U_1 \cong \mathbf{F}_p^\times, \quad U_n/U_{n+1} \cong \mathbf{F}_p \quad (n > 0).$$

For when $n > 0$, we have $(1 + p^n x)(1 + p^n y) \equiv 1 + p^n(x + y) \pmod{p^{n+1}}$.

Next consider the equation

$$x^{p-1} = 1. \quad (12)$$

It has $p - 1$ roots in \mathbf{F}_p , all distinct, and so each can be lifted uniquely to a solution in \mathbf{Z}_p . Thus (12) has $p - 1$ roots in \mathbf{Z}_p ; these roots form a subgroup L of U , called the *set of multiplicative representatives* of \mathbf{F}_p^\times in \mathbf{Z}_p . We assert that

$$U = L \times U_1. \quad (13)$$

For any $x \in U$ can be written $c(1 + a_1 p + \dots) \in LU_1$, and this representation is unique, because the only element of L with constant term 1 is 1, so that $L \cap U_1 = 1$. This establishes (13).

It remains to find the structure of U_1 ; here we need a homomorphism from the additive to the multiplicative group of \mathbf{Q}_p . In the case of \mathbf{R} such a mapping is provided by the exponential function. Now $\exp x$ can still be defined as a power series over \mathbf{Q}_p , but it will no longer converge for all x , because the coefficients $1/n!$ do not tend to 0 as $n \rightarrow \infty$. Instead we shall use the binomial series. We shall need

LEMMA 3.5 For any $z \in \mathbf{Z}_p$ and $n \in \mathbf{N}$, $\binom{z}{n} \in \mathbf{Z}_p$.

Proof. Given any integer N , we can write $z = z_N + p^N z'$, where $z_N \in \mathbf{Z}$ and $z' \in \mathbf{Z}_p$. Now for any polynomial f , $f(x) - f(y) = (x - y)g(x, y)$, where g is a polynomial in x and y which only depends on f . Taking $x = z$, $y = z_N$, $f(x) = \binom{x}{n}$, we find that

$$\binom{z}{n} - \binom{z_N}{n} = p^N z' g(z, z_N). \quad (14)$$

Here the coefficients of g lie in \mathbf{Q}_p and depend only on n , not on N , so for large enough N the right-hand side of (14) lies in \mathbf{Z}_p . Since clearly $\binom{z_N}{n} \in \mathbf{Z}_p$, we have $\binom{z}{n} \in \mathbf{Z}_p$, as claimed. ■

Now consider the binomial series in \mathbf{Z}_p :

$$(1 + p)^x = \sum \binom{x}{n} p^n. \quad (15)$$

By the lemma $\binom{x}{n} \in \mathbf{Z}_p$, hence the series converges for all $x \in \mathbf{Z}_p$, and as in elementary analysis one proves that

$$(1 + p)^x (1 + p)^y = (1 + p)^{x+y}.$$

So we have a homomorphism $\varphi: \mathbf{Z}_p \rightarrow U_1$ given by $x \mapsto (1 + p)^x$. To find the kernel, we write $x = ap^r + \dots$, where $a \not\equiv 0 \pmod{p}$. Then

$$(1 + p)^x = 1 + xp + \binom{x}{2} p^2 + \dots \equiv 1 + ap^{r+1} \pmod{p^{r+2}},$$

at least for $p \neq 2$. This is impossible if $(1 + p)^x = 1$, hence the mapping φ is injective for $p \neq 2$. When $p = 2$, suppose that $(1 + p)^x = 1$; then $1 = (1 + 2)^{2x} = (1 + 2^3)^x$, and now we can argue as before: if $x = a2^r + \dots$, $a \not\equiv 0 \pmod{2}$, then $(1 + 2^3)^x \equiv 1 + a2^{r+3} \pmod{2^{r+4}}$, which is again a contradiction. Thus we have an injection even for $p = 2$.

The mapping $\varphi: \mathbf{Z}_p \rightarrow U_1$ induces homomorphisms

$$\varphi_n: \mathbf{Z}_p \rightarrow U_1/U_{n+1} \quad (n = 0, 1, \dots),$$

and it is clear from the previous argument that $\ker \varphi_n = p^n \mathbf{Z}_p$ for $p \neq 2$. Hence the induced mapping

$$\bar{\varphi}_n: \mathbf{Z}_p/p^n \mathbf{Z}_p \rightarrow U_1/U_{n+1}$$

is an injection for $p \neq 2$. Here both sides have the same finite order p^n , hence $\bar{\varphi}_n$ is an isomorphism. In particular, this shows $\bar{\varphi}_n$ and with it φ_n to be surjective. It follows from the construction of U_1 that $\varphi: \mathbf{Z}_p \rightarrow U_1$ is surjective, hence it is an isomorphism. This proves the case $p \neq 2$ of

THEOREM 3.6 *For any odd prime p , $\mathbf{Z}_p \cong U_1$, while for $p = 2$, $\mathbf{Z}_2 \cong U_2$ and $U_1 \cong \langle -1 \rangle \times U_2$.*

It only remains to consider the case $p = 2$. The formula (15) still applies and it gives an injection $\mathbf{Z}_2 \rightarrow U_1$, but this is no longer surjective. We note that now U_1/U_3 is not cyclic: $(1 + 2)^2 \equiv 1 \pmod{2^3}$, so U_1/U_3 is isomorphic to the 4-group. Instead of φ we shall use the mapping $\psi: \mathbf{Z}_2 \rightarrow U_2$ given by

$$\psi(x) = (1 + 4)^x = \sum \binom{x}{n} 2^{2n}.$$

Clearly this is injective, and as before we get an isomorphism

$$\psi_n: \mathbf{Z}_2/2^n \mathbf{Z}_2 \rightarrow U_2/U_{n+2}.$$

Hence ψ is an isomorphism, and clearly $U_1 = U_2 \times \langle -1 \rangle$. ■

If we combine the expression for U_1 obtained in this theorem with earlier results, we obtain

COROLLARY 3.7 *The multiplicative structure of \mathbf{Q}_p is given by*

$$\begin{aligned}\mathbf{Q}_p^\times &\cong \mathbf{Z} \times \mathbf{Z}_p \times \mathbf{C}_{p-1} & \text{for } p \neq 2, \\ \mathbf{Q}_2^\times &\cong \mathbf{Z} \times \mathbf{Z}_2 \times \mathbf{C}_2.\end{aligned}\blacksquare$$

The study of functions on \mathbf{Q}_p shows many features quite different from the familiar case of functions on \mathbf{R} . We shall not go into detail, but content ourselves with giving a criterion for the continuity of functions on \mathbf{Z}_p (cf. K. Mahler (1981)).

Let f be any function on \mathbf{Z}_p with values in \mathbf{Q}_p . With f we associate a series of coefficients $\{a_n\}$, defined as

$$a_n = \sum (-1)^k \binom{n}{k} f(n-k). \quad (16)$$

These are just the iterated differences of f at 0; let us define the translation operator T by

$$Tf(x) = f(x+1),$$

and the difference operator D as $T - 1$:

$$Df(x) = (T - 1)f(x) = f(x+1) - f(x).$$

Then (16) is equivalent to the formula

$$a_n = D^n f(0).$$

We remark that the definition of a_n uses only the values of f on \mathbf{N} ; moreover, we regain the $f(n)$ by the formula

$$f(n) = \sum_k \binom{n}{k} a_k. \quad (17)$$

For on substituting the values of a_k we find

$$\sum \binom{n}{k} D^k f(0) = (1 + D)^n f(0) = T^n f(0) = f(n).$$

Now the continuity criterion states that f is continuous iff the a_n tend to 0:

THEOREM 3.8 *A function f from \mathbf{N} to \mathbf{Q}_p is continuous if and only if its coefficients a_n , given by (16) tend to 0 as $n \rightarrow \infty$. Moreover, when this is so, then f has a unique continuous extension f^* to \mathbf{Z}_p , given by*

$$f^*(x) = \sum \binom{x}{n} a_n. \quad (18)$$

Proof. Let us state the condition for f to be continuous; since \mathbf{Z}_p is compact, continuity and uniform continuity mean the same thing on \mathbf{Z}_p :

f is continuous on \mathbf{Z}_p iff, given $s \in \mathbb{N}$, there exists $t \in \mathbb{N}$ such that

$$v(f(x) - f(y)) \geq s \quad \text{for all } x, y \in \mathbf{Z}_p \text{ such that } v(x - y) \geq t, \quad (19)$$

and of course the same condition applies if f is only defined on \mathbf{Z} . In particular, if f is periodic* with period p^t , then $f(x) = f(y)$ whenever $v(x - y) \geq t$, and it follows that such a function is always continuous. We begin by proving the result for periodic functions; thus we show that for a periodic function f , with period p^t say, the coefficients $a_n = D^n f(0)$ tend to 0. The translation operator T , restricted to the space of functions with period $q = p^t$, satisfies $T^q = 1$; hence by the binomial theorem,

$$D^q = (T - 1)^q = T^q + pG - 1 = pG,$$

where G is an operator with integer coefficients. It follows that $D^{qs} = p^s G^s$ and for any $n > qs$,

$$a_n = D^n f(0) = p^s (G^s D^{n-qs} f)(0),$$

because f has period q . Since D, G have integer coefficients, it follows that $v(a_n) \geq s$ for $n \geq p^t s$, and so $a_n \rightarrow 0$, as claimed.

Next we show that any continuous function can be approximated by periodic functions with sufficiently small period. Let f be continuous, so that (19) holds, and fix $s \geq 1$. For each $x \in \mathbf{Z}_p$ there is a unique rational integer $g(x)$ in the range

$$0 \leq g(x) < p^s \quad (20)$$

such that

$$v(f(x) - g(x)) \geq s. \quad (21)$$

Since f is continuous, we can choose t so that (19) holds; if $x, y \in \mathbf{Z}_p$ are such that $v(x - y) \geq t$, then by (19) and the definition of g , we have

$$v(g(x) - g(y)) \geq \min \{v(f(x) - g(x)), v(f(x) - f(y)), v(f(y) - g(y))\} \geq s.$$

Thus $g(x) \equiv g(y) \pmod{p^s}$, and by (20) it follows that $g(x) = g(y)$. On writing $y = x + q$ we have $g(x + q) = g(x)$, so g has period q , and from (21) we see that it approximates f .

If f, g have coefficients a_n, b_n respectively, then by (21) and Lemma 3.5,

$$v(a_n - b_n) = v\left(\sum (-1)^k \binom{n}{k} [f(n-k) - g(n-k)]\right) \geq s.$$

*From another point of view such a function can be described as a *step function*.

Since g has period q , $b_n \rightarrow 0$, so for large enough n , $v(a_n) \geq s$. But s was arbitrary, so this shows that $a_n \rightarrow 0$, as claimed.

Conversely, if $\{a_n\}$ are the coefficients for f and $a_n \rightarrow 0$, then the series (18) for f^* converges for all $x \in \mathbf{Z}_p$, because $\binom{x}{n} \in \mathbf{Z}_p$ by Lemma 3.5, and it is continuous, for, given $s \in \mathbf{N}$, we can choose n_0 such that $v(a_n) \geq s$ for $n > n_0$ and then choose $v(x - y)$ so large that $v\left(\binom{x}{n} - \binom{y}{n}\right) \geq s$ for $n = 1, 2, \dots, n_0$. Comparing (17) and (18), we see that f^* agrees with f on \mathbf{N} . It follows that f is continuous and f^* is its extension to \mathbf{Z}_p , unique because \mathbf{N} is dense in \mathbf{Z}_p . ■

Exercises

- (1) Solve $x^3 = 4$ in \mathbf{Z}_5 , \mathbf{Z}_2 , \mathbf{Z}_3 when possible.
- (2) How can the positive expression (5) be reconciled with the negative expression for $1/2$ in the line above? (Hint. Remember (4).)
- (3) Show that $\exp x$ converges for $v(x) > 1/(p-1)$. Use \exp instead of the binomial function to prove Th. 3.6. (Hint. Show first that $v(n!) = [n/p] + [n/p^2] + \dots$, where $[\xi]$ is the greatest integer $\leq \xi$.)
- (4) Show that a p -adic number $\sum a_i p^i$ is rational iff the a_i from some index onwards are periodic. Describe the set of all p -adic numbers represented by a finite series $\sum a_i p^i$.
- (5) Show that the equation $x^2 + 1 = 0$ is irreducible over \mathbf{Q}_2 . (Hint. Try $x = a + 2b$.) Do the same for $x^2 - 2 = 0$.
- (6) Let p be an odd prime. Show that if α, β are units in \mathbf{Z}_p , then $\alpha x^2 + \beta y^2 = 1$ has a solution in \mathbf{Z}_p . Deduce that two regular quadratic forms f, g of the same rank over \mathbf{Q}_p are equivalent iff $\det f$ and $\det g$ define the same residue class in $\mathbf{Q}_p^\times / \mathbf{Q}_p^{\times 2}$.
- (7) Let $x \in \mathbf{Q}_p^\times$ have the form $x = p^n u$ ($n \in \mathbf{Z}$, $u \in U$). Show that x is a square iff n is even and u is a quadratic residue mod p when p is odd, and $u \equiv 1 \pmod{8}$ when $p = 2$. Hence find $\mathbf{Q}_p^\times / \mathbf{Q}_p^{\times 2}$.
- (8) Let $p \neq 2$ and $v(a_1) = v(a_2) = v(a_3)$. Show that $a_1 x_1^2 + a_2 x_2^2 + a_3 x_3^2$ is isotropic over \mathbf{Q}_p . Deduce that any quadratic form in at least five variables over \mathbf{Q}_p is isotropic. Conclude that \mathbf{Q}_p cannot be ordered.
- (9) Let f be a p -adic function with coefficients a_n . Show that $\sum f(n)x^n = (1+t)\sum a_n t^n$, where $t = x(1-x)^{-1}$. Deduce that the function g defined by $g(n) = (-1)^n a_n$ has coefficients $b_n = (-1)^n f(n)$.

8.4 Integral elements

Let S be any commutative ring and R a subring of S . An element of S is said to be *integral over R* if it satisfies a monic equation with coefficients in R :

$$x^n + a_1x^{n-1} + \cdots + a_n = 0 \quad (a_i \in R). \quad (1)$$

This generalizes the definition of an algebraic integer given in 5.1, which was the special case $R = \mathbf{Z}$. Even in that case we have, besides the obvious instances such as $\sqrt{2}$, also integers like $\frac{1}{2}(-1 + \sqrt{-3})$, which arises as root of $x^3 = 1$.

There are several equivalent ways of expressing the definition, which are often useful:

PROPOSITION 4.1 *Let $R \subseteq S$ be rings. For any $c \in S$ the following conditions are equivalent:*

- (a) *c is integral over R ,*
- (b) *$R[c]$ is finitely generated as R -module,*
- (c) *there is a finitely generated R -submodule M of S with zero annihilator in $R[c]$, such that $cM \subseteq M$.*

When c is a unit in S , all are equivalent to

- (d) $c \in R[c^{-1}]$.

Proof. (a) \Rightarrow (b). Let c satisfy (1); we claim that $R[c]$ is generated as R -module by $1, c, c^2, \dots, c^{n-1}$. For it is clearly generated by all the powers c^v , and for $v \geq n$ we have on multiplying (1) for $x = c$ by c^{v-n} and rearranging:

$$c^v = -(a_1c^{v-1} + \cdots + a_nc^{v-n}).$$

By induction on v this expresses all powers of c in terms of $1, c, \dots, c^{n-1}$. We note that when c is a unit in S , we find in the same way

$$c = -(a_1 + a_2c^{-1} + \cdots + a_nc^{1-n}). \quad (2)$$

Thus when a unit c of S is integral over R , then $c \in R[c^{-1}]$. Conversely, if $c \in R[c^{-1}]$, we have an equation of the form (2) with $a_i \in R$, and this can be rearranged as (1) (with $x = c$); hence c is then integral over R . This shows that (a) \Leftrightarrow (d) when c is a unit in S .

(b) \Rightarrow (c) is clear, since $R[c]$ is a submodule of the required sort, and to prove (c) \Rightarrow (a), let M be generated by u_1, \dots, u_n . Then

$$cu_i = \sum a_{ij}u_j, \quad \text{where } a_{ij} \in R.$$

Hence $\sum_j (c\delta_{ij} - a_{ij})u_j = 0$; if we write $A = (a_{ij})$ and multiply by $\text{adj}(cI - A)$, we obtain

$$\det(cI - A)u_i = 0, \quad i = 1, \dots, n.$$

By hypothesis, the annihilator of all the u_i in $R[c]$ is 0, and it follows that $\det(cI - A) = 0$, which on expansion gives a monic equation for c over R . ■

Given a ring S with a subring R , S is said to be *integral over R* if every element of S is integral over R . We note that if $c_1, \dots, c_n \in S$ are such that c_i is integral over $R[c_1, \dots, c_{i-1}]$ ($i = 1, \dots, n$), then $R[c_1, \dots, c_n]$ is integral over R . For on writing $R_i = R[c_1, \dots, c_i]$, we may, by induction on n , assume that R_{n-1} is finitely generated as R -module, say $R_{n-1} = \sum' Ru_\lambda$. By hypothesis, R_n is finitely generated as R_{n-1} -module, say $R_n = \sum^s R_{n-1}v_\mu$; hence $R_n = \sum Ru_\lambda v_\mu$ and this shows R_n to be generated by the rs elements $u_\lambda v_\mu$ over R . It follows that all the elements of $R_n = R[c_1, \dots, c_n]$ are integral over R . This fact has a useful consequence:

COROLLARY 4.2 *Let S be a ring and R a subring of S . Then the elements of S that are integral over R form a subring containing R .* ■

The subring \bar{R} of all elements of S integral over R is called the *integral closure* of R in S , and R is said to be *integrally closed* in S if $\bar{R} = R$. Generally an integral domain is called ‘integrally closed’ (without qualification) if it is integrally closed in its field of fractions. For example, $\mathbf{Z}[\sqrt{-1}]$ is integrally closed, as is easily seen, but $\mathbf{Z}[\sqrt{-3}]$ is not, for its integral closure contains the element $\frac{1}{2}(-1 + \sqrt{-3})$.

There is a close connexion between the integral closure and general valuation rings; to describe it we shall need an existence lemma for valuation rings, which is also useful elsewhere. On any field K we consider pairs $P = (R, \alpha)$ consisting of a subring R of K and a proper ideal α in R . Given two such pairs $P_i = (R_i, \alpha_i)$ ($i = 1, 2$), we shall say that P_1 dominates P_2 , in symbols $P_1 \geq P_2$, if $R_1 \supseteq R_2$ and $\alpha_1 \supseteq \alpha_2$.

LEMMA 4.3 *Let K be a field, R a subring of K and α a non-zero proper ideal in R . Then there is a subring V with an ideal \mathfrak{p} such that (V, \mathfrak{p}) is maximal among pairs dominating (R, α) . Further, any such maximal pair (V, \mathfrak{p}) consists of a valuation ring $\neq K$ and its maximal ideal.*

Proof. The family of pairs dominating (R, α) is clearly inductive and so by Zorn’s lemma it has a maximal element (V, \mathfrak{p}) . It remains to show that V is a valuation ring in K and \mathfrak{p} its maximal ideal. Let $c \in K$; we must show that $c \in V$ or $c^{-1} \in V$. Assume the contrary; then since $c \notin V$, we have $V[c] \supset V$, and if the ideal \mathfrak{p}' generated by \mathfrak{p} in $V[c]$ is proper, then $(V[c], \mathfrak{p}') > (V, \mathfrak{p})$, which contradicts the maximality. Hence $\mathfrak{p}' = V[c]$, i.e. we have an equation

$$1 = a_0 + a_1c + \cdots + a_mc^m, \quad a_i \in \mathfrak{p}. \quad (3)$$

Similarly, since $c^{-1} \notin V$, we have an equation

$$1 = b_0 + b_1c^{-1} + \cdots + b_nc^{-n}, \quad b_j \in \mathfrak{p}. \quad (4)$$

We may assume that m, n are chosen to be as small as possible, and by symmetry we may take $m \geq n$. Multiplying (4) by c^m we get

$$(1 - b_0)c^m = b_1c^{m-1} + \cdots + b_nc^{m-n}. \quad (5)$$

If we now multiply (3) by $1 - b_0$ and substitute for $(1 - b_0)c^m$ from (5), we obtain an equation of the form (3), but with a lower value of m . This is a contradiction, and it shows that either c or c^{-1} lies in V , i.e. V is a valuation ring in K . Further, \mathfrak{p} is a maximal ideal of V , by the maximality of the pair (V, \mathfrak{p}) . Finally, $V \neq K$, for if $V = K$, then its maximal ideal would be 0 and so could not contain \mathfrak{a} . ■

In this lemma we assumed that $\mathfrak{a} \neq 0$. If $\mathfrak{a} = 0$, we can still apply the lemma, replacing \mathfrak{a} by any non-zero ideal, as long as R is not a field. Even when R is a field, if K is transcendental over R , we can enlarge R to a subring $R[x]$ not a field and proceed as before. However, if R is a field and K is algebraic over R , then any subring between R and K is a field; that case any pair dominating $(R, 0)$ is of the form $(L, 0)$, where L is a field between R and K .

We can now characterize the integral closure of a subring as follows.

THEOREM 4.4 *Let K be a field and A a subring. Then the integral closure \bar{A} of A in K is the intersection of all general valuation rings of K containing A .*

Proof. Let $\{V_\lambda\}$ be the family of all valuation rings in K that contain A , and denote by \mathfrak{p}_λ the maximal ideal of V_λ . If c is integral over A , then either $c = 0$ or its monic equation over A may be written

$$1 + a_1c^{-1} + \cdots + a_nc^{-n} = 0. \quad (6)$$

Suppose that $c \notin V_\lambda$; then $c^{-1} \in \mathfrak{p}_\lambda$ and now (16) shows that $1 \in \mathfrak{p}_\lambda$, a contradiction; hence $c \in \bigcap V_\lambda$. Conversely, if c is not integral over A , then $c \neq 0$ and by Prop. 4.1, $c \notin A[c^{-1}]$, hence (c^{-1}) is a proper ideal in $A[c^{-1}]$. By Lemma 4.3 there exists a valuation ring $V \supseteq A[c^{-1}]$ with maximal ideal containing (c^{-1}) . It follows that $c \notin V$, so $c \notin \bigcap V_\lambda$; therefore we have $\bar{A} = \bigcap V_\lambda$ as claimed. ■

We observe that Th. 4.4 provides another proof that the elements of a field integral over a given subring themselves form a ring. The theorem also shows that a subring of a field K is integrally closed iff it is the intersection of all the valuation rings containing it. Here it is of course important to include all valuations, not merely the principal ones.

We note that a UFD may also be described as the intersection of the valuation rings associated with the various atoms. Hence we have

COROLLARY 4.5 *A unique factorization domain is integrally closed (in its field of fractions).* ■

For some applications it is useful to have a characterization of real-valued valuations. This is given by

THEOREM 4.6 *Let K be a field with a subring V . Then V is the valuation ring of a non-trivial real-valued valuation on K if and only if V is a maximal subring of K which is not a field.*

Proof. Let V be the valuation ring of a non-trivial real-valued valuation v on K , and let \mathfrak{p} be its maximal ideal. Any ring strictly containing V contains an element c such that $v(c) < 0$. Now for any $x \in K$ there exists $n \in \mathbb{N}$ such that $nv(c) < v(x)$, hence $v(xc^{-n}) > 0$, so $a = xc^{-n} \in V$ and therefore $x = c^n a \in V[c]$. Hence $V[c] = K$ and this shows V to be a maximal proper subring of K .

Conversely, let V be a maximal subring of K , not a field. By Lemma 4.3, V is a valuation ring. We complete the proof by showing that its value group Γ is archimedean ordered, for then it can be embedded in \mathbf{R} (cf. 6.6). Let $a, b \in V$, $a^{-1} \notin V$; then $v[a^{-1}] = K$, hence $b^{-1} = ca^{-n}$ for some $c \in V$ and some $n \geq 0$, and so $v(a^n b^{-1}) = v(c) \geq 0$, i.e. $nv(a) \geq v(b)$. Thus Γ is indeed archimedean ordered and we can embed Γ in \mathbf{R} by fixing $\alpha \in \Gamma$, $\alpha > 0$, and defining

$$\varphi(\beta) = \inf \{m/n \mid m\alpha \geq n\beta\}.$$

Hence v is equivalent to a real-valued valuation. ■

Another important consequence of Lemma 4.3 is the extension theorem; to describe it we need yet another way of looking at valuations.

Let K be a field. By a *place* of K in a field k we understand a mapping

$$\varphi: K \rightarrow k \cup \{\infty\},$$

such that $k\varphi^{-1} = V$ is a subring of K and the restriction $\varphi|V$ is a ring homomorphism. Thus with the usual operations on ∞ , we have

$$(x - y)\varphi = x\varphi - y\varphi, \quad 1\varphi = 1, \quad (xy)\varphi = x\varphi \cdot y\varphi,$$

whenever the right-hand side is defined (i.e. different from $0 \cdot \infty$ and $\infty - \infty$). The set $k\varphi^{-1}$ is called the set of elements where φ is *finite*.

For example, a valuation on K , with valuation ring V and maximal ideal \mathfrak{p} , defines a place on the residue class field V/\mathfrak{p} , which is finite on V . Conversely, if φ is a place of K in k and V is the set where φ is finite, then if $x \notin V$, we have $x^{-1} \in V$, for otherwise we should have $1 = x\varphi \cdot x^{-1}\varphi = \infty \cdot \infty = \infty$, a contradiction. Hence V is a valuation ring and $\varphi|V: V \rightarrow V/\mathfrak{p}$ is the residue class homomorphism. It is clear that two places $\varphi_i: K \rightarrow k_i$ ($i = 1, 2$) correspond to the same valuation iff there is an isomorphism $\alpha: k_1 \rightarrow k_2$ such that $\varphi_1 \alpha = \varphi_2$; in this case the places are said to be *equivalent*. The situation is summed up in

PROPOSITION 4.7 *The valuations on a field K correspond to the places of K , with equivalent valuations corresponding to equivalent places.* ■

For the next result we need a special case of localization, which forms the subject of 9.3. We shall describe this briefly here and refer the reader who wants to know more to 9.3.

Let R be a ring and S a subset of R . A ring homomorphism $f: R \rightarrow R'$ is called *S-inverting* if it maps the elements of S to invertible elements in R' . It is easily seen that the set of all elements inverted by a given homomorphism is *multiplicative*, i.e. it contains 1 and is closed under multiplication. If R is an integral domain with field of fractions K , and S is a multiplicative subset of R , then the set

$$R_S = \{x \in K \mid x = a/u, \text{ where } a \in R, u \in S\}$$

is easily seen to be a subring of K containing R ; it is called the *ring of fractions with denominators in S* . In particular, when S is the complement of a prime ideal \mathfrak{p} , then R_S is a *local ring*, i.e. the set of all non-units forms an ideal. In this case one usually writes (somewhat illogically) $R_{\mathfrak{p}}$ to mean R_S ; the risk of confusion is small, since the multiplicative set S does not usually contain 0, whereas \mathfrak{p} always does.

THEOREM 4.8 (Extension theorem) *Let K be a field containing a subring A and let $f: A \rightarrow E$ be a homomorphism of A into a field E . Then f can be extended to a place of K .*

Proof. Write $\mathfrak{p} = \ker f$ and $S = A \setminus \mathfrak{p}$; then f is S -inverting and so can be extended to a homomorphism of the local ring $A_{\mathfrak{p}}$ into E . If $A_{\mathfrak{p}}$ is a field and K is algebraic over it, then we can extend f further to K , by Lemma 3.5.3. Otherwise there exists by Lemma 4.3 (and the remark following it) a valuation ring V with maximal ideal \mathfrak{m} dominating $A_{\mathfrak{p}}$ and its maximal ideal \mathfrak{p}_1 , say. Thus $\mathfrak{m} \cap A_{\mathfrak{p}} \supseteq \mathfrak{p}_1$ and in fact equality holds here, because every element of $A_{\mathfrak{p}}$ not in \mathfrak{p}_1 is a unit. Thus we have

$$\begin{array}{ccccc} & & V/\mathfrak{m} & & \\ & & \uparrow & & \\ A/\mathfrak{p} & \longrightarrow & A_{\mathfrak{p}}/\mathfrak{p}_1 & \longrightarrow & E \end{array}$$

the diagram shown: here $A_{\mathfrak{p}}/\mathfrak{p}_1$ is a subfield of V/\mathfrak{m} , hence the embedding can be extended to one of V/\mathfrak{m} in a field $F \supseteq E$, and this is a place of K in F . ■

A closer analysis shows that the field F can be taken within an algebraic closure of E , but this fact will not be needed here.

Exercises

- (1) Verify that $\mathbf{Z}[\sqrt{-1}]$ is integrally closed.
- (2) Let $R \subseteq S$ be rings. Show that an element c integral over R is a non-zerodivisor iff the constant term in the monic equation of least degree for c is a non-zerodivisor in S .

- (3) Let $R \subset R'$ be integral domains such that R' is integral over R . Show that for every prime ideal \mathfrak{p} of R there is a prime ideal \mathfrak{P} of R' such that $\mathfrak{P} \cap R = \mathfrak{p}$.
- (4) Let K/k be a finitely generated field extension. Show that there exists a valuation trivial on k but non-trivial on K iff K/k is not algebraic.
- (5) In Lemma 4.3 show that if \mathfrak{a} is prime, then \mathfrak{p} can be chosen so that $\mathfrak{p} \cap R = \mathfrak{a}$. (Hint. See the proof of Th. 4.8.)
- (6) Let A be an integral domain, integrally closed in its field of fractions K . Show that for any monic polynomials f, g over K , if $f, g \in A[x]$, then $f, g \in A[x]$. Give an example to show that the condition (on integral closure) cannot be omitted.
- (7) Let A be an integrally closed domain and L any field containing A . If $a_1, \dots, a_n \in L$ are such that for each $i = 1, \dots, n$ there exists $m_i \in \mathbb{N}$ such that $a_i^{m_i} = f_i(a_1, \dots, a_n)$, where f_i is a polynomial of degree $< m_i$ with coefficients in A , then a_1, \dots, a_n are integral over A . Deduce that any extension of A finitely generated as A -module is integral over A .

8.5 Extension of valuations

We now turn to ways of extending a valuation given on a field to a larger field. There are essentially two cases to consider, the algebraic and the transcendental extensions. It is possible to treat both at once, but we shall take these cases separately, since it is then possible to deal with the algebraic case more simply, while the transcendental case can be treated more explicitly.

Let L/K be any field extension, and let v be a valuation on K , with an extension w to L . We shall write V, \mathfrak{p}, Γ for the valuation ring, maximal ideal and value group of v , and W, \mathfrak{P}, Δ for the corresponding entities for w . Clearly we have $\Gamma = v(K) = w(K) \subseteq w(L) = \Delta$, hence Γ is a subgroup of Δ . The index

$$(\Delta : \Gamma) = e \tag{1}$$

is always denoted by e and is called the *ramification index*. Next we have $\mathfrak{P} \cap V = \mathfrak{p}$, hence the natural homomorphism

$$V/\mathfrak{p} \rightarrow W/\mathfrak{P} \tag{2}$$

is an embedding; so denoting the residue class fields by \bar{K}, \bar{L} respectively, we see that \bar{L} is an extension of \bar{K} . The degree

$$[\bar{L} : \bar{K}] = f \tag{3}$$

is called the *residue degree*. From the properties of the index and degree it is clear that both the ramification index and the residue degree are multiplicative under extensions.

For example, let $K = \mathbb{Q}$, $L = \mathbb{Q}(\sqrt{2})$. The 2-adic valuation v_2 has an extension to L for which $e = 2, f = 1$ (2 is ‘ramified’ in L). If p is an odd prime, either 2 is a quadratic residue mod p (i.e. $x^2 \equiv 2 \pmod{p}$ has a solution); then 2 is a square in

Q_p and v_p has two extensions to L , and $e = f = 1$. Or 2 is a quadratic non-residue mod p ; then 2 is not a square in Q_p (p is ‘inert’ in L); now v_p has just one extension to L , $e = 1$, $f = 2$ and the residue field is enlarged from F_p to $F_p(\sqrt{2})$. This illustrates the general situation to be described in Th. 5.7 below.

In comparing the values of v and w the following relation is often useful: Let $u_1, \dots, u_r \in W$ be such that their residues mod \mathfrak{P} , $\bar{u}_1, \dots, \bar{u}_r$, say, are linearly independent over $\bar{K} = V/\mathfrak{p}$. Then we claim that for any $\alpha_1, \dots, \alpha_r \in K$, we have

$$w(\alpha_1 u_1 + \dots + \alpha_r u_r) = \min \{v(\alpha_1), \dots, v(\alpha_r)\}. \quad (4)$$

To prove this assertion we may assume, after appropriate renumbering, that $v(\alpha_1) = \dots = v(\alpha_k) < v(\alpha_j)$ ($j > k$), so that $\alpha_i/\alpha_1 \in V$ for all i . Since $\bar{u}_i \neq 0$, it follows that $w(u_i) = 0$, so that left-hand side of (4) has a value $\geq v(\alpha_1)$. If this value is $> v(\alpha_1)$, then after dividing by α_1 and reducing mod \mathfrak{P} we get

$$\bar{u}_1 + (\overline{\alpha_2/\alpha_1}) \bar{u}_2 + \dots + (\overline{\alpha_r/\alpha_1}) \bar{u}_r = 0.$$

This contradicts the linear independence of the \bar{u}_i over \bar{K} and it establishes equality in (4).

An important relation between e and f is given by

THEOREM 5.1 *Let L/K be a finite field extension of degree n , and let v be a valuation on K with an extension w to L . Then the ramification index e and residue degree f satisfy the relation*

$$ef \leq n. \quad (5)$$

If v is real-valued, or principal, or trivial, then so is w .

If moreover K is complete and v is principal, then equality holds in (5).

Proof. Denote the valuation rings of K , L by V , W respectively, so that $W \cap K = V$, and denote by \mathfrak{p} , \mathfrak{P} the corresponding maximal ideals. Let $u_1, \dots, u_r \in W$ be such that their residues mod \mathfrak{P} are linearly independent over V/\mathfrak{p} , and take $\pi_1, \dots, \pi_s \in L$ such that the $w(\pi_i)$ are incongruent mod Γ , the value group of v . We assert that the rs elements $u_i \pi_j$ are linearly independent over K . For if there is a relation

$$\sum \alpha_{ij} u_i \pi_j = 0, \quad \text{where } \alpha_{ij} \in K, \quad (6)$$

let us write $a_j = \sum_i u_i \alpha_{ij}$, so that (6) reads $\sum a_j \pi_j = 0$. If the a_j are not all zero, it follows that $w(a_h \pi_h) = w(a_k \pi_k)$ for some h, k , $1 \leq h < k \leq s$. This means that $w(\pi_h/\pi_k) = w(\pi_h) - w(\pi_k) = w(a_k) - w(a_h)$. Here the right-hand side is in Γ , by (4), but the left-hand side is not, which is a contradiction. Hence all the a_i vanish, and so, again by (4), the α_{ij} must also vanish, which shows the $u_i \pi_j$ to be linearly independent. It follows that $rs \leq n$; in particular, e and f must be finite and taking $s = e$, $r = f$, we obtain (5).

From the definition of e it follows that $e\Delta \subseteq \Gamma$, where Δ is the value group of w ;

thus if Γ is embedded in \mathbf{R} , so is Δ , and if $\Gamma \subseteq \mathbf{Z}$, then by renormalizing we can ensure that $\Gamma \subseteq e\mathbf{Z}$; it follows that $e\Delta \subseteq e\mathbf{Z}$, i.e. $\Delta \subseteq \mathbf{Z}$. If v is trivial, then $\Gamma = 0$, hence $e\Delta = 0$ and so w is also trivial.

Now assume that v is principal and K is complete. Then so is L , as a finite-dimensional K -space, by Prop. 2.7. Let π, Π be uniformizers for v, w respectively; then every element of L can be written as

$$x = \sum \alpha_i \Pi^i,$$

where α_i belongs to a transversal of the residue class field \bar{L} in W . Instead of the powers Π^i we can also use $\pi^i, \Pi\pi^i, \dots, \Pi^{e-1}\pi^i$, and instead of a transversal of \bar{L} in W we can take elements u_1, \dots, u_f of W such that $\bar{u}_1, \dots, \bar{u}_f$ form a basis of \bar{L} over \bar{K} and use $\sum \alpha_i u_i$, where the α_i run over a transversal of \bar{K} in V . Then every element of L may be written

$$x = \sum_{i=-N}^{\infty} \sum_{\mu\nu} a_{\mu\nu i} u_{\mu} \Pi^{\nu} \pi^i, \quad \text{where } a_{\mu\nu i} \in K.$$

This shows that L is spanned by the ef elements $u_{\mu} \Pi^{\nu}$ over K . Hence $[L:K] \leq ef$ and it follows that we have equality in (5). ■

We begin by looking at transcendental extensions; here we need not postulate completeness, nor even assume that the valuation is principal.

PROPOSITION 5.2 *Let K be a field with a valuation v and residue class field \bar{K} . If $K(t)$ is a purely transcendental extension of K , then there is just one extension of v to $K(t)$ such that t remains transcendental over \bar{K} . This valuation on $K(t)$ has the same value group as v and has the residue class field $\bar{K}(t)$.*

The valuation on $K(t)$ determined in this way is called the *Gaussian extension* of v to $K(t)$.

Proof. If there is such an extension w of v , then on $K[t]$ it is given by

$$w(a_0 + a_1 t + \dots + a_n t^n) = \min \{v(a_0), \dots, v(a_n)\}, \quad (7)$$

by (4). This is already enough to determine w on the field of fractions $K(t)$. Thus there can be at most one such extension, and there is one, since w given by (7) clearly satisfies all the conditions. It is clear from (7) that v and w have the same value group. Further, the residue class field clearly contains $\bar{K}(t)$, and for any element f/g such that $w(f/g) \geq 0$ we can find $u \in K(t)$, $\bar{u} \in \bar{K}(t)$ such that $w(f/g - u) > 0$; this is most easily seen by writing f/g as a Laurent series in t . This shows $\bar{K}(t)$ to be the residue class field. ■

Next we prove that extensions always exist in the algebraic case.

PROPOSITION 5.3 *Let L/K be a field extension of finite degree and let v be a valuation on K . Then v has an extension to L ; moreover, any such extension is real-valued, principal or trivial whenever v is.*

If v is real-valued and K is complete, then there is a unique extension w of v to L , given in terms of the norm by

$$w(\gamma) = [L:K]^{-1}v(N_{L/K}(\gamma)) \quad \text{for } \gamma \in L. \quad (8)$$

Proof. We already know from Th. 5.1 that if v is real-valued, principal or trivial, then so is any extension to L . It only remains to prove the existence.

Let V be the valuation ring of v on K . Then $V \subseteq L$ and by the extension theorem (Th. 4.8) the natural homomorphism $V \rightarrow V/\mathfrak{p} = K$ can be extended to a place of L in an extension field of \bar{K} . This gives the required valuation on L . When K is complete (for a real-valued valuation v), then the extension w to L is unique, by Th. 2.9. Now take any $\gamma \in L$, suppose that f is its minimal polynomial over K , of degree n say, and let E be a splitting field of f over L ; w has a unique extension to E , still denoted by w . Over E we can write $f = \prod(x - \gamma_i)$, where $\gamma_1 = \gamma$, say. Since f is irreducible over K , we have $K(\gamma_i) \cong K(\gamma)$ for $i = 1, \dots, n$, hence by uniqueness, $w(\gamma_i) = w(\gamma)$ and so $v(N_{K(\gamma)/K}(\gamma)) = \sum w(\gamma_i) = nw(\gamma)$. If we put $[L:K] = rn$, then $v(N_{L/K}(\gamma)) = rv(N_{K(\gamma)/K}(\gamma)) = rnw(\gamma)$ and on dividing by rn this yields (8). ■

We observe that this result has an analogue in the archimedean case: then K must be \mathbf{R} or \mathbf{C} , by Ostrowski's second theorem, and $[L:K]$ is 1 or 2. The only non-trivial case is that where $K = \mathbf{R}$, $L = \mathbf{C}$ and the absolute value on \mathbf{C} is given by

$$\|x\| = |x\bar{x}|^{1/2},$$

which is the multiplicative analogue of (8).

The following necessary condition for irreducibility over a complete field is an easy consequence of Prop. 5.3.

COROLLARY 5.4 *Let K be a field, complete under a real-valued valuation v . If $f = a_0x^n + a_1x^{n-1} + \dots + a_n$ is an irreducible polynomial over K , then*

$$v(a_i) \geq \min \{v(a_0), v(a_n)\}. \quad (9)$$

Proof. Suppose first that $v(a_0) \leq v(a_n)$, so that $v(a_n/a_0) \geq 0$, and consider the monic polynomial $a_0^{-1}f$. If its zeros in a splitting field E are $\alpha_1, \dots, \alpha_n$, and w denotes the unique extension of v to E , then

$$w(\alpha_i) = \frac{1}{n}v(a_n/a_0) \geq 0.$$

Hence the elementary symmetric functions of the α_i are integers, and so

$v(a_i/a_0) \geq 0$, i.e. $v(a_i) \geq v(a_0)$, and this proves (9) in this case. When $v(a_0) > v(a_n)$, we can make a reduction to the previous case by the substitution $y = 1/x$, hence (9) holds generally. ■

For complete fields there is a useful method of lifting factorizations over the residue class field, known as Hensel's lemma. It is usually proved by a somewhat lengthy verification; but it can be obtained more directly from Th. 4.8 and Th. 2.9 (asserting the existence and uniqueness of extensions). We single out the essential step as a lemma.

We remark that a polynomial f over a valuation ring V is primitive (i.e. its coefficients have no common factor) precisely when its image in the residue class field V/\mathfrak{p} is non-zero.

LEMMA 5.5 *Let K be a complete field under a real-valued valuation v , with valuation ring V , maximal ideal \mathfrak{p} and residue class field $V/\mathfrak{p} = \bar{K}$. Write $a \mapsto \bar{a}$ for the residue class map $V \rightarrow V/\mathfrak{p}$. If*

$$f = a_0 x^n + a_1 x^{n-1} + \cdots + a_n \quad (10)$$

is any irreducible primitive polynomial over V , then one of the following three cases will arise:

- (i) a_0, a_n are both units and \bar{f} is a power of an irreducible polynomial,
- (ii) a_0 is a unit and $\bar{f} = \bar{a}_0 x^n$,
- (iii) a_n is a unit and $\bar{f} = \bar{a}_n$.

Proof. Suppose first that a_0 is a unit. Let E be a splitting field of f over K and w the unique extension of v to E . We write W for the valuation ring of w in E and \mathfrak{P} for its maximal ideal, so that its residue class field is $W/\mathfrak{P} = \bar{E}$. If σ is an automorphism of E over K and $\gamma \in E$, then $w(\gamma^\sigma) = w(\gamma)$ because γ^σ and γ are conjugate over K . This shows that σ maps W into itself, and likewise \mathfrak{P} , so it induces an automorphism $\bar{\sigma}$ of $W/\mathfrak{P} = \bar{E}$ over \bar{K} . Now f is irreducible over V , hence it is also irreducible over K , by inertia (Th. 3.7.2), and if its zeros are $\alpha_1, \dots, \alpha_n$, then

$$w(\alpha_i) = \frac{1}{n} v(a_n) \geq 0$$

by Prop. 5.3. This shows that $\alpha_i \in W$ and we can therefore factorize the image \bar{f} of f over \bar{E} as $\bar{f} = \bar{a}_0 \prod (x - \bar{\alpha}_i)$. For any i, j there is an automorphism σ of E/K such that $\alpha_i^\sigma = \alpha_j$, hence $\bar{\alpha}_i^\sigma = \bar{\alpha}_j$, and so the zeros of \bar{f} are permuted transitively by the automorphisms of \bar{E}/\bar{K} . It follows that \bar{f} is a power of an irreducible polynomial. If a_n is a unit, we are in case (i), while for $a_n \in \mathfrak{p}$, \bar{f} has the factor x and so it must be a power of x , and we have (ii). Now suppose that a_0 is a non-unit. Since f is primitive, some a_i is a unit, hence a_n is a unit by (9). Consider $g(y) = y^n f(y^{-1})$; this

is again irreducible and case (ii) applies because the highest coefficient is a unit, while the constant term is not, so $\bar{g} = \bar{a}_n y^n$; therefore $\bar{f} = \bar{a}_n$. ■

THEOREM 5.6 (Hensel's lemma) *Let K be a complete field for a real-valued valuation v , with valuation ring V , maximal ideal \mathfrak{p} and residue class field $V/\mathfrak{p} = K$. Given a primitive polynomial f over V , suppose that $g_0, h_0 \in V[x]$ are such that*

$$\bar{f} = \bar{g}_0 \bar{h}_0, \quad (11)$$

and that \bar{g}_0, \bar{h}_0 are relatively prime and g_0 is monic; then there exist unique polynomials $g, h \in V[x]$ such that g is monic and g, h satisfy

$$f = gh, \quad g = \bar{g}_0, \quad h = \bar{h}_0. \quad (12)$$

Moreover $\deg g = \deg g_0$.

Proof. We factorize f over $V[x]$ as $f = \prod_1^m p_i^{r_i}$, where the p_i are distinct and irreducible over V , hence by inertia also irreducible over K . Since f is primitive, each p_i is primitive, and by Lemma 5.5, either $\bar{p}_i = \pi_i^{s_i}$ for some irreducible polynomial π_i over \bar{K} and $s_i \geq 1$ (cases (i), (ii)), or $\bar{p}_i = \pi_i \in \bar{K}$ (case (iii)).

Suppose first that the highest coefficient of f is a unit. On dividing by it we may take f monic, and all the p_i (and hence the π_i) may also be taken monic. We have $\bar{g}_0 \bar{h}_0 = \prod_1^m \pi_i^{r_i s_i}$ and since \bar{g}_0, \bar{h}_0 are relatively prime we can number the p_i so that $\bar{g}_0 = \prod_1^t \pi_i^{r_i s_i}$, $\bar{h}_0 = \prod_{t+1}^m \pi_i^{r_i s_i}$; here the π_i need not be distinct, but if $\pi_i = \pi_j$ then $i, j \leq t$ or $i, j > t$. Hence there is just one way to satisfy (12), namely by putting $g = \prod_1^t p_i^{r_i}$, $h = \prod_{t+1}^m p_i^{r_i}$.

If the highest coefficient of f is a non-unit, then case (iii) will occur, i.e. $\bar{p}_i \in \bar{K}$ for some i . Now we can number the p 's so that p_1, \dots, p_u are monic, while p_{u+1}, \dots, p_m have a non-unit as highest coefficient, and of course $u < m$. We can further renumber the p_1, \dots, p_u so that $g_0 = \prod_1^t \pi_i^{r_i s_i}$, $h_0 = \prod_{t+1}^m \pi_i^{r_i s_i}$, where $t \leq u$. Now there are several ways of satisfying (12), by taking one or more of the factors p_{u+1}, \dots, p_m to g , but if g is to be monic, we must have $g = \prod_1^t p_i^{r_i}$, $h = \prod_{t+1}^m p_i^{r_i}$. This then is the unique solution; since $\bar{g} = \bar{g}_0$ and both g, g_0 are monic, they have the same degree. ■

We remark that in the proof of Hensel's lemma we assumed the field K to be complete. This is a 'topological' condition, involving limiting processes; thus the completion of a field generally has a higher cardinal than the field itself. Sometimes it is preferable, instead of completeness, to assume the conclusion of Hensel's lemma. A field with this property is called *henselian*, and every field with a real-valued valuation has a least henselian extension, its 'henselization'. This is obtained by a purely algebraic process, analogous to forming the algebraic closure. We also note that a field K is henselian iff its valuation has a unique extension to any finite field extension (cf. Endler (1972) and Ex. (9)).

We next consider algebraic extensions of incomplete fields. From 5.7 we recall

that for two fields E, F over a common subfield k , such that F/k is separable of degree n , we have

$$E \otimes_k F \cong K_1 \times \dots \times K_r, \quad (13)$$

where K_1, \dots, K_r are the different composites of E and F ; each is a field containing a copy of E and of F . Let $\alpha \in F$ and write f, g_i for the characteristic polynomial of α over k , and of its image in K_i over E respectively. Then $f = g_1 \dots g_r$; for by definition $f(x) = \det(xI - A)$, where $A = (a_{ij})$ and $\alpha v_i = \sum a_{ij}v_j$ for a basis v_1, \dots, v_n of F/k . Clearly f is also the characteristic polynomial of α as an element of $E \otimes F$ over E , because the v 's will be a basis for $E \otimes F$ over E . We now change the basis in $E \otimes F$ by choosing a basis adapted to the decomposition on the right of (13). Each K_i is mapped into itself by α , hence we have $f = g_1 \dots g_r$. In particular, comparing last coefficients and second coefficients in this equation we find

$$N_{F/K}(\alpha) = \prod N_{K_i/E}(\alpha), \quad (14)$$

$$T_{F/K}(\alpha) = \sum T_{K_i/E}(\alpha). \quad (15)$$

With these preparations we can describe the relation between the ramification indices and residue degrees which takes the place of (5) in the incomplete case:

THEOREM 5.7 *Let k be a field with a principal valuation v , and let K/k be a separable extension of degree n . Then there are r extensions w_1, \dots, w_r of v to K , where $1 \leq r \leq n$, and if w_i has ramification index e_i and residue degree f_i , then*

$$\sum e_i f_i = n. \quad (16)$$

Moreover, if k, K have completions \tilde{k}, \tilde{K} under v, w_i respectively, then

$$\tilde{k} \otimes_k K \cong K_1 \times \dots \times K_r. \quad (17)$$

Proof. Since we are dealing with principal valuations whose restrictions to k are all equal, two extensions of v to K are equivalent iff they are actually equal.

If \tilde{K} is a completion of K with respect to a valuation w extending v , then \tilde{k} is isomorphic to a subfield to \tilde{K} . We shall identify it with its image; then $\tilde{k}K$ is a dense subset of \tilde{K} , but since K/k is finite, it is a complete \tilde{k} -space by Prop. 2.7, so $\tilde{K} = \tilde{k}K$. Conversely, given a composite $\hat{K} = \tilde{k}K$, there is by Prop. 5.3 a unique valuation \hat{w} on \hat{K} extending the valuation v defined on \tilde{k} , and \hat{K} is complete by Prop. 2.7. Since k is dense in \tilde{k} , K is dense in \hat{K} , so \hat{K} is the completion of K with respect to the restriction of \hat{w} to K . Thus we have a bijection between the valuations of K extending v and the composites of \tilde{k} and K .

By (13) we have a decomposition of the form (17), where the K_i are the composites of \tilde{k} and K , and hence are the completions of K with respect to the different valuations extending v .

Finally let $[K_i : \tilde{k}] = n_i$. For a principal valuation the residue class field and value group are not affected by passing to completions, hence by Th. 5.1, $n_i = e_i f_i$ and by (17), $n = [K : k] = [\tilde{k} \otimes K : \tilde{k}] = \sum n_i = \sum e_i f_i$. ■

We again observe that this result has an analogue for archimedean absolute values. There $\tilde{k} = \mathbf{R}$ or \mathbf{C} and if e.g. $\mathbf{R} \otimes K = K_1 \times \cdots \times K_r$, we then have $[K_i : \mathbf{R}] = 1$ or 2 according as K_i is real or complex.

For Galois extensions the statement of Th. 5.7 can be simplified:

COROLLARY 5.8 *If in Th. 5.7 K/k is Galois, then the automorphisms of $\tilde{k} \otimes K$ induced by $\text{Gal}(K/k)$ permute the K_i transitively. Hence $e_i = e, f_i = f$ independently of i , and*

$$[K:k] = efr,$$

where r is the number of extensions of v to K .

This will follow if we can show that for some $\sigma \in \text{Gal}(K/k)$, $w_1 = \sigma w_2$, i.e. $w_1(x) = w_2(x^\sigma)$ for all $x \in K$. To this end consider the valuations of K defined by

$$x \mapsto w_1(x^\sigma), \quad x \mapsto w_2(x^\sigma),$$

as σ ranges over $\text{Gal}(K/k) = G$. If these two sets of valuations are disjoint, then by the approximation theorem there exists $a \in K$ such that $w_1(a^\sigma) > 0, w_2(a^\sigma - 1) > 0$ for all $\sigma \in G$. Hence $w_2(a^\sigma) = 0$ and $v(N(a)) = \sum w_2(a^\sigma) = 0$, but also $v(N(a)) = \sum w_1(a^\sigma) > 0$, which is a contradiction. Hence the sets are not disjoint, say $w_1(x^\sigma) = w_2(x^\tau)$ for some $\sigma, \tau \in G$, and so $w_1 = \sigma^{-1} \cdot \tau w_2$ as desired. This shows that the K_i are permuted transitively and the rest is clear. ■

We remark that the factors r, f, e correspond to the three kinds of extension: decomposition, residue class field extension and ramification respectively (cf. Vol. 3).

Exercises

- (1) Let K be a field with a principal valuation. Show that the residue class field has the same characteristic as K iff the valuation is trivial on the prime subfield of K .
- (2) Prove the first part of Prop. 5.3 for real-valued valuations by finding a maximal proper subring containing the valuation ring of v and using Th. 4.6.
- (3) Use Prop. 5.2 and Th. 4.4 to show that if a domain R is integrally closed, then so is the polynomial ring $R[x]$.
- (4) Let K be a valued field with residue class field k of characteristic 0. Show that K contains a subfield k_0 isomorphic to k and mapped to k by the residue class mapping. (Hint. Use Lemma 3.2.)
- (5) Let K be a valued field with valuation v and consider the rational function field $K(t)$. If x is another generator, say $x = (at + b)(ct + d)^{-1}$, find the conditions under which x and t define the same Gaussian extension of v .

- (6) Let K be a field with a principal valuation v and let $K(t)$ be the rational function field. With a positive real number λ define for $f = a_0 + a_1t + \dots + a_nt^n$, a function $w(f) = \min_i\{v(a_i) + i\lambda\}$. Show that w is a valuation on $K[t]$ which can be extended to $K(t)$. Determine its value group and residue class field.
- (7) Let E/K be a field extension of degree n and suppose that E is complete under a principal valuation v . Denote the valuation rings in K, E by V, W and their residue class fields by \bar{K}, \bar{E} . Show that if \bar{E}/\bar{K} is a separable extension then W is a free V -module with basis $1, \alpha, \dots, \alpha^{n-1}$ for a suitable $\alpha \in W$. (Hint. Use Th. 5.1 and the fact that a separable extension is simple.)
- (8) Use Hensel's lemma to find a root of $x^5 = 2$ in \mathbf{Q}_7 .
- (9) Let K be a field with a real-valued valuation v such that v has a unique extension to any finite extension of K . Show that formula (8) of Prop. 5.3 holds, and that Cor. 5.4, Lemma 5.5 and Th. 5.6 all still hold. (This exercise suggests that we can obtain the henselization of K by taking each irreducible polynomial f over K , primitive and with coefficients in V , say, and if \bar{f} can be factorized into relatively prime factors over \bar{K} , say $\bar{f} = \bar{g}\bar{h}$, adjoining enough elements from the algebraic closure of K to K to enable us to lift \bar{g}, \bar{h} to K ; cf. Endler (1972).)
- (10) Let K be a field with a real-valued valuation v . Show that v has a unique extension to any purely inseparable extension of K .

Further exercises on Chapter 8

- (1) Let Γ be any totally ordered abelian group. Verify that the group algebra $k\Gamma$ (for any field k) is an integral domain. (If Γ is written additively, it is convenient to write the elements of $k\Gamma$ as formal ‘polynomials’ $\sum c_\alpha x^\alpha$, where $c_\alpha \in k$, $\alpha \in \Gamma$, with rule $x^\alpha x^\beta = x^{\alpha+\beta}$.) Show that the field of fractions K of $k\Gamma$ has a valuation with value group Γ . Determine the valuation ring and maximal ideal of this valuation. Examine the particular case $\Gamma = \mathbf{Z}$.
- (2) Show that Th. 2.3 still holds for non-commutative rings.
- (3) Show that a function f from \mathbf{Z}_p to \mathbf{Q}_p takes integer values on \mathbf{Z} iff all its coefficients $D^n f(0)$ are integers.
- (4) Show that \mathbf{Q}_p and \mathbf{Q}_q for distinct primes p, q are not isomorphic. (Hint. Look for solutions of $x^2 = p$.)
- (5) Let E, F, G be fields, let α be a place of E in F and β a place of F in G . Show that $\alpha\beta$ (suitably defined) is a place of E in G .
- (6) Let $k \subset K$ be fields and V a minimal valuation ring in K containing k . Using Ex. (5), show that the corresponding place is algebraic over k . Use Zorn's lemma to show that such minimal valuation rings always exist. Deduce that if A is a subring of a field K , then any

homomorphism of A into an algebraically closed field E can be extended to a place of K in E .

- (7) Let A be an integral domain and K its field of fractions. We call A *completely integrally closed* in K if for any $x \in K$ for which there exists $d \neq 0$ in K such that $dx^n \in A$ for all $n > 0$, it is true that $x \in A$. Verify that a completely integrally closed ring is integrally closed, and show that the intersection of any set of principal valuation rings is completely integrally closed.
- (8) Show that a valuation ring is completely integrally closed iff its value group is archimedean ordered (cf. Ex. (7) and **6.6**).
- (9) A finite extension E/K of complete valued fields is said to be *totally ramified* if $[E:K] = e$ is the ramification index. Given an extension E/K with a principal valuation v and with uniformizers p, π in K, E respectively, show that E/K is totally ramified iff v satisfies an equation $x^e + a_1px^{e-1} + \cdots + a_{e-1}px + a_ep = 0$, where $v(a_i) \geq 0$ for $i = 1, \dots, e-1$, $v(a_e) = 0$, and $E = K(\pi)$. (Such an equation is called an *Eisenstein equation*, and the corresponding polynomial is an *Eisenstein polynomial*.)
- (10) (a) Show that any polynomial over \mathbb{Q} close (in the usual absolute value) to a polynomial with n simple real zeros itself has n simple real zeros. (b) Show that a polynomial over \mathbb{Q} close (in the p -adic valuation) to an Eisenstein polynomial for the prime p is itself Eisenstein and hence irreducible. (c) For any integers $0 \leq n \leq m$ construct an irreducible polynomial over \mathbb{Q} of degree m with exactly n real zeros.
- (11) (Krasner's lemma) Let L be a complete field under a non-archimedean absolute value and K a complete subfield. Given $\alpha, \beta \in L$, where α is algebraic over K and β is separable algebraic over $K(\alpha)$, show that if α is closer to β than are any of the conjugates of β over K , then β is fixed under all automorphisms of $K(\alpha, \beta)/K(\alpha)$ and deduce that $\beta \in K(\alpha)$.
- (12) Let K be a complete valued field. Show that if f is monic, irreducible and separable over K , then any polynomial g sufficiently close to f is also irreducible, and to any zero α of f there corresponds a zero β of g such that $K(\alpha) = K(\beta)$.
- (13) Let K be an algebraically closed field with a real-valued valuation. Show that its completion \tilde{K} is again algebraically closed. (*Hint.* First show that the zeros of a polynomial are continuous functions of the coefficients. Now any polynomial f over \tilde{K} may be approximated by polynomials f_v over K , and if f is separable of degree d , then the zeros of these polynomials f_v can be arranged in d Cauchy sequences.)
- (14) Let K be a complete valued field and F its algebraic closure. Show that the valuation has a unique extension to F , which is real-valued if the original valuation was so, and in that case the completion of F is again algebraically closed.
- (15) In the algebraic closure F of \mathbb{Q}_p take any sequence (a_n) such that a_r is a power of a_s for $r \leq s$ (e.g. a_r could be taken to be a primitive $(r!)$ th root of 1); denote the degree of a_r over \mathbb{Q}_p

by n_r . Show that for any $n < n_r$, and any $h \in \mathbb{N}$ there exists $k > h$ such that a_r satisfies no congruence of degree $n \pmod{p^k}$. Deduce the existence of a sequence of integers $k_1 < k_2 < \dots$ such that a_r satisfies no congruence of degree $< n_r \pmod{p^{k_r}}$, and hence show that the series $\sum a_r p^{k_r}$ converges to an element transcendental over \mathbb{Q}_p . Deduce that F is not complete (its completion is algebraically closed, by Ex. (13)).

9

Commutative rings

Commutative ring theory has its origins in number theory and algebraic geometry in the 19th century. Today it is of particular importance in algebraic geometry, and there has been an interesting interaction of algebraic geometry and number theory, using the methods of commutative algebra. Here we can do no more than describe the basic techniques and take the first steps in the subject. Throughout this chapter all rings will be commutative, unless otherwise stated.

9.1 Operations on ideals

Historically the first ring to be studied was the ring \mathbf{Z} of integers; the term ‘ring’ was first used by Hilbert (1897) in his ‘Zahlbericht’ for a ring of algebraic integers (Zahlring). In \mathbf{Z} every ideal is principal; in fact *ideals* were first introduced (by Kummer) as ‘ideal numbers’ in rings of algebraic integers which lacked unique factorization. In \mathbf{Z} we can from any two numbers a, b form their highest common factor (HCF; also greatest common divisor, GCD) (a, b) , their product ab and their least common multiple (LCM) $[a, b]$. These operations correspond to operations on ideals in any ring.

We shall denote ideals in a ring R by small gothic letters $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \dots$. If \mathfrak{a} is an ideal in a commutative ring R , generated by elements a_1, \dots, a_n , we write $\mathfrak{a} = (a_1, \dots, a_n)$; the elements of \mathfrak{a} are all the linear combinations $\sum r_i a_i$ ($r_i \in R$). In particular, when $n = 1$, \mathfrak{a} is principal; thus (a) consists of all elements ra ($r \in R$).

Given ideals $\mathfrak{a}, \mathfrak{b}$ in R , we define their *sum* as

$$\mathfrak{a} + \mathfrak{b} = \{x + y \mid x \in \mathfrak{a}, y \in \mathfrak{b}\}.$$

It is easily seen that this is an ideal, the least ideal containing \mathfrak{a} and \mathfrak{b} . If $\mathfrak{a}, \mathfrak{b}$ are principal, say $\mathfrak{a} = (a)$, $\mathfrak{b} = (b)$, then their sum, if principal, has the form (d) , where d is an HCF of a and b . For example in \mathbf{Z} , $(36) + (10) = (2)$; on the other hand in $\mathbf{Z}[x]$, $(x) + (2) = (x, 2)$ and this cannot be simplified.

Similarly the product is defined as

$$\mathfrak{a}\mathfrak{b} = \{ \sum x_i y_i \mid x_i \in \mathfrak{a}, y_i \in \mathfrak{b} \}.$$

This operation is again associative and commutative, so that ideals form a commutative monoid under multiplication, with $(1) = R$ as neutral element. A

third operation is the intersection $\mathfrak{a} \cap \mathfrak{b}$, which corresponds to the LCM when all ideals are principal.

Let us note the form these definitions take for finitely generated ideals. If $\mathfrak{a} = (a_1, \dots, a_r)$, $\mathfrak{b} = (b_1, \dots, b_s)$, then $\mathfrak{a} + \mathfrak{b} = (a_1, \dots, a_r, b_1, \dots, b_s)$, while $\mathfrak{ab} = (a_1 b_1, a_1 b_2, \dots, a_1 b_s, a_2 b_1, \dots, a_r b_s)$. This shows in particular that the sum and product of finitely generated ideals are again finitely generated. By contrast there is no simple expression for $\mathfrak{a} \cap \mathfrak{b}$, and it need not be finitely generated, even when both \mathfrak{a} and \mathfrak{b} are. But we note that there is a short exact sequence relating $\mathfrak{a} \cap \mathfrak{b}$ to $\mathfrak{a} + \mathfrak{b}$:

$$0 \rightarrow \mathfrak{a} \cap \mathfrak{b} \xrightarrow{\lambda} \mathfrak{a} \oplus \mathfrak{b} \xrightarrow{\mu} \mathfrak{a} + \mathfrak{b} \rightarrow 0. \quad (1)$$

Here μ is defined as $(x, y)\mu = x - y$ ($x \in \mathfrak{a}$, $y \in \mathfrak{b}$) and $\ker \mu \cong \mathfrak{a} \cap \mathfrak{b}$; this is easily verified and may be left to the reader.

There is also a form of division for ideals. Given ideals $\mathfrak{a}, \mathfrak{b}$, we define

$$\mathfrak{a} : \mathfrak{b} = \{x \in R \mid xb \subseteq \mathfrak{a}\}.$$

The reader will have no difficulty in verifying that this is an ideal. This is so even if \mathfrak{b} is an arbitrary subset of R , not necessarily an ideal. Thus for any subset S of R we have

$$\mathfrak{a} : S = \bigcap \{\mathfrak{a} : (x) \mid x \in S\}.$$

The addition, multiplication and intersection of ideals are associative and commutative and satisfy the distributive law: $\mathfrak{a}(\mathfrak{b} + \mathfrak{c}) = \mathfrak{ab} + \mathfrak{ac}$, as is easily checked. Division is related to the other operations by the formulae:

$$(\mathfrak{a} : \mathfrak{b})\mathfrak{b} \subseteq \mathfrak{a} \subseteq \mathfrak{a} : \mathfrak{b}, \quad (2)$$

$$(\bigcap \mathfrak{a}_i) : \mathfrak{b} = \bigcap (\mathfrak{a}_i : \mathfrak{b}), \quad \mathfrak{a} : (\sum \mathfrak{b}_i) = \bigcap (\mathfrak{a} : \mathfrak{b}_i), \quad (3)$$

$$(\mathfrak{a} : \mathfrak{b}) : \mathfrak{c} = \mathfrak{a} : \mathfrak{bc}. \quad (4)$$

For example, to prove (4), we have for any $x \in R$, $x \in (\mathfrak{a} : \mathfrak{b}) : \mathfrak{c} \Leftrightarrow xc \in \mathfrak{a} : \mathfrak{b} \Leftrightarrow xbc \in \mathfrak{a} \Leftrightarrow x \in \mathfrak{a} : \mathfrak{bc}$. The other rules are established similarly.

Exercises

- (1) Prove the formulae (2)–(4).
- (2) Show that $(\mathfrak{a} \cap \mathfrak{b})(\mathfrak{a} + \mathfrak{b}) \subseteq \mathfrak{ab}$, and give an example in $\mathbb{Z}[x]$ where the inequality is strict.
- (3) Show that $(\mathfrak{a} + \mathfrak{c})(\mathfrak{b} + \mathfrak{c}) \subseteq \mathfrak{ab} + \mathfrak{c}$, and give examples where the inequality is strict.
- (4) Let $\mathfrak{a}, \mathfrak{b}$ be finitely generated ideals in an integral domain. Show that if $\mathfrak{a} + \mathfrak{b}$ is principal then $\mathfrak{a} \cap \mathfrak{b}$ is finitely generated. (Hint. Use the exact sequence (1).)

- (5) Let $K = k(x, y, z_1, z_2, \dots)$, where k is a field and x, y, z_i are indeterminates. Show that in the k -subalgebra of K generated by $x, y, z_i, xy^{-1}z_i (i = 1, 2, \dots)$ neither $(x) \cap (y)$ nor $(x):(y)$ is finitely generated.

9.2 Prime ideals and factorization

When discussing unique factorization domains, briefly UFDs, in Vol. 1 we defined atoms and primes. We recall that an *atom* in an integral domain R is a non-unit which cannot be written as a product of two non-units, and a *prime* is an element p , not zero or a unit, such that for any $a, b \in R$, $p|ab$ implies $p|a$ or $p|b$. Here $p|a$ (' p divides a ') means that $a = pc$ for some $c \in R$. Any prime is necessarily an atom and in a UFD the converse holds; in fact a UFD may be characterized as an integral domain with the properties:

- (i) *every element not zero or a unit can be written as a product of atoms;*
- (ii) *every atom is prime.*

An integral domain is said to be *atomic* if it satisfies (i). We note that a UFD may also be characterized as an integral domain in which every element not zero or a unit can be written as a product of primes. Moreover, an atomic domain in which the sum of any two principal ideals is principal is a UFD (Th. 3 of **10.5**, p. 321; Vol. 1).

In the study of algebraic number theory it was found that rings of algebraic integers always satisfy (i) but not necessarily (ii), and so they may not be UFD's; certain pairs of elements α, β failed to have an HCF, so that (α, β) was non-principal. It was this fact that led Kummer and Dedekind to develop ideal theory. Here the analogue of an atom is a *maximal* ideal, i.e. an ideal which is maximal among proper ideals (cf. 2.2). Even more important is the notion of prime ideal. In any commutative ring R , a *prime ideal* is a proper ideal \mathfrak{p} such that $xy \in \mathfrak{p}$ implies $x \in \mathfrak{p}$ or $y \in \mathfrak{p}$; note that R itself is not a prime ideal. To illustrate the definition, an element p is prime iff (p) is a non-zero prime ideal. An illustration of non-principal prime ideals is the following chain of prime ideals in the polynomial ring $k[x_1, \dots, x_n]$, where k is a field:

$$0 \subset (x_1) \subset (x_1, x_2) \subset \cdots \subset (x_1, x_2, \dots, x_n).$$

We note that an ideal \mathfrak{p} in a ring R is prime iff R/\mathfrak{p} is an integral domain. Since \mathfrak{p} is maximal iff R/\mathfrak{p} is a field, we deduce

PROPOSITION 2.1 *In a commutative ring R every maximal ideal is prime. In particular, every non-trivial commutative ring has prime ideals.*

To obtain the second assertion we note that a non-trivial ring has maximal ideals, by Krull's theorem (Th. 2.2.11). ■

Krull's theorem has a useful generalization. Let us call a subset S of R *multiplicative* if $1 \in S$ and $x, y \in S$ implies $xy \in S$.

THEOREM 2.2 *Let R be a commutative ring, S a multiplicative subset and \mathfrak{a} an ideal of R disjoint from S . Then there exists an ideal \mathfrak{m} in R which contains \mathfrak{a} , is disjoint from S and is maximal with respect to these properties. Any such ideal \mathfrak{m} is prime.*

Proof. Let \mathcal{A} be the set of all ideals \mathfrak{a}' with the properties $\mathfrak{a}' \supseteq \mathfrak{a}$, $\mathfrak{a}' \cap S = \emptyset$. \mathcal{A} is inductive, for if $\{\mathfrak{c}_\lambda\}$ is a chain of ideals in \mathcal{A} , their union \mathfrak{c} is an ideal containing \mathfrak{a} (of course, if the chain is empty, then $\mathfrak{c} = \mathfrak{a}$). If $\mathfrak{c} \cap S \neq \emptyset$, let $x \in \mathfrak{c} \cap S$; then $x \in \mathfrak{c}_\lambda$ for some λ and so $\mathfrak{c}_\lambda \cap S \neq \emptyset$, a contradiction. Thus \mathcal{A} is inductive; by Zorn's lemma it has a maximal member \mathfrak{m} and this is an ideal with the required properties.

Now let \mathfrak{m} be as stated and assume $b, c \notin \mathfrak{m}$, $bc \in \mathfrak{m}$. By the maximality of \mathfrak{m} , $(\mathfrak{m} + (b)) \cap S \neq \emptyset$, say $s = bu + x \in S$, where $u \in R$, $x \in \mathfrak{m}$, and similarly $t = cv + y \in S$, where $v \in R$, $y \in \mathfrak{m}$. Then S contains

$$st = (bu + x)(cv + y) = bcuv + x(cv + y) + buy \in \mathfrak{m},$$

and this contradicts the fact that $\mathfrak{m} \cap S = \emptyset$. Therefore $b, c \notin \mathfrak{m}$ implies $bc \notin \mathfrak{m}$ and since $1 \in S$, $1 \notin \mathfrak{m}$; this shows \mathfrak{m} to be prime. ■

If S is a multiplicative subset of R such that $0 \notin S$, then we can apply Th. 2.2 with $\mathfrak{a} = 0$ and obtain

COROLLARY 2.3 *Given a multiplicative set S not containing 0 in a commutative ring R , there exist ideals in R that are maximal subject to being disjoint from S .* ■

As examples of multiplicative sets we mention (i) the multiplicative set generated by an element f of R , i.e. $\{1, f, f^2, \dots\}$, (ii) the complement of a prime ideal and (iii) the set $1 + \mathfrak{a} = \{1 + a | a \in \mathfrak{a}\}$, where \mathfrak{a} is any ideal.

A multiplicative set S is said to be *saturated* if $ab \in S$ implies $a \in S$. Since $ab = ba$, it then also follows that $b \in S$. Clearly the complement of any prime ideal is multiplicative and saturated. In the opposite direction we have

PROPOSITION 2.4 *Let R be a commutative ring and S a subset. Then S is multiplicative and saturated if and only if its complement $R \setminus S$ is a union of prime ideals.*

Proof. Clearly the complement of any union of prime ideals is multiplicative and saturated. Conversely, if S is multiplicative and saturated, let $a \notin S$; then $ab \notin S$ for all $b \in R$, hence $(a) \cap S = \emptyset$, so by Th. 2.2 there is a prime ideal \mathfrak{p} containing a and disjoint from S . It follows that the union of all prime ideals disjoint from S is precisely $R \setminus S$. ■

We can also use Th. 2.2 to describe the intersection of all prime ideals in a ring.

PROPOSITION 2.5 *In a commutative ring R , the set \mathfrak{N} of all nilpotent elements is an ideal, equal to the intersection of all prime ideals.*

Proof. We have to show that

$$\mathfrak{N} = \bigcap \mathfrak{p}, \quad (1)$$

where the intersection is over all prime ideals of R . If $a \in \mathfrak{N}$, then $a^n = 0$ for some $n \geq 1$. Hence for any prime ideal \mathfrak{p} , $a^n \in \mathfrak{p}$, therefore $a \in \mathfrak{p}$; thus $\mathfrak{N} \subseteq \bigcap \mathfrak{p}$. Now let $a \notin \mathfrak{N}$; then $a^n \neq 0$ for all $n \geq 1$, hence $0 \notin \{1, a, a^2, \dots\}$ and by Th. 2.2 there is a prime ideal \mathfrak{p} disjoint from $\{1, a, a^2, \dots\}$. Hence the right-hand side of (1) does not contain a and it follows that we have equality. ■

The ideal \mathfrak{N} consisting of all nilpotent elements of R is called the *nilradical* of R . Given any ideal \mathfrak{a} of R , we can define the *radical* of \mathfrak{a} in R as

$$\sqrt{\mathfrak{a}} = \{x \in R \mid x^n \in \mathfrak{a} \text{ for some } n \geq 1\}.$$

Thus $\sqrt{\mathfrak{a}}$ is the inverse image, under the natural homomorphism $R \rightarrow R/\mathfrak{a}$, of the nilradical of R/\mathfrak{a} . Applying Prop. 2.5, and bearing in mind that the prime ideals of R/\mathfrak{a} are images of prime ideals of R , by the third isomorphism theorem, we see that $\sqrt{\mathfrak{a}}$ is the intersection of all prime ideals containing \mathfrak{a} .

In terms of prime ideals there is another criterion for unique factorization.

THEOREM 2.6 *An integral domain is a unique factorization domain if and only if every non-zero prime ideal contains a prime element.*

Proof. In any domain R let S be the set of all products of prime elements and units. Clearly S is multiplicative; it is also saturated, for if $ab \in S$, say $ab = p_1 \cdots p_r$ (p_i prime), then $p_1 | ab$, hence $p_1 | a$ or $p_1 | b$, say the former, $a = p_1 a_1$. Then $a_1 b = p_2 \cdots p_r$; now by an induction on r we see that $a_1, b \in S$, hence also $a = p_1 a_1 \in S$, so S is indeed saturated. By Prop. 2.4 its complement is a union of prime ideals. But any prime element belongs to S , so the prime ideals disjoint from S contain no prime elements. Hence if R satisfies the conditions of the theorem, then S consists of all non-zero elements and so R is then a UFD.

Conversely, let R be a UFD and \mathfrak{p} a non-zero prime ideal. Take $0 \neq a \in \mathfrak{p}$; by unique factorization, $a = p_1 \cdots p_r$, where p_i is prime, $r \geq 1$ and $p_1 \cdots p_r \in \mathfrak{p}$, hence $p_i \in \mathfrak{p}$ for some i , because \mathfrak{p} is a prime ideal. This shows the condition to be necessary. ■

In a UFD every minimal non-zero prime ideal is principal. For let \mathfrak{p} be a minimal non-zero prime ideal and p a prime element in \mathfrak{p} ; then $0 \subset (p) \subseteq \mathfrak{p}$, hence

$\mathfrak{p} = (p)$, by minimality, and so \mathfrak{p} is principal. In Noetherian rings the converse holds: a Noetherian domain in which each minimal non-zero prime ideal is principal is a UFD (cf. e.g. Nagata 1962).

Exercises

- (1) Show that $\sqrt{\sqrt{\mathfrak{a}}} = \sqrt{\mathfrak{a}}$ and $\sqrt{\mathfrak{a}^n} = \sqrt[n]{\mathfrak{a}}$, for any $n \geq 1$.
- (2) Show that $\sqrt{(\mathfrak{a}\mathfrak{b})} = \sqrt{(\mathfrak{a} \cap \mathfrak{b})} = \sqrt{\mathfrak{a}} \cap \sqrt{\mathfrak{b}}$ and $\sqrt{(\mathfrak{a} + \mathfrak{b})} = \sqrt{(\sqrt{\mathfrak{a}} + \sqrt{\mathfrak{b}})}$.
- (3) Show that in an Artinian ring every prime ideal is maximal.
- (4) Show that the intersection (and union) of any chain of prime ideals is again prime. Deduce the existence of minimal prime ideals in any non-trivial ring.
- (5) Show that if $a^m = b^n = 0$, then $(a + b)^{m+n-1} = 0$. Hence give a direct proof that the set of all nilpotent elements in a commutative ring is an ideal.
- (6) Show that an ideal \mathfrak{a} in a ring R is prime iff it is proper and for any ideals $\mathfrak{b}, \mathfrak{c} \supset \mathfrak{a}$ implies $\mathfrak{bc} \not\subseteq \mathfrak{a}$.
- (7) Show that for any ideal \mathfrak{a} in a Noetherian ring, $(\sqrt{\mathfrak{a}})^n \subseteq \mathfrak{a}$, for some n depending on \mathfrak{a} . Give an example to show that this may fail without the Noetherian condition.
- (8) Let R be a ring and \mathfrak{N} its nilradical. Show that the following conditions are equivalent:
 (a) R/\mathfrak{N} is an integral domain, (b) $\mathfrak{ab} = 0 \Rightarrow \mathfrak{a}^2 = 0$ or $\mathfrak{b}^2 = 0$, (c) $xy = 0 \Rightarrow x$ or y is nilpotent.

9.3 Localization

When we constructed the field of fractions of an integral domain in Ch. 6 of Vol. 1, we found the essential property of the set of denominators to be its closure under multiplication. Let us now take any commutative ring R , not necessarily an integral domain, and let S be a multiplicative subset of R . Then we can construct fractions a/s with denominators in S as follows. We define a relation on the product set $R \times S$ by setting

$$(a, s) \sim (a', s') \Leftrightarrow (as' - sa')t = 0 \quad \text{for some } t \in S. \quad (1)$$

This relation is clearly reflexive and symmetric; to prove that it is transitive, let $(a, s) \sim (a', s')$ and $(a', s') \sim (a'', s'')$, say $(as' - sa')t = 0$, $(a's'' - s'a'')t' = 0$, where $t, t' \in S$. Then we have

$$\begin{aligned} (as'' - sa'')s'tt' &= as's''tt' - sa's''tt' + sa's''tt' - sa''s'tt' \\ &= (as' - sa')t.s''t' + (a's'' - s'a'')t'.st \\ &= 0. \end{aligned}$$

Since $s'tt' \in S$, this proves that $(a, s) \sim (a'', s'')$. The proof shows why we had to introduce t in the definition (1). If S consists entirely of non-zerodivisors, we can replace the condition in (1) by $as' - sa' = 0$, and the definition reduces essentially to that given in 6.2, p. 141 of Vol. 1.

Thus the relation ' \sim ' defined by (1) is reflexive, symmetric and transitive and so is an equivalence on $R \times S$. Let us write $\frac{a}{s}$ or a/s for the equivalence class containing (a, s) and define addition and multiplication by the rules

$$\frac{a}{s} + \frac{a'}{s'} = \frac{as' + sa'}{ss'}, \quad \frac{a}{s} \cdot \frac{a'}{s'} = \frac{aa'}{ss'}. \quad (2)$$

It is routine to verify that these operations are well-defined and that the set R_S of all equivalence classes forms a ring under the operations (2), with $0/1$ as zero and $1/1$ as unit element. The natural mapping $\lambda: R \rightarrow R_S$ given by

$$\lambda: x \mapsto x/1 \quad (3)$$

is clearly a homomorphism which maps every element of S to a unit in R_S , for if $s \in S$, then $(s/1)(1/s) = s/s = 1$. This ring R_S is called the *ring of fractions* with denominators in S .

Given any ring R and a subset X of R , we say that a homomorphism $f: R \rightarrow R'$ is *X -inverting* if it maps the elements of X to invertible elements of R' . As we have just seen, the homomorphism (3) is S -inverting, but we can say more than that:

THEOREM 3.1 *Let R be a commutative ring and S a multiplicative subset of R . Then there exists a ring R_S and a homomorphism $\lambda: R \rightarrow R_S$ which is universal S -inverting, i.e. it is S -inverting and for every S -inverting homomorphism $f: R \rightarrow R'$ there is a unique homomorphism $f': R_S \rightarrow R'$ such that $f = \lambda f'$. Moreover, this property determines R_S up to isomorphism.*

The elements of R_S can be written as fractions a/s ($a \in R$, $s \in S$), where $a/s = a'/s'$ if and only if $(as' - sa')t = 0$ for some $t \in S$; the addition and multiplication in R_S are defined by (2) and λ is given by (3), with kernel

$$\ker \lambda = \{a \in R \mid at = 0 \text{ for some } t \in S\}.$$

Proof. We saw that λ as defined in (3) is S -inverting. Now let $f: R \rightarrow R'$ be any S -inverting homomorphism and define a mapping $f_1: R \times S \rightarrow R'$ by

$$(a, s)f_1 = (af)(sf)^{-1}.$$

This is possible because f is S -inverting, and f_1 takes the same value on equivalent pairs: if $(as' - sa')t = 0$, then $(af.s'f - sf.a'f)tf = 0$ and hence $af.(sf)^{-1} = a'f.(s'f)^{-1}$. Thus we obtain a well-defined mapping $f': R_S \rightarrow R'$ by putting $(a/s)f' = af.(sf)^{-1}$. This mapping is easily seen to be a homomorphism, using (3),

and it has the property

$$(a/1)f' = af, \quad (4)$$

i.e. $\lambda f' = f$. Moreover, it is the only such mapping, for (4) determines the values of f' on the elements $a/1$ and its value on $1/s$ must then be the inverse of its value on $s/1$. The uniqueness of R_S follows as usual by universality. Finally take $a \in \ker \lambda$; by (3) this means that $a/1 = 0/1$, i.e. $at = 0$ for some $t \in S$. ■

We see in particular that the mapping $\lambda: R \rightarrow R_S$ is injective precisely when S consists of non-zerodivisors. For example, when R is an integral domain and $S = R^\times$, then R_S is just the field of fractions, as constructed in 6.2, Vol. 1. At the other extreme, if $0 \in S$, then $a/s = 0$ for all $a \in R$, $s \in S$, because $(a.1 - s.0)0 = 0$; hence $R_S = 0$. Leaving this trivial case aside, we may suppose that $0 \notin S$.

An important application is to the case where S is the complement of a prime ideal p . Here one often writes R_p in place of R_S , as we have done in 8.4. The ring R_p just constructed is a *local ring*; in the canonical homomorphism $R \rightarrow R_p$ the prime ideal p of R corresponds to the unique maximal ideal of R_p consisting of all the non-units. This ring R_p is also called the *local ring of R at p* , and the process of forming R_p is called *localization*.

It is essential to be clear about the distinction between R/p and R_p . In rough terms we may think of R/p as being obtained from R by ‘putting the elements in p equal to 0’, while R_p is formed by ‘making the elements outside p invertible’. These two rings arise whenever we have a homomorphism from a ring into a field, $f: R \rightarrow k$ say, such that k is generated, as field, by the image of f . The kernel of this homomorphism is a prime ideal p say, and there are two ways of analysing f , which form a commutative diagram:

$$\begin{array}{ccc} R & \xrightarrow{\quad} & R \\ \downarrow & \searrow f & \downarrow \\ R/p & \xrightarrow{\quad} & k \end{array}$$

We can either take R/p , an integral domain, and form its field of fractions, or we can take the local ring R_p and form the quotient by its maximal ideal, called the *residue class field* of R_p . Each time we get the same field k . Although one habitually uses the first route, it turns out that the second route (via R_p) is easier to generalize to non-commutative rings (cf. Cohn 1985, Ch. 7).

It is important to know the relation between ideals in R and in R_S . With every ideal a in R we associate an *expanded ideal* a_S or a^e of R_S generated by the image $a\lambda$:

$$a^e = \{a/s \mid a \in a, s \in S\}.$$

To verify that α^e is an ideal, we note that if $a/s, a'/s' \in \alpha^e$, then $a/s + a'/s' = (as' + sa')/ss' \in \alpha^e$ and for any $b/t \in R_S$, $a/s \cdot b/t = ab/st \in \alpha^e$. Now take an ideal \mathfrak{A} in R_S and define the corresponding contracted ideal in R by

$$\mathfrak{A}^c = \mathfrak{A}\lambda^{-1} = \{x \in R \mid x\lambda \in \mathfrak{A}\}.$$

Clearly this is an ideal, and the next result shows when these operations are inverse:

PROPOSITION 3.2 *Let R be a commutative ring, S a multiplicative subset and R_S the corresponding ring of fractions, with the natural homomorphism $\lambda: R \rightarrow R_S$. Then*

- (i) *for any ideal \mathfrak{A} of R_S , $\mathfrak{A}^{ce} = \mathfrak{A}$,*
- (ii) *for any ideal α of R , $\alpha \subseteq \alpha^{ec}$, with equality if and only if no element of S is a zero divisor on R/α .*

When the condition (ii) holds, we shall say that S is α -regular. Thus S is α -regular iff for any $s \in S$, $x \in R$, $sx \in \alpha$ implies $x \in \alpha$.

Proof. (i) \mathfrak{A}^{ce} is the ideal of R_S generated by $(\mathfrak{A}\lambda^{-1})\lambda$; clearly $(\mathfrak{A}\lambda^{-1})\lambda \subseteq \mathfrak{A}$, hence $\mathfrak{A}^{ce} \subseteq \mathfrak{A}$ and we must prove the reverse inclusion. Let $a/s \in \mathfrak{A}$; then $a \in \mathfrak{A}^c$ and $a/s \in \mathfrak{A}^{ce}$, hence $\mathfrak{A} = \mathfrak{A}^{ce}$ as claimed.

(ii) For any $x \in \alpha$, $x/1 \in \alpha^e$ and so $x \in \alpha^{ec}$; this shows that $\alpha \subseteq \alpha^{ec}$. Now assume that $\alpha = \alpha^{ec}$ or more generally, $\alpha = \mathfrak{A}^c$ for some ideal \mathfrak{A} of R_S and let $sx \in \alpha$, where $x \in R$, $s \in S$. Then $sx/1 \in \mathfrak{A}$, hence $x/1 = sx/s \in \mathfrak{A}$, so $x \in \alpha$; this shows that S must be α -regular. Conversely, if S is α -regular, then $x \in \alpha^{ec} \Leftrightarrow x/1 \in \alpha^e \Leftrightarrow x/1 = a/s$ for some $a \in \alpha$, $s \in S$. This means that $(xs - a)t = 0$ for some $t \in S$, hence $xst \in \alpha$ and by regularity we may cancel st and find that $x \in \alpha$. Thus $\alpha^{ec} \subseteq \alpha$, and equality follows. ■

This result shows that there is a bijection between the ideals of R_S and the ideals α of R for which S is α -regular. In particular, if α is a prime ideal, S is α -regular iff $S \cap \alpha = \emptyset$ and we obtain

COROLLARY 3.3 *With the notation of Prop. 3.2 the correspondence $\mathfrak{A} \mapsto \mathfrak{A}^c$ is a bijection between the prime ideals of R_S and the prime ideals of R disjoint from S . In particular, for any prime ideal \mathfrak{p} of R the correspondence $\mathfrak{A} \mapsto \mathfrak{A}^c$ is a bijection between the prime ideals of $R_{\mathfrak{p}}$ and the prime ideals of R contained in \mathfrak{p} .* ■

This result shows that the localization of a local ring need not be local (e.g. invert $x + y$ in the power series ring $k[[x, y]]$). But this does hold for local Bezout domains, i.e. valuation rings.

The formation of fractions can also be extended to modules. Let R be a

commutative ring and M an R -module. Given a multiplicative subset S of R , we can define M_S as the set of equivalence classes on $M \times S$, where $(m, s) \sim (m', s')$ iff $(ms' - m's)t = 0$ for some $t \in S$. As before we can verify that this is indeed an equivalence and M_S becomes an R_S -module relative to the operations

$$m/s + m'/s' = (ms' + m's)/ss', \quad m/s \cdot a/t = ma/st. \quad (5)$$

Of course we can equally well regard M_S as R -module, using the equation

$$x \cdot a = x(a\lambda) \quad x \in M_S, a \in R,$$

to define the R -action in terms of the R_S -action. This is described as ‘pulling the action of R_S back along λ ’, or more briefly, defining the action of R by *pullback*. In detail, the action of R_S on M_S is defined by a homomorphism $R_S \rightarrow \text{End}(M_S)$ (given by the second equation (5)); we define the action of R by pullback along λ , using the composite map $R \xrightarrow{\lambda} R_S \rightarrow \text{End}(M_S)$.

There is a canonical mapping $\mu: M \rightarrow M_S$ given by $m \mapsto m/1$, whose kernel is given by

$$\ker \mu = \{m \in M \mid ms = 0 \text{ for some } s \in S\}. \quad (6)$$

The universal property of M_S is described in

PROPOSITION 3.4 *Let R be a commutative ring, S a multiplicative subset and M an R -module. Then there is an R -module M_S with a homomorphism $\mu: M \rightarrow M_S$ which is universal for homomorphisms of M into R -modules on which S acts by automorphisms.*

The proof is straightforward and may be left to the reader. ■

We remark that the correspondence $M \mapsto M_S$ is a functor; this is not hard to verify, and also follows from the fact (easily checked) that

$$M_S \cong M \otimes_R R_S. \quad (7)$$

In general the functor $M \mapsto M_S$ need not be faithful, since e.g. S may contain an element annihilating M . But we can obtain a faithful functor by localizing at all prime ideals, or even at all maximal ideals.

PROPOSITION 3.5 *Let R be a commutative ring and X the set of all maximal ideals of R . Then the functor*

$$M \mapsto \prod_{m \in X} M_m, \quad (8)$$

is faithful.

Proof. Let $\alpha: M \rightarrow N$ be a non-zero mapping, say $x\alpha \neq 0$, and define $\text{Ann}(x\alpha) = \{a \in R \mid (x\alpha)a = 0\}$. This is an ideal of R , proper because $x\alpha \neq 0$, and so contained

in some $m \in X$. Hence $x \notin \ker(N \rightarrow N_m)$ and so the induced mapping $\alpha_m: M_m \rightarrow N_m$ is non-zero; this shows (8) to be faithful. ■

For any R -module M , the set of prime ideals p in R such that $M_p \neq 0$ is called the *support* of M , written $\text{Supp}(M)$. By (6) any $x \in M$ maps to zero in M_p iff $\text{Ann}(x) \subseteq p$. Hence we see that for a finitely generated R -module M , $\text{Supp}(M)$ consists of those p for which $p \supseteq \text{Ann}(M)$.

A further useful property of our functor is given by

PROPOSITION 3.6 *Let R be a commutative ring and S a multiplicative subset. Then the localization functor $M \mapsto M_S$ is exact.*

Proof. Given a sequence

$$A \xrightarrow{\alpha} B \xrightarrow{\beta} C, \quad (9)$$

exact at B , consider the corresponding sequence

$$A_S \xrightarrow{\alpha_S} B_S \xrightarrow{\beta_S} C_S. \quad (10)$$

We have to show that $\text{im } \alpha_S = \ker \beta_S$. Clearly $\alpha_S \beta_S = 0$; conversely, let $b/s \in B_S$ and suppose that $(b/s)\beta_S = 0$, i.e. $b\beta/s = 0$. Then $b\beta \cdot t = 0$ for some $t \in S$, hence $(bt)\beta = 0$. By exactness of (9), $bt = a\alpha$ for some $a \in A$ and so $b/s = (a/st)\alpha_S$, and this shows (10) to be exact. ■

If we recall (7), we see that tensoring with R_S is an exact functor, by Prop. 3.6. This is expressed by saying that R_S as left R -module is *flat*, a fact which can also be verified directly (cf. Vol. 3).

We end this section with a useful relation between factorization in R and in R_S . The first part is well-known, the converse is essentially due to Nagata.

THEOREM 3.7 *Let R be an integral domain and S a multiplicative subset of R^\times . (i) If R is a UFD, then so is R_S . (ii) If R is an atomic domain, S consists of products of primes and R_S is a UFD, then R is a UFD.*

Proof. (i) We may regard R as a subring of R_S , because the natural mapping $\lambda: R \rightarrow R_S$ is injective, and since the elements of S are units in R_S , every element of R_S is associated in R_S to an element of R . In R each element has a factorization into primes, so (i) will follow if we show that every prime p in R either becomes a unit in R_S or it stays prime, depending on whether or not p divides an element of S .

If $p|s \in S$, then clearly p becomes a unit in R_S . Otherwise let $p|ab$ in R_S ; here a, b may be taken in R , by passing to associates. Then $ab = pd$, where $d = c/s$ ($c \in R$, $s \in S$), hence $abs = pc$. By hypothesis, $p \nmid s$, but p is prime in R , hence $p|a$ or $p|b$, and

this shows that p is prime in R_S . Thus every element of R_S has a factorization into primes, and this shows R_S to be a UFD.

(ii) We must show that every atom in R is prime. Thus let p be an atom in R ; we claim that either p remains an atom in R_S or it becomes a unit. For let $p = ab$ in R_S ; we shall show that a or b is a unit in R_S . Write $a = a'/s$, $b = b'/t$, where $a', b' \in R$, $s, t \in S$, so that $pst = a'b'$. By hypothesis st can be written as a product of primes: $st = q_1 \cdots q_r$, and $pq_1 \cdots q_r = a'b'$. Each q_i divides either a' or b' in R , and cancelling them one by one we obtain

$$p = a''b'', \quad (11)$$

where a'', b'' are the quotients of a', b' respectively, after division by the q 's. In R_S all the q 's are units, hence a'' is associated to a' and so also to a within R_S ; similarly b'' is associated to b in R_S . Now p is an atom in R ; by (11) either a'' or b'' is a unit, and accordingly, a or b is a unit in R_S . This shows p to be an atom or a unit in R_S .

We now treat these two cases separately:

(x) p remains an atom in R_S , hence p is prime in R_S because the latter is a UFD. If $p|ab$ in R , then $p|ab$ in R_S , hence p divides a or b in R_S , say the former: $a = pd$, $d = c/s$ ($c \in R$, $s \in S$), so

$$as = pc. \quad (12)$$

Now s is a product of primes in R : $s = s_1 \cdots s_r$, say, and no s_i divides p in R , for if $s_i|p$, then p would be associated to s_i in R and so would become a unit in R_S , which is not the case. Hence by (12), $s_i|c$ in R for $i = 1, \dots, r$ and cancelling s_1, \dots, s_r in turn from (12), we are left with the equation $a = pc'$, i.e. $p|a$ in R . This shows p to be prime in R .

(β) p becomes a unit in R_S . Then p divides some $s \in S$, say $s = pc$, and on cancelling the prime factors of s one by one we find that p is divisible by (and hence associated to) some prime factor s' of s ; therefore p is itself prime. ■

Our aim is to show that a polynomial ring over a UFD is again a UFD. We recall Gauss's lemma (Lemma 3.7.1): If A is an integral domain, then every prime in A stays prime in $A[x]$.

THEOREM 3.8 *If R is a UFD, then so is the polynomial ring $R[x]$.*

Proof. Let $K = R_R^\times$ be the field of fractions of R ; by the Euclidean algorithm $K[x]$ is a UFD (Th. 6, 6.5, p. 159 of Vol.1). Now $K[x] = R[x]_{R^\times}$ and R^\times consists of products of primes, because R is a UFD. Moreover, by Gauss's lemma, these primes stay prime in $R[x]$. It only remains to show that $R[x]$ is atomic. Given $f = a_0x^n + \cdots + a_n$ ($a_0 \neq 0$), if a_0 can be written as a product of k prime factors, then f can be factorized into at most $k + n$ non-unit factors, for the leading term of each factor must either contain x or a non-unit factor of a_0 . Thus $R[x]$ satisfies all the conditions of Th. 3.7 relative to the multiplicative set R^\times , and so $R[x]$ is a UFD, as claimed. ■

Since a field is trivially a UFD, we obtain by induction,

COROLLARY 3.9 *For any field k , the polynomial ring $k[x_1, \dots, x_n]$ in a number of indeterminates is a UFD.* ■

Exercises

- (1) Show that the set inverted in a homomorphism $R \rightarrow R'$ is multiplicative and saturated.
- (2) Show that every ring of fractions of \mathbf{Z}/n has the form \mathbf{Z}/m . When n is given, what integers m can occur?
- (3) Let R be a ring and S a multiplicative set. Show that if R is Noetherian, then so is R_S .
- (4) Let \mathfrak{a} be an ideal in R and S a multiplicative subset. Show that in R_S , $(\sqrt{\mathfrak{a}})_S = \sqrt{\mathfrak{a}_S}$.
- (5) Prove the isomorphism (7) by defining a bilinear mapping from M and R_S to M_S .
- (6) Verify that for a finitely generated R -module M , $\text{Supp}(M)$ consists of the prime ideals of R containing $\text{Ann}(M)$. Give an example to show that this fails for general R -modules.
- (7) Let M be an R -module and A, B submodules of M . Show that $A \subseteq B$ iff $A_{\mathfrak{m}} \subseteq B_{\mathfrak{m}}$ for all maximal ideals \mathfrak{m} of R .
- (8) Let R be an integral domain with field of fractions K , so that every localization of R is embedded in K . Show that $\bigcap R_{\mathfrak{m}} = R$, where \mathfrak{m} runs over all maximal ideals of R . (*Hint.* Apply Prop. 3.5 to the R -linear maps from R to K/R .)
- (9) In Prop. 3.2 show that $\mathfrak{a}^{ec} = \{a \in R \mid sa \in \mathfrak{a} \text{ for some } s \in S\}$.
- (10) Show that the following are equivalent, for any ring R : (a) $R_{\mathfrak{p}}$ is a domain for each prime ideal \mathfrak{p} , (b) $R_{\mathfrak{m}}$ is a domain for each maximal ideal \mathfrak{m} , (c) if $ab = 0$ in R , then $\text{Ann}(a)$ and $\text{Ann}(b)$ are comaximal in R .

9.4 Noetherian rings

We recall from 2.2 that a ring R is *Noetherian* if every ascending chain of ideals breaks off, or equivalently, if every ideal is finitely generated. For example, any principal ideal domain is Noetherian, and so are many of the rings encountered in number theory and algebraic geometry.

In a Noetherian ring R the nilradical \mathfrak{N} is nilpotent, i.e. there exists an integer n such that $\mathfrak{N}^n = 0$. For let $\mathfrak{N} = (a_1, \dots, a_r)$ and suppose that $a_i^{v_i+1} = 0$; write $v = \sum v_i$ and consider \mathfrak{N}^{v+1} . It is generated by the products of $v+1$ factors a_i , hence some a_i must occur to a power $\geq v_i + 1$ and so each product is zero; thus $\mathfrak{N}^{v+1} = 0$. More generally this argument shows that for any ideal \mathfrak{a} of R , some power of $\sqrt{\mathfrak{a}}$ is contained in \mathfrak{a} .

PROPOSITION 4.1 *Every Noetherian domain is atomic.*

Proof. We must show that every element not zero or a unit can be expressed as a product of a finite number of atoms. Let S be the set of all units and products of atoms in R ; if R contains a non-zero element c not in S , choose such c subject to the further condition that the ideal (c) is maximal among all principal ideals (c') with $c' \notin S$. Then c is not an atom or a unit, so we can write $c = ab$, where $(c) \subset (a), (b)$. By the maximality of c it follows that $a, b \in S$, say $a = a_1 \cdots a_r$, $b = b_1 \cdots b_s$ where the a_i, b_j are atoms (or units). But then $c = a_1 \cdots a_r b_1 \cdots b_s \in S$, which is a contradiction. Hence $S = R^\times$, and so R is atomic. ■

We remark that this result is just an instance of the decomposition lemma (Lemma 2.2.12), but it seemed clearer to have a separate proof.

The most important source of Noetherian rings is the following result.

THEOREM 4.2 (Hilbert basis theorem) *Let R be a Noetherian ring. Then the polynomial ring $R[x]$ is again Noetherian.*

Proof. (H. Sarges) We assume that $A = R[x]$ is not Noetherian and show that R cannot be Noetherian. Let \mathfrak{a} be an ideal in A which is not finitely generated, and take a non-zero polynomial f_1 of least degree in \mathfrak{a} . By induction, if we have found $f_1, \dots, f_k \in \mathfrak{a}$, we take $f_{k+1} \in \mathfrak{a} \setminus \sum_1^k f_i A$ of least possible degree. Since \mathfrak{a} is not finitely generated, we thus obtain an infinite sequence of polynomials f_1, f_2, \dots in \mathfrak{a} . Let f_i have degree n_i and leading coefficient a_i . We have $n_1 \leq n_2 \leq \dots$ by construction; we claim that

$$a_1 R \subset a_1 R + a_2 R \subset \dots \quad (1)$$

is an infinite ascending chain, which will contradict the fact that R is Noetherian. If the chain (1) breaks off, we have $a_{k+1} = \sum_1^k a_i b_i$ for some $b_i \in R$, but then $f_{k+1} - \sum f_i x^{n_{k+1}-n_i} b_i$ would be an element in $\mathfrak{a} \setminus \sum_1^k f_i A$ of degree less than n_{k+1} , which contradicts the definition of f_{k+1} . This establishes the result. ■

By induction on n we obtain

COROLLARY 4.3 *If R is Noetherian, then so is $R[x_1, \dots, x_n]$. In particular, $k[x_1, \dots, x_n]$ is Noetherian for any field k and any $n \geq 1$.* ■

Of course this result does not extend to infinitely many indeterminates; thus for any non-zero ring R we have

$$R \subset R[x_1] \subset R[x_1, x_2] \subset \dots$$

and we can form the union $R[x_1, x_2, \dots]$ as a polynomial ring in countably many

indeterminates $x_i (i \in \mathbb{N})$. This is an integral domain if R is, but it is not Noetherian, for the ideal (x_1, x_2, \dots) is not finitely generated.

PROPOSITION 4.4 *Let K be a commutative Noetherian ring. Then any commutative ring R which is finitely generated as K -algebra is Noetherian.*

Proof. Let R be generated by c_1, \dots, c_n over K . Then R is a homomorphic image of the polynomial ring $K[x_1, \dots, x_n]$ obtained by mapping $x_i \mapsto c_i$, say $R \cong K[x_1, \dots, x_n]/\mathfrak{a}$. Thus the ideals of R correspond to the ideals of $K[x_1, \dots, x_n]$ which contain \mathfrak{a} , and hence satisfy the maximum condition, because the polynomial ring does. ■

Exercises

- (1) Prove Prop. 4.1 by considering the monoid of classes of associated elements and applying Lemma 2.2.12.
- (2) Show that if R is a Noetherian ring, then so is $R[[x]]$, the ring of formal power series in x .
- (3) Show that the polynomial ring $k[x_1, x_2, \dots]$ in countably many indeterminates over a field k is a UFD.
- (4) Let R be a ring and $\mathfrak{a}_1, \dots, \mathfrak{a}_r$ ideals of R such that $\bigcap \mathfrak{a}_i = 0$. Show that if R/\mathfrak{a}_i is Noetherian for each i , then R is Noetherian.
- (5) Let \mathfrak{a} be a finitely generated ideal in a ring R . Show that if $\mathfrak{a}^2 = \mathfrak{a}$, then \mathfrak{a} is generated by a single idempotent element.
- (6) Show that the relation $\mathfrak{a} = \text{Ann}(b)$ between ideals \mathfrak{a} and b in a ring R defines a Galois connexion which is a lattice anti-isomorphism between annihilator ideals. Show that if every ideal of R is the annihilator of a finite set, then R is Artinian.

9.5 Dedekind domains

The phenomenon of non-unique factorization was first encountered by E. Kummer in his work on Fermat's last theorem, a famous conjecture which states that the equation

$$x^n + y^n = z^n, \quad \text{where } n > 2, \tag{1}$$

has no solution in non-zero integers x, y, z . It was asserted by Fermat without proof and no proof is known, despite intensive efforts over 350 years. It should be remarked however, that the impossibility of (1) has been proved for a very large number of values of n (certainly for all $n < 100\,000$). Moreover, G. Faltings in 1983 showed that for any $n \geq 3$ there are at most a finite number of solutions with x, y, z coprime.

An early method of attack was to factorize the left-hand side of (1) in $\mathbf{Z}[\zeta_n]$, where ζ_n is a primitive n th root of 1. If $\mathbf{Z}[\zeta_n]$ is a UFD, this leads to a proof of Fermat's last theorem for this value of n . But in general $\mathbf{Z}[\zeta_n]$ is not a UFD, and Kummer's investigations of such rings led him to the creation of his theory of 'ideal numbers' and hence Dedekind to his ideal theory. One of Kummer's discoveries was that every non-zero ideal in $\mathbf{Z}[\zeta]$ can be written uniquely as a finite product of prime ideals. Later, Dedekind showed that the same is true in the ring of integers of any algebraic number field, i.e. the integral closure of \mathbf{Z} in a finite extension field of \mathbf{Q} . This property is also of interest in algebraic geometry, where it describes the rings of non-singular curves.

Before examining these rings abstractly let us look at some concrete examples. In Vol. 1 (Th. 1 of **10.5**, p. 319) we saw that every Euclidean domain is a UFD; this enables us to show that the ring $\mathbf{Z}[i]$ of Gaussian integers is a UFD. We shall do this by verifying that the ring is Euclidean with respect to the usual norm

$$N(x + iy) = x^2 + y^2.$$

We have to show that for any $a, b \in \mathbf{Z}[i]$, $b \neq 0$, there exists c such that

$$N(a - bc) < N(b). \quad (2)$$

We note that $\mathbf{Z}[i]$ is the precise ring of integers in $\mathbf{Q}(i)$. On dividing by b and remembering that N is multiplicative, we obtain $N(a/b - c) < 1$. In other words, we have to show that for each $\gamma \in \mathbf{Q}(i)$ there exists $c \in \mathbf{Z}[i]$ such that

$$N(\gamma - c) < 1. \quad (3)$$

Put $\gamma = x + iy$, where x, y are rational numbers. Every rational number is within $\frac{1}{2}$ of an integer, so we can find rational integers u, v such that $|x - u| \leq \frac{1}{2}, |y - v| \leq \frac{1}{2}$, and on writing $c = u + iv$, we have

$$N(\gamma - c) = (x - u)^2 + (y - v)^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2} < 1.$$

Thus (3) is established and $\mathbf{Z}[i]$ is Euclidean. The same method can be used on the ring of integers in $\mathbf{Q}(\sqrt{-d})$, for $d = 2, 3, 7, 11$ (cf. Ex. (6)).

Next consider $\mathbf{Q}(\sqrt{-5})$; its ring of integers is $\mathbf{Z}[\sqrt{-5}]$, and we have

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}). \quad (4)$$

By considering norms we see that $2, 3, 1 \pm \sqrt{-5}$ are atoms. For example, $N(2) = 4$, so if 2 is composite, then a proper factor must have norm 2, but the equation $u^2 + 5v^2 = 2$ has no solutions in integers. Similarly for 3, $1 \pm \sqrt{-5}$; hence (4) represents two distinct factorizations of 6, and this shows that $\mathbf{Z}[\sqrt{-5}]$ is not a UFD. However, the ideal (6) can be expressed uniquely as a product of maximal ideals

$$(6) = (2, 1 + \sqrt{-5})^2 \cdot (3, 1 + \sqrt{-5}) \cdot (3, 1 - \sqrt{-5}).$$

We begin with the most general case. Let \mathfrak{o} be an integral domain and K its field of fractions. By a *fractional ideal* of \mathfrak{o} we understand an \mathfrak{o} -submodule \mathfrak{A} of A such that

$$z\mathfrak{o} \subseteq \mathfrak{A} \subseteq u\mathfrak{o}, \quad \text{for some } z, u \in K^\times. \quad (5)$$

We note that (5) certainly holds when $0 \neq \mathfrak{A} \subseteq \mathfrak{o}$. Thus an ordinary ideal \mathfrak{a} of \mathfrak{o} is a fractional ideal iff it is non-zero; the non-zero ideals of \mathfrak{o} will be called the *integral ideals*. The usual multiplication can be defined for fractional ideals:

$$\mathfrak{AB} = \{\sum x_i y_i \mid x_i \in \mathfrak{A}, y_i \in \mathfrak{B}\}, \quad (6)$$

and it is clear from (5) that this product is again a fractional ideal. Hence in any integral domain \mathfrak{o} the fractional ideals form a monoid (with \mathfrak{o} itself as neutral), which will be denoted by $I = I(\mathfrak{o})$.

For each fractional ideal \mathfrak{A} we can define an ‘inverse’

$$(\mathfrak{o} : \mathfrak{A}) = \{x \in K \mid x\mathfrak{A} \subseteq \mathfrak{o}\}.$$

If $z\mathfrak{o} \subseteq \mathfrak{A} \subseteq u\mathfrak{o}$, then $u^{-1}\mathfrak{o} \subseteq (\mathfrak{o} : \mathfrak{A}) \subseteq z^{-1}\mathfrak{o}$, and if $c \in \mathfrak{o}$, then $x\mathfrak{A} \subseteq \mathfrak{o}$ implies $cx\mathfrak{A} \subseteq x\mathfrak{A} \subseteq \mathfrak{o}$. This shows that $(\mathfrak{o} : \mathfrak{A})$ is again a fractional ideal. Any fractional ideal \mathfrak{A} satisfies

$$\mathfrak{A}(\mathfrak{o} : \mathfrak{A}) \subseteq \mathfrak{o}, \quad (7)$$

but here equality need not hold. If it does hold, then \mathfrak{A} is said to be *invertible* and we also write \mathfrak{A}^{-1} in place of $(\mathfrak{o} : \mathfrak{A})$. For example, any non-zero principal ideal $a\mathfrak{o}$ is invertible, with inverse $a^{-1}\mathfrak{o}$. We remark that if $\mathfrak{AB} = \mathfrak{o}$ for some fractional ideal \mathfrak{B} , then $\mathfrak{B} \subseteq (\mathfrak{o} : \mathfrak{A})$, hence $\mathfrak{o} = \mathfrak{AB} \subseteq \mathfrak{A}(\mathfrak{o} : \mathfrak{A})$, and it follows that $\mathfrak{A}(\mathfrak{o} : \mathfrak{A}) = \mathfrak{o}$, so that \mathfrak{A} is then invertible. Thus the invertible fractional ideals are just the units of $I(\mathfrak{o})$.

We note that in an integral domain \mathfrak{o} an ideal is isomorphic to \mathfrak{o} iff it is non-zero principal; further, any fractional ideal is isomorphic to an integral ideal, for if $\mathfrak{A} \subseteq u\mathfrak{o}$, then $u^{-1}\mathfrak{A} \subseteq \mathfrak{o}$ and the ideals \mathfrak{A} , $u^{-1}\mathfrak{A}$ are isomorphic via the mapping $x \mapsto u^{-1}x$.

We first give some characterizations of Dedekind domains, including one in homological terms.

PROPOSITION 5.1 *For any integral domain \mathfrak{o} , the following conditions are equivalent:*

- (a) *the set $I(\mathfrak{o})$ of fractional ideals is a group under the multiplication (6),*
- (b) *every fractional ideal of \mathfrak{o} is invertible,*
- (c) *every integral ideal of \mathfrak{o} is invertible,*
- (d) *every ideal of \mathfrak{o} is projective.*

Moreover, any ring satisfying (a)–(d) is Noetherian.

Proof. (a) \Leftrightarrow (b) \Rightarrow (c) are clear.

(c) \Rightarrow (d). Let \mathfrak{a} be an ideal in \mathfrak{o} ; we may assume that $\mathfrak{a} \neq 0$. By hypothesis \mathfrak{a} is invertible, say $\mathfrak{ab} = \mathfrak{o}$. Hence there exist $a_i \in \mathfrak{a}$, $b_i \in \mathfrak{b}$ ($i = 1, \dots, n$) such that $\sum a_i b_i = 1$ and $a_i b_i \subseteq \mathfrak{o}$. It follows that multiplication by b_i defines a homomorphism $\mathfrak{a} \rightarrow \mathfrak{o}$. Now for any $x \in \mathfrak{a}$, $b_i x \in \mathfrak{o}$ and

$$x = \sum a_i (b_i x); \quad (8)$$

therefore, by the dual basis lemma (Prop. 4.5.5), \mathfrak{a} is finitely generated projective.

(d) \Rightarrow (a). Let $(a_i), (f_i)$ ($i \in I$) be dual bases for the projective ideal \mathfrak{a} ; thus

$$x = \sum a_i (f_i, x). \quad (9)$$

Given $x, y \in \mathfrak{a}$, we have for any $i \in I$, $(f_i, x)y = (f_i, xy) = (f_i, y)x$. If $x \neq 0$, we can write $b_i = (f_i, x)x^{-1}$; the b_i lie in K , almost all are 0 and $b_i y = (f_i, y)$ for all $y \in \mathfrak{a}$; hence $b_i \mathfrak{a} \subseteq \mathfrak{o}$ and (9) reduces to (8). Dividing by x we have $\sum a_i b_i = 1$, so on putting $\mathfrak{b} = \sum b_i \mathfrak{o}$, we have $\mathfrak{ab} = \mathfrak{o}$, and this shows \mathfrak{a} to be invertible.

This proves (a)–(d) to be equivalent; moreover, we saw that every invertible ideal is finitely generated, so \mathfrak{o} must be Noetherian. ■

We remark that the proof shows an ideal in an integral domain to be invertible iff it is non-zero projective.

An integral domain satisfying any of these equivalent properties is called a *Dedekind domain*. It is clear from (c) or (d) that any principal ideal domain is a Dedekind domain. A ring in which every ideal is projective is also called *hereditary*; thus Dedekind domains may be described as hereditary integral domains.

If S is a multiplicative set in a Dedekind domain \mathfrak{o} , then the ideals of the ring of fractions \mathfrak{o}_S correspond to the contracted ideals of \mathfrak{o} , by Prop. 3.2. Using (c) of Prop. 5.1, we obtain

COROLLARY 5.2 *Let \mathfrak{o} be a Dedekind domain and S a multiplicative subset. Then \mathfrak{o}_S is again a Dedekind domain.* ■

Dedekind domains have been characterized by E. Noether as Noetherian integrally closed domains in which all non-zero prime ideals are maximal. To prove this result we need a variant of the decomposition lemma (Lemma 2.2.12), which has a similar proof.

LEMMA 5.3 *Any ideal \mathfrak{a} of a non-trivial Noetherian ring R contains a finite product of non-zero prime ideals:*

$$\mathfrak{a} \supseteq \mathfrak{p}_1 \dots \mathfrak{p}_r, \quad \text{where } \mathfrak{p}_i \text{ is a prime ideal } \neq 0,$$

unless $\mathfrak{a} = 0$ and R is an integral domain, or $\mathfrak{a} = R$ and R is a field.

Proof. Assume the contrary and let \mathfrak{a} be a ‘maximal offender’, i.e. \mathfrak{a} is maximal among the ideals that do not contain a finite product of non-zero prime ideals. If $\mathfrak{a} = R$, there is nothing to prove, since every ring not zero or a field has non-zero prime ideals. If $\mathfrak{a} \neq R$, then either \mathfrak{a} is prime, but then $\mathfrak{a} = 0$ and this leads to the excluded case, or there exist $\mathfrak{b}_1, \mathfrak{b}_2 \supset \mathfrak{a}$ such that $\mathfrak{b}_1 \mathfrak{b}_2 \subseteq \mathfrak{a}$. By maximality, $\mathfrak{b}_1 \supseteq \mathfrak{p}_1 \dots \mathfrak{p}_r$, $\mathfrak{b}_2 \supseteq \mathfrak{p}_{r+1} \dots \mathfrak{p}_s$ where the \mathfrak{p}_i are non-zero prime ideals. Hence $\mathfrak{a} \supseteq \mathfrak{b}_1 \mathfrak{b}_2 \supseteq \mathfrak{p}_1 \dots \mathfrak{p}_s$; this is a contradiction, and it establishes the desired conclusion. ■

We can now state E. Noether’s result; it is convenient to include several characterizations.

THEOREM 5.4 *Let \mathfrak{o} be an integral domain. Then the following conditions are equivalent:*

- (a) \mathfrak{o} is a Dedekind domain,
- (b) \mathfrak{o} is Noetherian and $\mathfrak{o}_{\mathfrak{m}}$ is a principal valuation ring (or a field) for all maximal ideals \mathfrak{m} of \mathfrak{o} ,
- (c) \mathfrak{o} is Noetherian, integrally closed and every non-zero prime ideal is maximal,
- (d) every non-zero prime ideal of \mathfrak{o} is invertible.

Proof. (a) \Rightarrow (b). Let \mathfrak{o} be a Dedekind domain; by Prop. 5.1 it is Noetherian, and by Cor. 5.2, $\mathfrak{o}_{\mathfrak{m}}$ is a Dedekind domain, for any maximal ideal \mathfrak{m} of \mathfrak{o} . By Prop. 5.1(c), if $a, b \in \mathfrak{o}_{\mathfrak{m}}$ are non-zero and $\mathfrak{A} = a\mathfrak{o}_{\mathfrak{m}} + b\mathfrak{o}_{\mathfrak{m}}$, then $a(\mathfrak{o}_{\mathfrak{m}}:\mathfrak{A}) + b(\mathfrak{o}_{\mathfrak{m}}:\mathfrak{A}) = \mathfrak{o}_{\mathfrak{m}}$. But this ring is local, so one of the two ideals must be the whole ring. If $a(\mathfrak{o}_{\mathfrak{m}}:\mathfrak{A}) = \mathfrak{o}_{\mathfrak{m}}$, then $a^{-1} \in (\mathfrak{o}_{\mathfrak{m}}:\mathfrak{A})$ and so $a^{-1}b \in \mathfrak{o}_{\mathfrak{m}}$. Thus either $a^{-1}b$ or $b^{-1}a$ is in $\mathfrak{o}_{\mathfrak{m}}$, so $\mathfrak{o}_{\mathfrak{m}}$ is a valuation ring; being Noetherian, it is either a field or a principal valuation ring.

(b) \Rightarrow (c). Assume (b), thus \mathfrak{o} is Noetherian and for every maximal ideal \mathfrak{m} , $\mathfrak{o}_{\mathfrak{m}}$ is a valuation ring. If $x \in \mathfrak{o}_{\mathfrak{m}}$, say $x = a/s$, where $a, s \in \mathfrak{o}$, $s \notin \mathfrak{m}$, then $s \in (\mathfrak{o}:x\mathfrak{o})$ and so $\mathfrak{o} \cap (\mathfrak{o}:x\mathfrak{o})$ meets $\mathfrak{o} \setminus \mathfrak{m}$. It follows that if $x \in \bigcap \mathfrak{o}_{\mathfrak{m}}$, then $\mathfrak{o} \cap (\mathfrak{o}:x\mathfrak{o})$ meets the complement of each maximal ideal and so must be the whole of \mathfrak{o} , hence $(\mathfrak{o}:x\mathfrak{o}) \supseteq \mathfrak{o}$, and so $x \in \mathfrak{o}$. This shows that $\mathfrak{o} = \bigcap \mathfrak{o}_{\mathfrak{m}}$, and so \mathfrak{o} is integrally closed, by Th. 8.4.4. Finally, if \mathfrak{p} is a non-zero prime ideal of \mathfrak{o} , then \mathfrak{p} is contained in a maximal ideal \mathfrak{m} and so $\mathfrak{p}\mathfrak{o}_{\mathfrak{m}}$ is a non-zero prime ideal in $\mathfrak{o}_{\mathfrak{m}}$, which can only be $\mathfrak{m}\mathfrak{o}_{\mathfrak{m}}$, by (b); therefore $\mathfrak{p} = \mathfrak{m}$ by Cor. 3.3 and (c) follows.

(c) \Rightarrow (d). Let \mathfrak{p} be a non-zero prime ideal of \mathfrak{o} and let $0 \neq a \in \mathfrak{p}$. By Lemma 5.3 we have $(\mathfrak{a}) \supseteq \mathfrak{p}_1 \dots \mathfrak{p}_r$, where the \mathfrak{p}_i are non-zero prime ideals. Let us choose r minimal; since $\mathfrak{p} \supseteq (\mathfrak{a}) \supseteq \mathfrak{p}_1 \dots \mathfrak{p}_r$ and \mathfrak{p} is prime, we have $\mathfrak{p} \supseteq \mathfrak{p}_i$ for some i , say $i = 1$. By (c), \mathfrak{p}_1 is maximal, so $\mathfrak{p} = \mathfrak{p}_1$ and by the minimality of r we can find $b \in \mathfrak{p}_2 \dots \mathfrak{p}_r$ such that $b \notin (\mathfrak{a})$. However, $b\mathfrak{p} \subseteq (\mathfrak{a})$, i.e. $a^{-1}b\mathfrak{p} \subseteq \mathfrak{o}$, so $a^{-1}b \in (\mathfrak{o}:\mathfrak{p})$. Now $\mathfrak{p} \subseteq (\mathfrak{o}:\mathfrak{p})\mathfrak{p} \subseteq \mathfrak{o}$, so $(\mathfrak{o}:\mathfrak{p})\mathfrak{p}$ is \mathfrak{p} or \mathfrak{o} . If $(\mathfrak{o}:\mathfrak{p})\mathfrak{p} = \mathfrak{p}$, then $a^{-1}b\mathfrak{p} = \mathfrak{p}$, and by Prop. 8.4.1(c), $a^{-1}b$ is then integral over \mathfrak{o} . Since \mathfrak{o} is integrally closed, this means

that $a^{-1}b \in \mathfrak{o}$, but this contradicts the fact that $b \notin (a)$. Therefore $(\mathfrak{o}:p)p = \mathfrak{o}$ and p is invertible, as claimed in (d).

(d) \Rightarrow (a). Suppose that (d) holds and (a) fails. Then the set \mathcal{P} of non-invertible integral ideals of \mathfrak{o} is non-empty. Under the ordering by inclusion \mathcal{P} is inductive, for if $\{a_\lambda\}$ is a chain in \mathcal{P} and $\mathfrak{a} = \bigcup a_\lambda$ is invertible, then it is finitely generated and so $\mathfrak{a} = a_\lambda$ for some λ , a contradiction. Thus \mathcal{P} is inductive and by Zorn's lemma it has a maximal member \mathfrak{m} , say. By (d), \mathfrak{m} is not prime or \mathfrak{o} , so there exist $a, b \notin \mathfrak{m}$ such that $ab \in \mathfrak{m}$, thus $b \in \mathfrak{m}:(a) \supset \mathfrak{m}$. By the maximality of \mathfrak{m} , $\mathfrak{m} + (a)$ and $\mathfrak{m}:(a)$ are invertible, hence so is $a(\mathfrak{m}:(a)) = \mathfrak{m} \cap (a)$. Now consider the short exact sequence

$$0 \rightarrow \mathfrak{m} \cap (a) \xrightarrow{\lambda} \mathfrak{m} \oplus (a) \xrightarrow{\mu} \mathfrak{m} + (a) \rightarrow 0, \quad (10)$$

where $(x, ya)\mu = x - ya$. The third term $\mathfrak{m} + (a)$ is invertible and hence projective; therefore (10) splits and we have $\mathfrak{m} \oplus (a) \cong [\mathfrak{m} \cap (a)] \oplus [\mathfrak{m} + (a)]$. Here the right-hand side is projective, hence \mathfrak{m} is projective, which is a contradiction. It follows that \mathcal{P} is empty, as we had to show. ■

The advantage of Dedekind domains over UFDs is that the former survive finite integral extensions, whereas the latter may not, as we saw at the beginning of this section. We now apply Noether's characterization to prove that a finite separable integrally closed integral extension of a Dedekind domain is again Dedekind; this still holds in the inseparable case, but with a different proof (cf. Ex. (8)).

THEOREM 5.5 *Let \mathfrak{o} be a Dedekind domain with field of fractions K . Given a finite separable extension L of K , the integral closure \mathfrak{D} of \mathfrak{o} in L is again a Dedekind domain.*

Proof. The ring \mathfrak{D} is integrally closed by construction. To show that non-zero prime ideals are maximal, let $\mathfrak{P} \subseteq \mathfrak{P}'$ be prime ideals in \mathfrak{D} such that $\mathfrak{P} \cap \mathfrak{o} = \mathfrak{P}' \cap \mathfrak{o}$. If $x \in \mathfrak{P}'$, take a monic equation for x over \mathfrak{o} and consider it mod \mathfrak{P} :

$$x^n + a_1 x^{n-1} + \cdots + a_n \equiv 0 \pmod{\mathfrak{P}}, \quad \text{where } a_i \in \mathfrak{o}.$$

Choose such a congruence for which n has its least value. We have $a_n \in \mathfrak{P}' \cap \mathfrak{o} = \mathfrak{P} \cap \mathfrak{o}$, hence $x(x^{n-1} + a_1 x^{n-2} + \cdots + a_{n-1}) \in \mathfrak{P}$, and since \mathfrak{P} is prime, we have $x \in \mathfrak{P}$ by the minimality of n . This shows that $\mathfrak{P}' = \mathfrak{P}$. Since every non-zero prime ideal of \mathfrak{o} is maximal, the same holds for \mathfrak{D} .

It remains to show that \mathfrak{D} is Noetherian. Let u_1, \dots, u_n be a K -basis for L ; on multiplying by suitable elements of K we may assume that $u_i \in \mathfrak{D}$. Since the trace $T(xy)$ is a non-singular pairing on L (cf. Prop. 3.9.5), we can find a dual basis v_1, \dots, v_n relative to T ; thus we have $T(u_i v_j) = \delta_{ij}$. Now if $x \in \mathfrak{D}$, we have $x = \sum \alpha_i v_i$ for some $\alpha_i \in K$, and in fact $\alpha_i = T(u_i x) \in \mathfrak{o}$, hence $\mathfrak{D} \subseteq \sum \mathfrak{o} v_i$. Therefore \mathfrak{D} is a

submodule of a finitely generated \mathfrak{o} -module, and since \mathfrak{o} is Noetherian, it follows that every ideal in \mathfrak{O} is finitely generated, hence \mathfrak{O} is Noetherian. ■

As a consequence, the ring \mathfrak{O} of integers of an algebraic number field L is a Dedekind domain, since \mathbf{Z} is clearly a Dedekind domain with field of fractions \mathbf{Q} . Similarly the ring of functions integral over $k[x]$ in a finite separable extension of $k(x)$ is also a Dedekind domain.

We now come to the unique factorization property of Dedekind domains mentioned at the beginning of this section; in fact this provides another characterization.

THEOREM 5.6 *For any integral domain \mathfrak{o} , the following properties are equivalent:*

- (a) \mathfrak{o} is a Dedekind domain,
- (b) every integral ideal of \mathfrak{o} can be expressed uniquely as a finite product of maximal ideals,
- (c) every integral ideal of \mathfrak{o} can be expressed as a finite product of prime ideals.

Proof. (a) \Rightarrow (b). Let \mathfrak{o} be a Dedekind domain and I_0 the set of integral ideals of \mathfrak{o} which can be expressed as a finite product of maximal ideals. Clearly I_0 contains \mathfrak{o} , as the empty product. If some integral ideal does not belong to I_0 , let \mathfrak{a} be a maximal such ideal; this exists because \mathfrak{o} is Noetherian. Since $\mathfrak{a} \neq \mathfrak{o}$, $\mathfrak{a} \subseteq \mathfrak{m} \subset \mathfrak{o}$ for some maximal ideal \mathfrak{m} . By Prop. 5.1(c), $\mathfrak{a} = \mathfrak{m}\mathfrak{b}$ for some integral ideal \mathfrak{b} , and since $I(\mathfrak{o})$ is a group, we have $\mathfrak{a} \neq \mathfrak{b}$. Thus $\mathfrak{a} \subset \mathfrak{b}$; $\mathfrak{b} \in I_0$ by the maximality of \mathfrak{a} , and hence $\mathfrak{a} = \mathfrak{m}\mathfrak{b} \in I_0$, a contradiction. Thus every integral ideal can be expressed as a finite product of maximal ideals. Any such expression is unique, for if $\mathfrak{p}_1 \cdots \mathfrak{p}_r = \mathfrak{q}_1 \cdots \mathfrak{q}_s$, then since $\mathfrak{p}_1 \supseteq \mathfrak{q}_1 \cdots \mathfrak{q}_s$ and \mathfrak{p}_1 is prime, we have $\mathfrak{p}_1 \supseteq \mathfrak{q}_i$ for some i . But \mathfrak{q}_i is maximal, so $\mathfrak{p}_1 = \mathfrak{q}_i$ and we may cancel these terms, because $I(\mathfrak{o})$ is a group. By induction it follows that $r = s$ and the expression is unique up to the order of the factors, and this proves (b).

(b) \Rightarrow (c) is clear. To prove (c) \Rightarrow (a), we first show that any invertible prime ideal \mathfrak{p} of \mathfrak{o} is maximal. If \mathfrak{p} is not maximal, then there exists $a \in \mathfrak{o}$ such that $\mathfrak{p} \subset \mathfrak{p} + (a) \subset \mathfrak{o}$, and by (c) we can write $\mathfrak{p} + (a) = \mathfrak{p}_1 \cdots \mathfrak{p}_r$, $\mathfrak{p} + (a^2) = \mathfrak{q}_1 \cdots \mathfrak{q}_s$, where $\mathfrak{p}_i, \mathfrak{q}_j$ are prime ideals containing \mathfrak{p} . In $\bar{\mathfrak{o}} = \mathfrak{o}/\mathfrak{p}$ we then have $(\bar{a}) = \bar{\mathfrak{p}}_1 \cdots \bar{\mathfrak{p}}_r$, $(\bar{a})^2 = \bar{\mathfrak{q}}_1 \cdots \bar{\mathfrak{q}}_s = \bar{\mathfrak{p}}_1^2 \cdots \bar{\mathfrak{p}}_r^2$. Let \bar{n} be minimal among the $\bar{\mathfrak{p}}_i, \bar{\mathfrak{q}}_j$, say $\bar{n} = \bar{\mathfrak{p}}_1$. Since $\bar{\mathfrak{p}}_1 \supseteq \bar{\mathfrak{q}}_j$ for some j , we find that $\bar{\mathfrak{p}}_1 = \bar{\mathfrak{q}}_j$ and so $\mathfrak{p}_1 = \mathfrak{q}_j$. Further, $\bar{\mathfrak{p}}_1$ is invertible as factor of the invertible ideal (\bar{a}) and so may be cancelled. By induction we find that $s = 2r$ and each $\bar{\mathfrak{p}}_i$ is equal to two of the $\bar{\mathfrak{q}}_j$, thus we have $(\mathfrak{p}_1 \cdots \mathfrak{p}_r)^2 = \mathfrak{q}_1 \cdots \mathfrak{q}_s$. It follows that

$$\mathfrak{p} \subseteq \mathfrak{p} + (a)^2 = [\mathfrak{p} + (a)]^2 \subseteq \mathfrak{p}^2 + (a);$$

since $a \notin \mathfrak{p}$, we have $\mathfrak{p} \cap (a) = \mathfrak{p}.(a)$, and so

$$\mathfrak{p} = \mathfrak{p} \cap (\mathfrak{p}^2 + (a)) = \mathfrak{p}^2 + (\mathfrak{p} \cap (a)) = \mathfrak{p}^2 + \mathfrak{p}.(a).$$

Since \mathfrak{p} is invertible, we can cancel it and find $\mathfrak{p} + (a) = \mathfrak{o}$, which is a contradiction. This shows \mathfrak{p} to be maximal.

Now let \mathfrak{m} be any non-zero prime ideal of \mathfrak{o} . If $0 \neq a \in \mathfrak{m}$, then $\mathfrak{m} \supseteq (a) = \mathfrak{p}_1 \cdots \mathfrak{p}_r$, hence $\mathfrak{m} \supseteq \mathfrak{p}_i$ for some i . But \mathfrak{p}_i as factor of (a) is invertible, hence maximal and so $\mathfrak{m} = \mathfrak{p}_i$, which shows \mathfrak{m} to be invertible. Thus every non-zero prime ideal of \mathfrak{o} is invertible, and so \mathfrak{o} is Dedekind, by Th. 5.4. ■

In spite of this unique factorization property Dedekind domains by no means include all UFDs, as is clear from Cor. 5.8 below.

COROLLARY 5.7 *A ring R is a principal ideal domain if and only if it is a UFD in which all non-zero prime ideals are maximal.*

Proof. The condition is clearly necessary. When it holds, we see from the remarks following Th. 2.6 that every minimal non-zero prime ideal in a UFD is principal. Hence every non-zero prime ideal of R is principal, so invertible, and by Th. 5.4, R is a Dedekind domain. By Th. 5.6, every non-zero ideal can be written as a product of (principal) prime ideals, and so is principal; thus R is a principal ideal domain. ■

As a further consequence we have

COROLLARY 5.8 *A Dedekind domain is a UFD if and only if it is a principal ideal domain.* ■

We can also use Th. 5.6 to describe the group of fractional ideals more closely:

PROPOSITION 5.9 *In a Dedekind domain \mathfrak{o} the set $I(\mathfrak{o})$ of fractional ideals is a free abelian group on the non-zero prime ideals.*

Proof. Let $\{\mathfrak{m}\}$ be the set of all non-zero prime ideals of \mathfrak{o} and take a fractional ideal $\mathfrak{A} \subseteq u\mathfrak{o}$. For some $v \neq 0$ we have $vu \in \mathfrak{o}$ and $v \in \mathfrak{o}$, hence $v\mathfrak{A} \subseteq \mathfrak{o}$ and we can write $v\mathfrak{A} = \prod \mathfrak{m}^{e_m}, (v) = \prod \mathfrak{m}^{f_m}$, where $e_m, f_m \geq 0$ and $\mathfrak{A} = \prod \mathfrak{m}^{e_m - f_m}$, hence the \mathfrak{m} 's generate $I(\mathfrak{o})$. If there is a relation between them, we may write this as $\prod \mathfrak{m}^{a_m} = \prod \mathfrak{m}^{b_m}$, where $a_m, b_m \geq 0$; by Th. 5.6, $a_m = b_m$, hence the \mathfrak{m} are indeed free generators of $I(\mathfrak{o})$. ■

We can now give a precise description of the valuations on a Dedekind domain. Let \mathfrak{o} be a Dedekind domain, not a field, and K its field of fractions. For each maximal ideal \mathfrak{p} of \mathfrak{o} there is a mapping $v_p: K^\times \rightarrow \mathbb{Z}$ given by $v_p(u) = e_p$ if $(u) = \prod \mathfrak{m}^{e_m}$. It is easily verified that v_p is a valuation on \mathfrak{o} , called the \mathfrak{p} -adic valuation. Let us show that every non-trivial valuation v whose valuation ring contains \mathfrak{o} is of this form. Let V be the valuation ring of v and \mathfrak{m} its maximal ideal;

then $\mathfrak{p} = \mathfrak{m} \cap \mathfrak{o}$ is a prime ideal of \mathfrak{o} , and $\mathfrak{p} \neq 0$, because v is non-trivial. By Th. 5.4(b), \mathfrak{p}_v is a principal valuation ring, and $\mathfrak{o}_{\mathfrak{p}} \subseteq V \subset K$. By Th. 8.4.6, $\mathfrak{o}_{\mathfrak{p}}$ is a maximal valuation ring of K , so $\mathfrak{o}_{\mathfrak{p}} = V$ and it follows that v is equivalent to $v_{\mathfrak{p}}$.

With this information we can derive an approximation theorem which strengthens Th. 8.2.5.

THEOREM 5.10 (Strong approximation theorem) *Let \mathfrak{o} be a Dedekind domain with field of fractions K . Given any non-trivial inequivalent valuations v_1, \dots, v_r on \mathfrak{o} , $n_1, \dots, n_r \in \mathbb{Z}$ and $x_1, \dots, x_r \in K$, there exists $x \in K$ such that $v_i(x - x_i) \geq n_i$ for $i = 1, \dots, r$ and $v(x) \geq 0$ for any valuation v on \mathfrak{o} not equivalent to any v_i .*

Proof. Without loss of generality we may assume that $n_i \geq 0$ ($i = 1, \dots, r$). By the remark preceding the theorem we know that $v_i = v_{\mathfrak{p}_i}$ for non-zero prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ of \mathfrak{o} . Let us write $x_i = y_i s^{-1}$, where $y_i, s \in \mathfrak{o}$. We have $(s) = \prod \mathfrak{p}^{v_{\mathfrak{p}}(s)}$, so $v_{\mathfrak{p}}(s) = 0$ for almost all \mathfrak{p} , and by increasing r we may assume that $v_{\mathfrak{p}}(s) = 0$ for $v_{\mathfrak{p}} \neq v_1, \dots, v_r$. We put $m_i = n_i + v_i(s)$; the $\mathfrak{p}_i^{m_i}$ are clearly pairwise comaximal, so by Th. 5.2.2, there exists $y \in \mathfrak{o}$ such that $y \equiv y_i \pmod{\mathfrak{p}_i^{m_i}}$. Now it follows that $x = ys^{-1}$ has all the desired properties. ■

COROLLARY 5.11 *If in Th. 5.10 the v_i are normalized valuations, then there exists $x \in K$ such that $v_i(x) = n_i$ and $v(x) \geq 0$ for all other valuations v .*

For we can find $x_i \in K$ such that $v_i(x_i) = n_i$; by the theorem there exists $x \in K$ such that $v_i(x - x_i) > n_i$, hence $v_i(x) = v_i(x_i) = n_i$. ■

Let \mathfrak{o} be a Dedekind domain with field of fractions K . The group $I(\mathfrak{o})$ has a subgroup consisting of the principal ideals, and the corresponding quotient group is called the *ideal class group* of \mathfrak{o} , written $C(\mathfrak{o})$. The mapping which associates with each element a of K^\times the fractional ideal (a) gives rise to an exact sequence

$$1 \rightarrow U \rightarrow K^\times \rightarrow I(\mathfrak{o}) \rightarrow C(\mathfrak{o}) \rightarrow 1,$$

where U denotes the group of units of \mathfrak{o} . It is clear that \mathfrak{o} is a principal ideal domain precisely when $C(\mathfrak{o}) = 1$, so this group measures the departure from being principal. It is a remarkable fact, proved by Kummer in 1847, that the ring of integers in an algebraic number field has a *finite* ideal class group (cf. e.g. Lang 1970). By contrast, the class groups of general Dedekind domains include all abelian groups (cf. Fossum 1973). The structure of U was determined by Dirichlet, who showed that for an algebraic number field with r_1 real and $2r_2$ complex conjugates, U is the direct product of a finite cyclic group and a free abelian group of rank $r_1 + r_2 - 1$.

If L is a finite extension of K and \mathfrak{O} is the integral closure of \mathfrak{o} in L , then by Th. 5.5 and Ex. (8), \mathfrak{O} is again a Dedekind domain. For any non-zero prime ideal

\mathfrak{p} in \mathfrak{o} we can write $\mathfrak{p}\mathfrak{O} = \mathfrak{P}_1^{e_1} \dots \mathfrak{P}_r^{e_r}$, $e_i > 0$. Thus \mathfrak{p} splits into a finite number of prime ideals in \mathfrak{O} ; we indicate this by writing $\mathfrak{P}_i|\mathfrak{p}$. If $e_i > 1$, \mathfrak{p} is said to be *ramified* at \mathfrak{P}_i ; in fact e_i is the ramification index of the extension of the \mathfrak{p} -adic valuation on K to the \mathfrak{P}_i -adic valuation on L . The homomorphism from $I(\mathfrak{o})$ to $I(\mathfrak{O})$ defined by $a \mapsto a\mathfrak{O}$ is called the *conorm mapping*; in terms of the free generators it sends \mathfrak{p} to $\prod \mathfrak{P}_i^{e_i}$. Further it maps principal ideals to principal ideals, and so induces a homomorphism $C(\mathfrak{o}) \rightarrow C(\mathfrak{O})$, which may be summed up in the commutative diagram

$$\begin{array}{ccccccc} 1 & \longrightarrow & U_{\mathfrak{o}} & \longrightarrow & K^{\times} & \longrightarrow & I(\mathfrak{o}) \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \longrightarrow & U_{\mathfrak{O}} & \longrightarrow & L^{\times} & \longrightarrow & I(\mathfrak{O}) \\ & & & & & & \downarrow \\ & & & & & & C(\mathfrak{o}) \\ & & & & & & \downarrow \\ & & & & & & 1 \end{array}$$

Exercises

- (1) Show that if $ab = (c) \neq 0$ is principal (in an integral domain), then a is invertible, and express its inverse in terms of b and c .
- (2) Show that in the ring $k[x, y]$ the ideal (x, y) is not invertible. (*Hint.* Find $(\mathfrak{o}:(x, y))$.)
- (3) Let d be a square-free integer and let \mathfrak{O} be the ring of integers in $\mathbf{Q}(\sqrt{d})$, i.e. the integral closure of \mathbf{Z} in $\mathbf{Q}(\sqrt{d})$. Show that if $d \equiv 2, 3 \pmod{4}$, \mathfrak{O} has a \mathbf{Z} -basis $1, \sqrt{d}$, and if $d \equiv 1 \pmod{4}$, \mathfrak{O} has a \mathbf{Z} -basis $1, (1 + \sqrt{d})/2$. Determine the ramified primes in each case.
- (4) Show that for an odd prime p , the p -adic valuation on \mathbf{Q} has two extensions to $\mathbf{Q}(\sqrt{-1})$ if $p \equiv 1 \pmod{4}$ and one if $p \equiv 3 \pmod{4}$. Show that all such extensions are unramified. By decomposing p into prime ideals over $\mathbf{Z}[\sqrt{-1}]$ show that p can be written as a sum of two squares iff $p \equiv 1 \pmod{4}$ or $p \equiv 2 \pmod{4}$ (Fermat). (*Hint.* Recall that $\mathbf{U}(p)$ is a cyclic group.)
- (5) Show that the group of units in the ring of integers in $\mathbf{Q}(\sqrt{-d})$, for a square-free positive integer d is $\mathbf{C}_2 = \{\pm 1\}$ except when $d = 1$ or 3 , and determine the group in these cases.
- (6) Show that the ring of integers in $\mathbf{Q}(\sqrt{-d})$ is Euclidean with respect to the usual norm, for $d = 2, 3, 7, 11$.
- (7) Show that the ring \mathfrak{o} of integers in $\mathbf{Q}(\sqrt{-d})$ is not Euclidean for any function φ (satisfying $\varphi(ab) \geq \varphi(a)$ for $a, b \neq 0$), when $d = 5$ or $d > 11$ and d is square-free. (*Hint.* Take γ not zero or a unit, with minimal $\varphi(\gamma)$ and show that $|\mathfrak{o}/\gamma\mathfrak{o}| = N(\gamma)$; deduce that γ must be a unit.)
- (8) Let \mathfrak{o} be a Dedekind domain with field of fractions K , let L be a finite extension of K and let \mathfrak{O} be the integral closure of \mathfrak{o} in L . Prove that \mathfrak{O} is a Dedekind domain. (*Hint.* By

Th. 5.5 reduce to the purely inseparable case, say $L = K^{1/q}$. Since $K \cong K^q$, the image $\mathfrak{o}^{1/q}$ of \mathfrak{o} is a Dedekind domain such that $\mathfrak{o}^{1/q}L = \mathfrak{D}$. If \mathfrak{A} is a fractional ideal in \mathfrak{D} , then we have $\sum a_i b_i = 1$, where $a_i \in \mathfrak{A}$, $b_i \in (\mathfrak{o}^{1/q} \cdot \mathfrak{A})$, because $\mathfrak{o}^{1/q}$ is Dedekind. Now raise this equation to the q th power.)

(9) For any field k of characteristic not 2 show that the k -algebra generated by x, y with the defining relation $x^2 + y^2 = 1$ is a Dedekind domain, but the k -algebra generated by x, y, z with the defining relation $x^2 + y^2 + z^2 = 1$ is not. Find the class group of $k[x, y]/(x^2 + y^2 - 1)$.

(10) Let \mathfrak{o} be a Dedekind domain with field of fractions K . Show that any ring between \mathfrak{o} and K is a Dedekind domain.

(11) Let K be a field and F a family of principal valuations such that (i) for any $x \in K$, $v(x) \geq 0$ for almost all $v \in F$ and (ii) given $v, v' \in F$ and $N > 0$, there exists $\alpha \in K$ such that $v(\alpha - 1) > N$, $v'(\alpha) > N$, and $w(x) \geq 0$ for all $w \neq v, v'$. If R is the intersection of the corresponding valuation rings, show how to describe a fractional ideal in terms of the valuations in F . Deduce that the fractional ideals form a group, and hence R is a Dedekind domain.

9.6 Modules over Dedekind domains

We have seen that an integral domain is Dedekind iff it is hereditary. This also leads to a convenient description of projective modules.

PROPOSITION 6.1 *Every finitely generated projective module over a Dedekind domain can be written as a direct sum of modules isomorphic to ideals.*

We remark that the ideals can be taken to be integral since every fractional ideal is isomorphic to an integral ideal.

Proof. Let M be a finitely generated projective \mathfrak{o} -module; M is a submodule of a free module which may be taken to have finite rank, say $M \subseteq \mathfrak{o}^r$. We project on the last factor \mathfrak{o} and denote the image, an ideal of \mathfrak{o} , by \mathfrak{a} . Thus we obtain an exact sequence

$$0 \rightarrow M' \rightarrow M \xrightarrow{\lambda} \mathfrak{a} \rightarrow 0, \quad (1)$$

where $M' = \ker \lambda = M \cap \mathfrak{o}^{r-1}$. By induction on r , M' is isomorphic to a direct sum of ideals, and since \mathfrak{a} is projective, the sequence (1) splits and we obtain the desired decomposition for M . ■

We note that the proof shows every submodule of a free module of finite rank to be projective; this actually holds for all submodules of free modules, and it accounts for the name ‘hereditary’.

To study modules over Dedekind domains more closely we need a reduction theorem, which allows us in many cases to replace the ring by a principal ideal domain (PID).

PROPOSITION 6.2 *A Dedekind domain with only finitely many prime ideals is principal.*

Proof. Let \mathfrak{o} be a Dedekind domain whose non-zero prime ideals are $\mathfrak{p}_1, \dots, \mathfrak{p}_r$, and denote by v_i the valuation associated with \mathfrak{p}_i . Given any integers $n_1, \dots, n_r \geq 0$, there exists $a \in \mathfrak{o}$ such that $v_i(a) = n_i$, by the strong approximation theorem (Th. 5.10), and clearly

$$(a) = \mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_r^{n_r};$$

hence every ideal of \mathfrak{o} is principal. ■

COROLLARY 6.3 *Let \mathfrak{o} be a Dedekind domain and \mathfrak{a} an integral ideal of \mathfrak{o} . Then $\mathfrak{o}/\mathfrak{a}$ is a principal ideal ring.*

Proof. Write $\mathfrak{a} = \mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_r^{n_r}$, where the \mathfrak{p}_i are the distinct prime ideals containing \mathfrak{a} . Let S be the set of elements which become units mod \mathfrak{a} , i.e. elements x such that $\mathfrak{a} + (x) = \mathfrak{o}$. Then S is multiplicative and \mathfrak{o}_S is a domain whose group of fractional ideals is freely generated by $\mathfrak{p}_1, \dots, \mathfrak{p}_r$. Hence it is a Dedekind domain with finitely many prime ideals, and so is a PID, by Prop. 6.2. Now the natural homomorphism $\mathfrak{o} \rightarrow \mathfrak{o}/\mathfrak{a}$ is S -inverting, and hence can be taken via \mathfrak{o}_S , and as homomorphic image of \mathfrak{o}_S , $\mathfrak{o}/\mathfrak{a}$ is principal. ■

Of course $\mathfrak{o}/\mathfrak{a}$ will not in general be a domain; in fact it has zero divisors unless it is a field (the case $r = 1, n_1 = 1$).

COROLLARY 6.4 *Every integral ideal \mathfrak{a} of a Dedekind domain \mathfrak{o} can be generated by two elements, one of which may be chosen arbitrarily as non-zero element of \mathfrak{a} . Moreover, this element may be chosen prime to a given element, itself prime to \mathfrak{a} .*

Proof. Take $0 \neq a \in \mathfrak{a}$; then $\mathfrak{a}/(a)$ is an ideal of $\mathfrak{o}/\mathfrak{a}$ and therefore principal; if $\mathfrak{a} \equiv (b) \pmod{(a)}$, then $\mathfrak{a} = (a, b)$. Moreover, given $c \in \mathfrak{o}$, prime to \mathfrak{a} , we have $cu + a = 1$ for some $u \in \mathfrak{o}$, $a \in \mathfrak{a}$ and we can start with this a . ■

COROLLARY 6.5 *Let $\mathfrak{a}, \mathfrak{b}$ be fractional ideals in a Dedekind domain \mathfrak{o} with field of fractions K , and suppose that \mathfrak{b} is integral. Then there exists $u \in K^\times$ such that*

$$u\mathfrak{a} + \mathfrak{b} = \mathfrak{o}. \quad (2)$$

Proof. Choose $c \in K^\times$ such that $c\mathfrak{a}$ integral. Then $c\mathfrak{a}/c\mathfrak{b}$ is an ideal in $\mathfrak{o}/c\mathfrak{b}$,

and hence is principal: $ca = (v) + cab$. On multiplying by $(ca)^{-1}$ we obtain $(vc^{-1})a^{-1} + b = 0$, and now (2) follows if we put $u = vc^{-1}$ and replace a by a^{-1} . ■

We recall from Vol. 1, p. 326, that an R -module M , for any integral domain R , is called a *torsion module* if for each $x \in M$ there exists $r \in R^\times$ such that $xr = 0$; if $xr = 0$ for $x \in M, r \in R$ implies $x = 0$ or $r = 0$, M is called *torsion-free*. We shall find that the theory of finitely generated modules over Dedekind domains runs parallel to the theory for principal ideal domains, with projective modules taking the place of free modules. Over a PID every finitely generated torsion-free module is free (Th. 2, Cor. of 10.6, p. 328, Vol. 1) and here we have a precise analogue.

PROPOSITION 6.6 *A finitely generated module over a Dedekind domain is torsion-free if and only if it is projective.*

Proof. Let \mathfrak{o} be the ring and K its field of fractions. If M is projective, then it is a submodule of a free module and hence torsion-free. Conversely, let M be finitely generated torsion-free. Then the natural mapping

$$M \rightarrow M \otimes_{\mathfrak{o}} K \tag{3}$$

is an embedding, and $M \otimes K$ is a finitely generated K -module, hence isomorphic to K^r , for some $r \geq 0$. Let M be generated by u_1, \dots, u_n and let c be a common denominator for the expressions of the images of the u_i in the embedding (3). Then $cM \subseteq \mathfrak{o}'$, hence M is isomorphic to a submodule of a free module; since \mathfrak{o} is hereditary, it follows that M is projective (cf. the remark after Prop. 6.1). ■

For any module M over an integral domain R , the set $tM = \{x \in M \mid xr = 0 \text{ for some } r \in R^\times\}$ is easily seen to be a submodule; moreover, it is a torsion module, the quotient M/tM is torsion-free, and these properties determine tM uniquely. It is called the *torsion submodule* of M .

COROLLARY 6.7 *Let M be a finitely generated module over a Dedekind domain \mathfrak{o} . Then $M = tM \oplus P$, where P is a torsion-free submodule of M .*

For we have the exact sequence

$$0 \rightarrow tM \rightarrow M \rightarrow M/tM \rightarrow 0;$$

here M/tM is torsion-free, and finitely generated because M is; hence it is projective and so the sequence splits: $M = tM \oplus P$, where $P \cong M/tM$. ■

This result shows that we may consider the torsion part and the torsion-free part separately; we begin with the torsion part. Thus let M be a finitely generated torsion module. If u_1, \dots, u_n is a generating set and \mathfrak{n}_i is the annihilator of

$u_i (i = 1, \dots, n)$, then $\mathfrak{n} = \mathfrak{n}_1 \cdots \mathfrak{n}_n$ annihilates every element of M . Let S be the set of elements of \mathfrak{o} prime to \mathfrak{n} ; then for any $s \in S$ there exists $t \in \mathfrak{o}$ and $a \in \mathfrak{n}$ such that $st + a = 1$, hence

$$xst = x \quad \text{for all } x \in M.$$

Thus the action of s on M is an automorphism with inverse t . It follows that the natural mapping

$$M \rightarrow M \otimes_{\mathfrak{o}} \mathfrak{o}_S \tag{4}$$

is an isomorphism of abelian groups. More precisely, (4) provides a natural transformation from \mathfrak{o} -modules to \mathfrak{o}_S -modules, which for modules annihilated by \mathfrak{n} is an isomorphism. Now by Prop. 6.2, \mathfrak{o}_S is a PID, so we can apply the basis theorem for modules over PIDs (Th. 2 of **10.6**, p. 327 of Vol. 1) and obtain

$$M \otimes_{\mathfrak{o}} \mathfrak{o}_S \cong \mathfrak{o}_S/(a_1) \oplus \cdots \oplus \mathfrak{o}_S/(a_r),$$

where $a_i | a_{i+1}$, and this decomposition is unique up to isomorphism. It follows that M admits a corresponding decomposition into cyclic \mathfrak{o} -modules, with ideals of \mathfrak{o} corresponding to the a_i . Thus we obtain

THEOREM 6.8 *Any finitely generated torsion module M over a Dedekind domain \mathfrak{o} is a direct sum of cyclic \mathfrak{o} -modules:*

$$M \cong \mathfrak{o}/\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{o}/\mathfrak{a}_r, \quad \mathfrak{a}_i \supseteq \mathfrak{a}_{i+1}, \tag{5}$$

and this decomposition is unique up to isomorphism. ■

As in the principal case we can, for any torsion module (indeed for any module), define for each non-zero prime ideal \mathfrak{p} , the \mathfrak{p} -primary part

$$M_{\mathfrak{p}} = \{x \in M \mid x\mathfrak{p}^n = 0 \text{ for some } n \geq 1\},$$

and in the same way as for PIDs (Th. 3, **10.6**, p. 328 of Vol. 1) prove

PROPOSITION 6.9 *Any torsion module M over a Dedekind domain is a direct sum of its primary parts, in a unique way:*

$$M = \bigoplus M_{\mathfrak{p}},$$

and when M is finitely generated, only finitely many terms on the right are different from zero. ■

It remains to classify the torsion-free modules. This is done in the next theorem, which is essentially due to Steinitz (1912). Before stating it we note a lemma which is needed in the proof.

LEMMA 6.10 *Let R be an integral domain with field of fractions K . If $\mathfrak{a}, \mathfrak{b}$ are*

fractional ideals of R and

$$f: \mathfrak{a} \rightarrow \mathfrak{b}$$

is a homomorphism of R -modules, then there exists $c \in K$ such that $xf = cx$. In particular, f is either zero or injective.

Proof. For $a, x \in \mathfrak{a}$ we have

$$af \cdot x = (ax)f = xf \cdot a.$$

Fix $a \neq 0$ and write $c = af \cdot a^{-1}$; then $c \in K$ and $xf = cx$. ■

In particular we see that $\mathfrak{a} \cong \mathfrak{b}$ iff $\mathfrak{a} = c\mathfrak{b}$ for some $c \in K^\times$.

THEOREM 6.11 *Let \mathfrak{o} be a Dedekind domain. Any finitely generated torsion-free \mathfrak{o} -module has the form $\mathfrak{o}^r \oplus \mathfrak{a}$, where \mathfrak{a} is an ideal, and*

$$\mathfrak{o}^r \oplus \mathfrak{a} \cong \mathfrak{o}^s \oplus \mathfrak{b} \Leftrightarrow r = s \text{ and } \mathfrak{a} \cong \mathfrak{b}.$$

More explicitly, if $\mathfrak{a}_1, \dots, \mathfrak{a}_r, \mathfrak{b}_1, \dots, \mathfrak{b}_s$ are any fractional ideals, then

$$\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r \cong \mathfrak{b}_1 \oplus \cdots \oplus \mathfrak{b}_s \Leftrightarrow r = s \text{ and } \mathfrak{a}_1 \cdots \mathfrak{a}_r \cong \mathfrak{b}_1 \cdots \mathfrak{b}_s. \quad (6)$$

Proof. By Props. 6.6 and 6.1, any finitely generated torsion-free module is a direct sum of ideals. Let us first prove the necessity of the conditions in (6). In Lemma 6.10 we have seen that any homomorphism $\mathfrak{a} \rightarrow \mathfrak{b}$ is obtained by multiplying by an element of K , hence any homomorphism

$$\gamma: \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r \rightarrow \mathfrak{b}_1 \oplus \cdots \oplus \mathfrak{b}_s$$

is obtained by multiplying by an $s \times r$ matrix over K : $\mathfrak{b}_i = \sum c_{ij} \mathfrak{a}_j$. Clearly γ is an isomorphism iff the matrix $C = (c_{ij})$ has an inverse; in particular we must then have $r = s$. We claim further that in this case

$$\mathfrak{b}_1 \cdots \mathfrak{b}_r = (\det C) \mathfrak{a}_1 \cdots \mathfrak{a}_r. \quad (7)$$

For, given $a_j \in \mathfrak{a}_j$, we have $c_{ij} a_j \in \mathfrak{b}_i$, hence $(\det C) \mathfrak{a}_1 \cdots \mathfrak{a}_r \subseteq \mathfrak{b}_1 \cdots \mathfrak{b}_r$. By symmetry we have the reverse inclusion, and hence equality in (7). This proves the necessity of the conditions (in either formulation); we observe that it holds in any integral domain, not necessarily Dedekind.

To establish the converse we shall show that $\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r \cong \mathfrak{o}^{r-1} \oplus \mathfrak{a}_1 \cdots \mathfrak{a}_r$; clearly the desired conclusion follows from this. Considered first the case $r = 2$:

$$\mathfrak{a} \oplus \mathfrak{b} \cong \mathfrak{o} \oplus \mathfrak{ab}. \quad (8)$$

On multiplying \mathfrak{b} by an element of K we may assume that \mathfrak{b} is integral, and then by Cor. 6.5 we may further assume that $\mathfrak{a} + \mathfrak{b} = \mathfrak{o}$. Then we have an exact sequence

$$0 \rightarrow \ker \lambda \rightarrow \mathfrak{a} \oplus \mathfrak{b} \xrightarrow{\lambda} \mathfrak{o} \rightarrow 0, \quad (9)$$

where $\lambda:(x, y) \mapsto x - y$ and $\ker \lambda = \mathfrak{a} \cap \mathfrak{b} = \mathfrak{ab}$, because $\mathfrak{a}, \mathfrak{b}$ are comaximal. Now (9) splits and so we obtain (8). It follows that

$$\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r \cong \mathfrak{o} \oplus \mathfrak{a}_1 \mathfrak{a}_2 \oplus \mathfrak{a}_3 \oplus \cdots \oplus \mathfrak{a}_r \cong \mathfrak{o}^{r-1} \oplus \mathfrak{a}_1 \cdots \mathfrak{a}_r,$$

by induction on r , which is what we had to show. ■

The result may be summed up by saying that a complete set of invariants for the projective module $\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r$, consists of the integer r and the isomorphism class of the fractional ideal $\mathfrak{a}_1 \cdots \mathfrak{a}_r$.

Exercises

- (1) Show that every projective module over a Dedekind domain, which is countably but not finitely generated is free. (This actually holds without the countability restriction, by a theorem of Kaplansky (1958).)
- (2) Show that if the definition of torsion-free module in the text is applied to general rings, then a ring which is not an integral domain has no non-zero torsion-free modules. (Over a general ring R a module is sometimes called torsion-free if multiplication by any non-zerodivisor of R is injective, but we shall have no need to use this definition.)
- (3) Show that any module over a Dedekind domain has the form $M = I \oplus E$, where I is injective and E has no non-zero injective submodules.
- (4) Let \mathfrak{o} be a Dedekind domain and K its field of fractions. Show that for any prime ideal $\mathfrak{p} \neq 0$ the \mathfrak{p} -primary part of K/\mathfrak{o} is indecomposable injective. It is called a module of type \mathfrak{p}^∞ .
Show that every divisible \mathfrak{o} -module is a direct sum of a K -space and of modules of type \mathfrak{p}^∞ , for different \mathfrak{p} . Deduce that every divisible \mathfrak{o} -module is injective.
- (5) Show that a finitely generated indecomposable module over a Dedekind domain \mathfrak{o} is either torsion-free or of type $\mathfrak{o}/\mathfrak{p}^n$.

9.7 Algebraic equations

Algebraic geometry may be defined as the study of the solutions of polynomial equations over fields. An equation $f(x) = 0$ in one variable has a finite number of roots, but the solutions of an equation in several variables form an algebraic variety, e.g. the equation $x^2 + y^2 = 1$ defines a circle in the (x, y) -plane. Our task will be to explain how rings arise in this context.

Let K be a commutative ring and consider a system of polynomial equations over K , in the indeterminates x_1, \dots, x_n . This means that we have a family of equations

$$E: \quad f_\lambda(x_1, \dots, x_n) = 0 \quad (\lambda \in I), \tag{1}$$

where the left-hand sides are members of the polynomial ring $K[x_1, \dots, x_n]$. Let L be a K -algebra; by a *solution* of the system E in L we understand an n -tuple (a_1, \dots, a_n) of elements of L such that $f_\lambda(a_1, \dots, a_n) = 0$ for all $\lambda \in I$. The set of all such solutions is denoted by $V_L(E)$.

Two systems E and F in the same variables over K are said to be *equivalent* if $V_L(E) = V_L(F)$ for all K -algebras L . Among all systems equivalent to E there is a maximal one, namely the ideal generated by E in the polynomial ring:

THEOREM 7.1 *Let E be a system of equations in x_1, \dots, x_n over K , denote by a the ideal of $K[x_1, \dots, x_n]$ generated by the left-hand sides of E and put $A = K[x_1, \dots, x_n]/a$. Then each solution of E in a K -algebra L may be described by a K -algebra homomorphism from A to L and the mapping so obtained is a natural bijection:*

$$V_L(E) \cong \text{Hom}_K(A, L). \quad (2)$$

A is called the *function ring* or *coordinate ring* of the system E .

The proof consists essentially in verifying the universal property of A : any homomorphism $A \rightarrow L$ gives rise to a homomorphism $\varphi: K[x_1, \dots, x_n] \rightarrow L$ mapping E to 0. If φ maps

$$x_i \mapsto \xi_i, \quad (3)$$

then (ξ_1, \dots, ξ_n) is the corresponding solution. Conversely, let (ξ_1, \dots, ξ_n) be a solution of the system E in L and let φ be the unique K -algebra homomorphism from $K[x_1, \dots, x_n]$ to L defined by (3); then $E \subseteq \ker \varphi$, because (ξ_1, \dots, ξ_n) was a solution of E . Hence φ can be factorized by the natural homomorphism $K[x_1, \dots, x_n] \rightarrow A$ to give a homomorphism $A \rightarrow L$, and these two constructions are evidently mutually inverse. ■

A system E is said to be *consistent* if it has a solution in some non-trivial K -algebra, i.e. $V_L(E) \neq \emptyset$ for some $L \neq 0$; otherwise E is *inconsistent*. From Th. 7.1 we obtain

COROLLARY 7.2 *A system E of equations over K is consistent if and only if the ideal generated by the left-hand sides is proper in $K[x_1, \dots, x_n]$.*

For the ideal in question is proper iff the algebra A in Th. 7.1 is non-trivial. Now assume $V_L(E) \neq \emptyset$; then by Th. 7.1, $\text{Hom}(A, L) \neq \emptyset$ for some non-trivial L , hence $A \neq 0$. Conversely, when $A \neq 0$, then $V_A(E) = \text{Hom}(A, A) \neq \emptyset$, since there is always the identity mapping. ■

We can apply the Hilbert basis theorem and obtain

PROPOSITION 7.3 *Let K be a Noetherian ring. Then every system of equations in x_1, \dots, x_n over K is equivalent to a finite system.*

Proof. By the Hilbert basis theorem (Th. 4.2) $R = K[x_1, \dots, x_n]$ is Noetherian. Let E be any subset of R and \mathfrak{a} the ideal generated by E . As a system of equations, E is equivalent to \mathfrak{a} , and since R is Noetherian, we can write $\mathfrak{a} = (f_1, \dots, f_r)$, so \mathfrak{a} is equivalent to the finite system f_1, \dots, f_r . ■

Th. 7.1 leads to a correspondence between systems of equations and function rings which have their counterparts in the geometrical objects formed by the solution sets. Thus let k be a field and E a system of equations in x_1, \dots, x_n over k . If L is a field containing k , then $V_L(E)$, the set of solutions of E in L , is a subset of L^n . When k is a subfield of the real numbers and $n = 2$ or 3 , we thus obtain a representation of E in the plane or in space, and the geometric language suggested by this example is used even when $n > 3$ and the coordinates lie in a general field.

To give an illustration, the equation

$$x^2 + y^2 = 1 \quad (4)$$

defines a circle in the plane and its function ring is generated by x, y over k , subject to the relation (4). This function ring is Noetherian, by Prop. 4.4, but not a UFD, for x^2 has the two factorizations $x \cdot x = (1 - y)(1 + y)$, and it is not hard to show (e.g. using the norm over $k(x)$) that $x, 1 + y, 1 - y$ are atoms in the ring, but not primes.

The circle in space corresponding to (4) is given by the system

$$x^2 + y^2 = 1, \quad z = 0.$$

Clearly it has the same function ring as the circle in the plane. The geometric object may also consist of several parts, e.g. the system

$$x(x^2 + y^2 + z^2 - 1) = y(x^2 + y^2 + z^2 - 1) = 0 \quad (5)$$

consists of a sphere and a line (the z -axis).

Let us consider more closely the correspondence between subsets of n -dimensional space and ideals in $R = k[x_1, \dots, x_n]$. For simplicity we shall work over k itself (rather than an extension), so that our space is k^n . To each subset E of R there corresponds the set $V_k(E)$, or simply $V(E)$, of solutions in k . A subset of k^n is said to be *closed* or an *algebraic set* if it has the form $V(E)$ for some $E \subseteq R$. Likewise with each subset S of k^n we associate the set $I(S)$ of all polynomials vanishing on S , and thus obtain a Galois connexion:

$$V(E) = \{p \in k^n \mid f(p) = 0 \text{ for all } f \in E\},$$

$$I(S) = \{f \in R \mid f(p) = 0 \text{ for all } p \in S\}.$$

It is clear that $I(S)$ is an ideal in R ; moreover, it coincides with its own radical, for if $f' \in I(S)$, then $f'(p) = 0$ for all $p \in S$ and so $f(p) = 0$, i.e. $f \in I(S)$. By Prop. 7.3, every algebraic set can be defined by a finite set of equations.

As in every Galois connexion we have the following rules:

- (i) $VI(S) \supseteq S$, $IV(\mathfrak{a}) \supseteq \mathfrak{a}$,
- (ii) $S_1 \subseteq S_2 \Rightarrow I(S_1) \supseteq I(S_2)$, $\mathfrak{a}_1 \subseteq \mathfrak{a}_2 \Rightarrow V(\mathfrak{a}_1) \supseteq V(\mathfrak{a}_2)$,
- (iii) $IVI(S) = I(S)$, $VIV(\mathfrak{a}) = V(\mathfrak{a})$.

Proof. (i) $f \in I(S)$ means that $f(p) = 0$ for all $p \in S$, hence $p \in S$ implies $p \in VI(S)$; similarly for the other inclusion.

(ii) If $f(p) = 0$ for all $p \in S_2$, then this holds for all $p \in S_1$, hence $I(S_1) \supseteq I(S_2)$; again the second relation follows similarly.

Now (iii) follows from (i) and (ii), as we have seen in 3.6. ■

What has been proved shows that the relation $f(p) = 0$ defines an order-inverting bijection between the subsets of k^n of the form $V(\mathfrak{a})$, the algebraic sets, and the subsets of R of the form $I(S)$. We have seen that each $I(S)$ is an ideal equal to its own radical; later we shall find, as a consequence of the Hilbert Nullstellensatz, in 9.10, that over an algebraically closed field, every ideal equal to its own radical is of this form.

We observe that the descending chains of closed sets in k^n correspond to ascending chains of ideals in $k[x_1, \dots, x_n]$, and since the latter ring is Noetherian, the ideal chains break off. It follows that k^n satisfies the descending chain condition on closed subsets. Of course the ascending chain condition will not generally hold, since e.g. every finite set is closed.

A closed set S in k^n is said to be *reducible* if $S = \emptyset$ or $S = S_1 \cup S_2$, where S_1, S_2 are proper closed subsets of S ; otherwise it is called *irreducible* or a *variety*. Thus the circle (4) is a variety, but the sphere- and-line (5) is reducible. However, even a variety may consist of several pieces in real space (cf. Ex. (5)). There is a simple algebraic criterion for reducibility:

PROPOSITION 7.4 *A closed set S in k^n is irreducible if and only if $I(S)$ is a prime ideal in $k[x_1, \dots, x_n]$.*

Proof. Write $\mathfrak{a} = I(S)$ and suppose that \mathfrak{a} is not prime; if $S = \emptyset$, there is nothing to prove; otherwise \mathfrak{a} is proper and there exist $f_1, f_2 \notin \mathfrak{a}$ but $f_1 f_2 \in \mathfrak{a}$. Hence there exist $p_1, p_2 \in S$ such that $f_i(p_i) \neq 0$. Put $S_i = V(\mathfrak{a} + (f_i))$ ($i = 1, 2$); then S_i is closed and a proper subset of S . Moreover, $S_1 \cup S_2 = V((\mathfrak{a} + (f_1))(\mathfrak{a} + (f_2))) = V(\mathfrak{a}^2 + (f_1 f_2)) = S$, so S is reducible. Conversely, suppose that $S = S_1 \cup S_2$, where $S_i \subset S$ and S_i is closed, and write $\mathfrak{a}_i = I(S_i)$. Then $\mathfrak{a}_i \supset \mathfrak{a}$, hence there exists $f_i \in \mathfrak{a}_i \setminus \mathfrak{a}$. Now $f_1 f_2 \in \mathfrak{a}_1 \cap \mathfrak{a}_2 = \mathfrak{a}$, and this shows that \mathfrak{a} is not prime. ■

If V is a variety in k^n , then $I(V)$ is a prime ideal and hence there is an extension field F of k containing the function ring of V , viz. the field of fractions of $k[x_1, \dots, x_n]/I(V)$. If ξ_i is the image of x_i in F under the natural homomorphism,

then every point of V can be obtained by ‘specializing’ the point $(\xi_1, \dots, \xi_n) \in F^n$ (in a sense which needs to be made precise); this point is therefore called a *generic point* of V . Conversely, any algebraic set with a generic point is a variety, e.g. a generic point for the circle (4) is $((1-t^2)/(1+t^2), 2t/(1+t^2))$.

Exercises

- (1) Show that a system of equations is consistent iff it has a common solution over some field.
- (2) Show that every algebraic set in k^n can be written as a union of a finite number of irreducible sets. (*Hint.* Use the Hilbert basis theorem.)
- (3) Show that over an algebraically closed field k , the irreducible subsets of k^2 are points, curves and k^2 .
- (4) Find the function ring of the hyperbola $xy = 1$ (over a field of characteristic not 2), and compare it with that of a straight line.
- (5) Verify that the curve $y^2 = x^3 - x$ is irreducible, but consists of two pieces (in the real plane).
- (6) Find the function ring of the ‘semicubical parabola’ $y^2 = x^3$, and show that it is not integrally closed. Find the integral closure, and a curve of which it is the function ring.
- (7) Verify that $V(ab) = V(a) \cup V(b)$, $V(\bigcup a_\lambda) = \bigcap V(a_\lambda)$, $V(\{1\}) = \emptyset$, $V(\{0\}) = k^n$. (This means that the algebraic sets form the closed sets of a topology on k^n , known as the *Zariski topology*, cf. 9.10.)
- (8) Show that every covering by open sets of k^n has a finite subcovering. (This is known as *quasicompactness* of k^n .)

9.8 The primary decomposition

In \mathbf{Z} or, more generally, in any UFD, each non-zero element has a factorization into primes:

$$a = up_1^{\alpha_1} \cdots p_r^{\alpha_r} \quad (u \text{ a unit, } p_i \text{ distinct primes, } \alpha_i > 0, r \geq 0). \quad (1)$$

In terms of ideals this can be written as

$$(a) = (p_1^{\alpha_1}) \cap \cdots \cap (p_r^{\alpha_r}). \quad (2)$$

Of course (2) is less precise than (1), but it has the advantage of holding for a much wider class of rings. We shall find that a decomposition of this form is true for any Noetherian ring. For $k[x_1, \dots, x_n]$ this was proved in 1905 by E. Lasker (also world chess champion from 1894 to 1921) and extended to general Noetherian

rings by E. Noether in 1921. Following Bourbaki we shall derive a decomposition for modules, which of course include ideals as a special case.

Let R be a commutative ring and M an R -module. A prime ideal \mathfrak{p} of R is said to be *associated* to M if it occurs as annihilator of an element: $\mathfrak{p} = \text{Ann}(u)$ for some $u \in M$. The set of all associated primes of M is denoted by $\text{Ass}(M)$ (the ‘assassinator’ of M). If $\mathfrak{p} = \text{Ann}(u)$, the mapping $x \mapsto ux$ of R into M shows that R/\mathfrak{p} is embedded in M , hence \mathfrak{p} is associated to M iff R/\mathfrak{p} is embedded in M . We also note that $\text{Ass}(R/\mathfrak{p}) = \{\mathfrak{p}\}$, since every non-zero element of R/\mathfrak{p} has annihilator \mathfrak{p} .

PROPOSITION 8.1 *Let M be any R -module; any maximal member of the set of annihilators of non-zero elements of M is a prime ideal.*

Proof. The annihilator of $u \neq 0$ is a proper ideal. Suppose that $\mathfrak{p} = \text{Ann}(u)$ is maximal and let $ab \in \mathfrak{p}$, $a \notin \mathfrak{p}$. Then $ua \neq 0$, $uab = 0$, hence $b \in \text{Ann}(ua) \supseteq \mathfrak{p}$. By the maximality of \mathfrak{p} we have equality, hence $b \in \mathfrak{p}$ and this shows \mathfrak{p} to be prime. ■

Of course there may be no such maximal elements (cf. Ex. (3)), but if R is Noetherian and $M \neq 0$, then we can be sure of such maximal annihilators, and they will be prime, by Prop. 8.1. Together with the obvious fact that $\text{Ass}(0) = \emptyset$, this proves

COROLLARY 8.2 *Let R be a Noetherian ring. Then for any R -module M , $\text{Ass}(M) = \emptyset$ if and only if $M = 0$.* ■

Let us call $a \in R$ a *zerodivisor on M* if it annihilates some non-zero element of M . Since every annihilator is contained in some maximal annihilator, we obtain

COROLLARY 8.3 *Let R be a Noetherian ring and M an R -module. Then the set of zerodivisors on M is $\bigcup \{\mathfrak{p} \mid \mathfrak{p} \in \text{Ass}(M)\}$.* ■

Let M be an R -module and M' a submodule; we have the formula

$$\text{Ass}(M') \subseteq \text{Ass}(M) \subseteq \text{Ass}(M') \cup \text{Ass}(M/M'). \quad (3)$$

The first inclusion is clear. To prove the second, let $\mathfrak{p} \in \text{Ass}(M)$ and suppose that $\mathfrak{p} = \text{Ann}(u)$. Either $uR \cap M' \neq 0$, and since any non-zero element of $uR \cong R/\mathfrak{p}$ has annihilator \mathfrak{p} , we see that $\mathfrak{p} \in \text{Ass}(M')$. Or $uR \cap M' = 0$, in which case the natural mapping $M \rightarrow M/M'$, restricted to uR , is injective, so uR is embedded in M/M' and $\mathfrak{p} \in \text{Ass}(M/M')$. ■

We remark that $\text{Ass}(M)$ is the set of prime ideals which *equal* some $\text{Ann}(x)$ ($x \in M$), while $\text{Supp}(M)$ is the set of prime ideals which *contain* some $\text{Ann}(x)$ ($x \in M$). A Noetherian module M has finite length iff $\text{Ass}(M)$, or equivalently, $\text{Supp}(M)$, consists entirely of maximal ideals.

We can now show that for a finitely generated module M over a Noetherian ring, $\text{Ass}(M)$ is finite. More precisely, we have

PROPOSITION 8.4 *Let R be a Noetherian ring and M a finitely generated R -module. Then there is a chain of submodules*

$$0 = M_0 \subset M_1 \subset \cdots \subset M_r = M, \quad (4)$$

such that $M_i/M_{i-1} \cong R/\mathfrak{p}_i$ for some prime ideal \mathfrak{p}_i of R , and

$$\text{Ass}(M) \subseteq \{\mathfrak{p}_1, \dots, \mathfrak{p}_r\}.$$

Proof. For $M = 0$ there is nothing to prove. Otherwise M has a submodule M_1 of the form R/\mathfrak{p} , by Cor. 8.2, and we can take $\mathfrak{p}_1 = \mathfrak{p}$. If we have found a chain $M_1 \subset M_2 \subset \cdots \subset M_r$, with $M_i/M_{i-1} \cong R/\mathfrak{p}_i$ and $M_r \neq M$, then M/M_r has a submodule of the form R/\mathfrak{p} , where $\mathfrak{p} \in \text{Ass}(M/M_r)$, say $M_{r+1}/M_r \cong R/\mathfrak{p}$. We can then put $\mathfrak{p}_{r+1} = \mathfrak{p}$ and by induction obtain a chain (4) of the required form. Now by another induction we have from (3),

$$\text{Ass}(M) \subseteq \text{Ass}(M_1/M_0) \cup \cdots \cup \text{Ass}(M_r/M_{r-1}) = \{\mathfrak{p}_1, \dots, \mathfrak{p}_r\}. \quad \blacksquare$$

Let M be an R -module; an element $a \in R$ is said to be *locally nilpotent* on M if for each $x \in M$ there exists $n = n(x)$ such that $xa^n = 0$. If M is finitely generated, by u_1, \dots, u_r , say, and $u_i a^{n_i} = 0$, then by taking $n = \max \{n_1, \dots, n_r\}$ we find that a^n annihilates M , so in this case the action of a on M is nilpotent.

When R is Noetherian, we observe that a is locally nilpotent on M iff $a \in \bigcap \{\mathfrak{p} \mid \mathfrak{p} \in \text{Ass}(M)\}$. For if a is locally nilpotent, then for any $\mathfrak{p} \in \text{Ass}(M)$, $a^n \in \mathfrak{p}$ for some n , hence $a \in \mathfrak{p}$. Conversely, if a is not locally nilpotent, then there exists $x \in M$ such that $xa^n \neq 0$ for all n . Among such x choose one, say x_1 , with maximal annihilator $\mathfrak{p} = \text{Ann}(x_1)$. We claim that \mathfrak{p} is prime. Clearly $\mathfrak{p} \neq R$; if $bc \in \mathfrak{p}$, then $x_1 bc = 0$, hence $c \in \text{Ann}(x_1 b) \supseteq \mathfrak{p}$. Either $x_1 ba^n \neq 0$ for all n ; then by maximality $c \in \mathfrak{p}$. Or $x_1 ba^n = 0$ for some n , in which case $b \in \text{Ann}(x_1 a^n) \supseteq \mathfrak{p}$, and $x_1 a^n \cdot a^m \neq 0$ for all m , so $b \in \mathfrak{p}$, again by maximality. This shows that \mathfrak{p} is prime. Clearly $\mathfrak{p} \in \text{Ass}(M)$ and $a \notin \mathfrak{p}$. This proves

PROPOSITION 8.5 *For a Noetherian ring R and any R -module M , an element a of R is locally nilpotent on M if and only if $a \in \bigcap \{\mathfrak{p} \mid \mathfrak{p} \in \text{Ass}(M)\}$. \blacksquare*

Let R be a Noetherian ring and M an R -module. A submodule Q of M is said to be *primary* (in M) if $\text{Ass}(M/Q)$ consists of a single element. If $\text{Ass}(M/Q) = \{\mathfrak{p}\}$, we also say that Q is \mathfrak{p} -primary. In particular, this defines primary ideals in R . Thus an ideal \mathfrak{q} in R is primary precisely if $\mathfrak{q} \neq R$ and every zero divisor in R/\mathfrak{q} is nilpotent; then $\sqrt{\mathfrak{q}} = \mathfrak{p}$ is a prime ideal and \mathfrak{q} is \mathfrak{p} -primary. For example, in \mathbf{Z} , the \mathfrak{p} -primary ideals are (p^r) , $r = 1, 2, \dots$, but in general the \mathfrak{p} -primary ideals need not be powers of \mathfrak{p} , e.g. in $k[x, y]$, (x, y^2) is (x, y) -primary, but $(x, y^2) \neq (x, y)^r$ for all r .

Neither is it true that each power of a prime ideal is necessarily primary (cf. Ex. (6)).

LEMMA 8.6 *Let R be a Noetherian ring and M an R -module. If Q_1, \dots, Q_r are submodules of M which are \mathfrak{p} -primary for the same \mathfrak{p} , then $Q_1 \cap \dots \cap Q_r$ is also \mathfrak{p} -primary.*

Proof. We have a natural homomorphism

$$M/(Q_1 \cap \dots \cap Q_r) \rightarrow M/Q_1 \oplus \dots \oplus M/Q_r, \quad (5)$$

obtained by composing the mappings from the left-hand side to M/Q_i . If the module on the right is written N , then $\text{Ass}(N) = \{\mathfrak{p}\}$ by (3), and since (5) is clearly injective, the same holds for the module on the left, hence $Q_1 \cap \dots \cap Q_r$ is \mathfrak{p} -primary, as claimed. ■

Our objective will be to obtain an embedding of our module into a direct sum of quotients by primary submodules:

$$M \rightarrow M/Q_1 \oplus \dots \oplus M/Q_r. \quad (6)$$

Such a mapping clearly exists for any primary submodules Q_1, \dots, Q_r , and (6) will be injective precisely when

$$Q_1 \cap \dots \cap Q_r = 0. \quad (7)$$

Such a representation of 0 as intersection of primary submodules is called a *primary decomposition* in M . To show that such decompositions exist let us define a submodule N of M to be *meet-reducible* if $N = M$ or $N = N_1 \cap N_2$, where $N_i \supset N$; otherwise N is *meet-irreducible*. By the decomposition lemma (Lemma 2.2.12), any submodule of a Noetherian module can be written as a finite intersection of meet-irreducible submodules. It remains to observe that meet-irreducible modules are primary:

LEMMA 8.7 *Let R be a Noetherian ring and M any R -module. Then any meet-irreducible submodule N of M is primary.*

Proof. We assume that N is not primary and show it to be meet-reducible. If N is not primary in M and $M \neq N$, then $\text{Ass}(M/N)$ contains at least two primes $\mathfrak{p}_1, \mathfrak{p}_2$ say; hence M has submodules $N_i \cong R/\mathfrak{p}_i$. Every element of $N_i \setminus N$ has annihilator \mathfrak{p}_i , hence $N_1 \cap N_2 = N$, and this shows N to be meet-reducible, as claimed. ■

Thus we find that every finitely generated module over a Noetherian ring has a primary decomposition (7). Of course the decomposition will not usually be unique, e.g. some terms in (6) might be redundant. To obtain some uniqueness we shall modify (6) as follows. Firstly, using Lemma 8.6, we collect together primary

submodules for the same prime ideal, so that if Q_i is \mathfrak{p}_i -primary, all the \mathfrak{p}_i are distinct. Secondly, we may suppose that

$$Q_i \not\supseteq \bigcap_{j \neq i} Q_j, \quad (8)$$

for if this were not so, we could omit Q_i from (8) without affecting the result. If the family $\{Q_1, \dots, Q_r\}$ satisfies these two conditions, the primary decomposition is called *irredundant*. For such decompositions we have the following first uniqueness theorem:

THEOREM 8.8 *Let R be a Noetherian ring and M a finitely generated R -module. Then there is a primary decomposition in M :*

$$0 = Q_1 \cap \cdots \cap Q_r, \quad \text{where } Q_i \text{ is } \mathfrak{p}_i\text{-primary,} \quad (9)$$

and (9) may be chosen irredundant, i.e. so that the \mathfrak{p}_i are distinct and (8) holds. When this is so, the \mathfrak{p}_i are uniquely determined as the members of $\text{Ass}(M)$.

Proof. We have seen that decompositions (9) always exist, and that we can modify such a decomposition so as to become irredundant. As a result we have the embedding (6) and it follows that $\text{Ass}(M) \subseteq \{\mathfrak{p}_1, \dots, \mathfrak{p}_r\}$. For $r \leq 1$ there is nothing to prove. If $r > 1$, put $N = Q_2 \cap \cdots \cap Q_r$; then $N \neq 0$ by irredundancy. We have $N = N/(Q_1 \cap N) \cong (Q_1 + N)/Q_1 \subseteq M/Q_1$, hence $\text{Ass}(N) \subseteq \text{Ass}(M/Q_1) = \{\mathfrak{p}_1\}$, so \mathfrak{p}_1 is associated with N and hence with M ; similarly for the other \mathfrak{p}_i . ■

For an ideal \mathfrak{a} in R we can by taking $M = R/\mathfrak{a}$ in Th. 8.8 obtain the usual primary decomposition for \mathfrak{a} , generalizing (2):

$$\mathfrak{a} = \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_r, \quad \text{where } \mathfrak{q}_i \text{ is } \mathfrak{p}_i\text{-primary.} \quad (10)$$

It follows that

$$\sqrt{\mathfrak{a}} = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r. \quad (11)$$

In particular, taking $\mathfrak{a} = 0$ and remembering the definition of the \mathfrak{p}_i as annihilators, we obtain the

COROLLARY 8.9 *In a Noetherian ring R the nilradical \mathfrak{N} can be written as intersection of finitely many prime ideals: $\mathfrak{N} = \bigcap_1^r \mathfrak{p}_i$, and then $\bigcup_1^r \mathfrak{p}_i$ is the set of all zero divisors in R .* ■

Theorem 8.8 does not tell us whether the primary components of \mathfrak{a} are unique, and in general the answer is ‘no’. Thus in the polynomial ring $k[x, y]$ we have

$$(x^2, xy) = (x) \cap (x, y)^2 = (x) \cap (x^2, y).$$

The primary component for (x) is the same in each decomposition, but not that

for (x, y) . Geometrically this ideal defines the y -axis and the origin with an ‘infinitesimal arrow’, represented by a nilpotent in the local ring. Without trying to make this idea more precise, we note that it was the non-maximal geometrical component which failed to be unique. To obtain a more precise uniqueness result, we shall need a lemma which is also useful elsewhere.

LEMMA 8.10 (i) *In any ring R , if \mathfrak{p} is a prime ideal and*

$$\mathfrak{p} \supseteq \bigcap_{i=1}^n \mathfrak{a}_i, \quad (12)$$

then $\mathfrak{p} \supseteq \mathfrak{a}_i$ for some i . Moreover, if $\mathfrak{p} = \bigcap \mathfrak{a}_i$, then $\mathfrak{p} = \mathfrak{a}_i$ for some i .

(ii) *In any ring R , if*

$$\mathfrak{a} \subseteq \bigcup_{i=1}^n \mathfrak{p}_i, \quad (13)$$

where each \mathfrak{p}_i is a prime ideal, then $\mathfrak{a} \subseteq \mathfrak{p}_i$ for some i .

Note that in (ii), $\bigcup \mathfrak{p}_i$ is not generally an ideal, and the result is false with $\sum \mathfrak{p}_i$ in place of $\bigcup \mathfrak{p}_i$.

Proof. (i) Suppose that $\mathfrak{p} \not\supseteq \mathfrak{a}_i$ for all i , and choose $x_i \in \mathfrak{a}_i \setminus \mathfrak{p}$. Then $x = x_1 \cdots x_n \in \bigcap \mathfrak{a}_i$ but $x \notin \mathfrak{p}$ because \mathfrak{p} is prime; this contradicts (12). If moreover, $\mathfrak{p} = \bigcap \mathfrak{a}_i$, then $\mathfrak{p} \subseteq \mathfrak{a}_i$ for all i , and by what has already been proved, equality must hold for some i .

(ii) We use induction on n , the case $n = 1$ being trivial. Let $\mathfrak{a} \not\subseteq \mathfrak{p}_i$ for all i ; by induction, for each i , $\mathfrak{a} \not\subseteq \bigcup_{j \neq i} \mathfrak{p}_j$, hence there exists $x_i \in \mathfrak{a}$, $x_i \notin \mathfrak{p}_j$ for all $j \neq i$, and so $x_i \in \mathfrak{p}_i$ by (13). Consider $y = x_1 + x_2 \cdots x_n$; by construction $x_i \in \mathfrak{a}$, so $y \in \mathfrak{a}$. But $x_1 \in \mathfrak{p}_1$, $x_2 \cdots x_n \notin \mathfrak{p}_1$, so $y \notin \mathfrak{p}_1$ and for $i > 1$, $x_i \notin \mathfrak{p}_i$ while $x_2 \cdots x_n \in \mathfrak{p}_i$, so again $y \notin \mathfrak{p}_i$. Hence $\mathfrak{a} \not\subseteq \bigcup \mathfrak{p}_i$, a contradiction, and the result follows. ■

For any ideal \mathfrak{a} of R , the minimal terms in $\text{Ass}(R/\mathfrak{a})$ are called the *minimal* or *isolated components* of \mathfrak{a} , while the other terms are said to be *embedded*. For example, in the above example, (x, y) is an embedded prime ideal and (x) is isolated. More generally, a subset Σ of $\text{Ass}(R/\mathfrak{a})$ is called *isolated* if any $\mathfrak{p} \in \text{Ass}(R/\mathfrak{a})$ such that $\mathfrak{p} \subseteq \mathfrak{p}' \in \Sigma$, satisfies $\mathfrak{p} \in \Sigma$; in other words, Σ is a lower segment in $\text{Ass}(R/\mathfrak{a})$.

We remark that the isolated primes of \mathfrak{a} are precisely the primes that are minimal among the primes containing \mathfrak{a} . For if \mathfrak{p} is any prime containing \mathfrak{a} , then $\mathfrak{p} \supseteq \mathfrak{a} = \bigcap \mathfrak{q}_i$, $\mathfrak{p} = \sqrt{\mathfrak{p}} \supseteq \sqrt{\bigcap \mathfrak{q}_i} = \bigcap \mathfrak{p}_i$, hence $\mathfrak{p} \supseteq \mathfrak{p}_i$ for some i , by Lemma 8.10.

We can now state the second uniqueness theorem.

THEOREM 8.11 *Let $\mathfrak{a} = \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_r$ be an irredundant primary decomposition of an ideal \mathfrak{a} in any ring R , and let $\sqrt{\mathfrak{q}_i} = \mathfrak{p}_i$ be the associated prime ideal. If*

$\{\mathfrak{p}_{i_1}, \dots, \mathfrak{p}_{i_t}\}$ is an isolated subset, then $\mathfrak{q}_{i_1} \cap \dots \cap \mathfrak{q}_{i_t}$ is independent of the choice of the decomposition for \mathfrak{a} . In particular, for any minimal component \mathfrak{p}_i , \mathfrak{q}_i is uniquely determined.

Proof. To simplify the notation, let us number the \mathfrak{p} 's so that the isolated subset is $\{\mathfrak{p}_1, \dots, \mathfrak{p}_t\}$. The set $S = R \setminus \mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_t$ is multiplicative and by hypothesis, if $\mathfrak{p}_i \cap S = \emptyset$, then $\mathfrak{p}_i \subseteq \mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_t$, hence $i \leq t$. Thus S meets $\mathfrak{p}_{t+1}, \dots, \mathfrak{p}_r$ and no others. Now form R_S ; since $\mathfrak{p}_i \cap S = \emptyset$ for $i \leq t$, we have $\mathfrak{q}_i^{e_i} = \mathfrak{q}_i$, (cf. Prop. 3.2) while for $i > t$, $\mathfrak{p}_i \cap S \neq \emptyset$. Take $a \in \mathfrak{p}_i \cap S$; then $a^n \in \mathfrak{q}_i$ for some n , but a^n is a unit in R_S , hence $(\mathfrak{q}_i)_S = (1)$ and it follows that

$$\mathfrak{a}_S = (\mathfrak{q}_1)_S \cap \dots \cap (\mathfrak{q}_t)_S, \quad \mathfrak{a}^{e_i} = \mathfrak{q}_1 \cap \dots \cap \mathfrak{q}_t.$$

This provides a description which is independent of the decomposition. ■

If \mathfrak{p} is a prime ideal in R and $S = R \setminus \mathfrak{p}$, then $(\mathfrak{p}^r)_S \cap R = \mathfrak{p}^{(r)}$ is sometimes called the *symbolic rth power* of \mathfrak{p} ; we note that it is the \mathfrak{p} -primary component of \mathfrak{p}^r .

Exercises

- (1) Find all possible primary decompositions of (x^2, xy) in $k[x, y]$. Do the same in $\mathbf{Z}[x, y]$.
- (2) Show that the members of $\text{Ass}(M)$ can also be defined as the ideals \mathfrak{a} for which there is a submodule L of M such that $\mathfrak{a} = \text{Ann}(L')$ for every non-zero submodule L' of L .
- (3) Let $R = k[x_1, x_2, \dots]$ and take M to be the R -module generated by elements u_1, u_2, \dots with defining relations $u_s x_s = 0$ ($s \leq r$). Show that $\text{Ass}(M)$ has no maximal elements.
- (4) Let R be a ring and S a multiplicative subset. Show that for any R -module M , any maximal ideal of the form $\text{Ann}(u)$ ($0 \neq u \in M$) and disjoint from S is prime. By taking a Noetherian localization of R show that such maximal ideals exist, and deduce that $\text{Ass}(M) \neq \emptyset$ unless $M = 0$.
- (5) In a ring R let \mathfrak{q}_i be \mathfrak{p}_i -primary ($i = 1, 2$), where $\mathfrak{p}_1 \supset \mathfrak{p}_2$ but neither of $\mathfrak{q}_1, \mathfrak{q}_2$ contains the other. Show that $\mathfrak{a} = \mathfrak{q}_1 \cap \mathfrak{q}_2$ is not primary although $ab \equiv 0 \pmod{\mathfrak{a}}$ implies $a^r \equiv 0$ or $b^s \equiv 0 \pmod{\mathfrak{a}}$ for some r, s . (An ideal \mathfrak{a} with this property is sometimes called *quasiprimary*.)
- (6) Consider the mapping $f: k[x, y, z] \rightarrow k[t]$ given by $x \mapsto t^3, y \mapsto t^4, z \mapsto t^5$. Show that its kernel is $\mathfrak{p} = (y^2 - xz, yz - x^3, z^2 - x^2y)$. Verify that \mathfrak{p} is prime and that \mathfrak{p}^2 is quasiprimary but not primary.
- (7) Show that if Q_1, Q_2 are primary, neither contains the other and $Q_1 \cap Q_2$ is \mathfrak{p} -primary, then Q_1, Q_2 are both \mathfrak{p} -primary.
- (8) Show that if $N = A \cap B$ is an irredundant primary decomposition and $B \subset B_1$, then B is meet-reducible.

- (9) Show that in Lemma 8.10(ii) the conclusion still holds if all but at most two of the \mathfrak{p}_i are prime.
- (10) Show that if \mathfrak{a} has no embedded components, then $\mathfrak{p} \in \text{Ass}(R/\mathfrak{a})$ iff $\mathfrak{a} \subseteq \mathfrak{p}$, $\mathfrak{a} \subset \mathfrak{a}:\mathfrak{p}$.
- (11) Show that a ring has a unique prime ideal iff it is completely primary (i.e. every non-unit is nilpotent).
- (12) Let R be a Noetherian ring and M a finitely generated R -module. Show that if an ideal \mathfrak{a} consists entirely of zero-divisors on M , then there exists $u \in M$, $u \neq 0$, such that $u\mathfrak{a} = 0$.
- (13) Let R be the function ring of the quadric $xy = z^2$. Show that $\mathfrak{p} = (x, z)$ is a prime ideal but that \mathfrak{p}^2 is not primary.
- (14) In any ring R , if \mathfrak{q} is \mathfrak{p} -primary, show that for any ideal \mathfrak{a} , (i) if $\mathfrak{a} \subset \mathfrak{q}$, then $\mathfrak{q}:\mathfrak{a} = (1)$, (ii) if $\mathfrak{a} \not\subseteq \mathfrak{p}$, then $\mathfrak{q}:\mathfrak{a} = \mathfrak{q}$, (iii) if $\mathfrak{a} \subseteq \mathfrak{p}$ but $\mathfrak{a} \not\subseteq \mathfrak{q}$, then $\mathfrak{q} \subset \mathfrak{q}:\mathfrak{a} \subset (1)$. Deduce that if \mathfrak{a} has the irredundant primary decomposition (10), then for any ideal \mathfrak{b} , $\mathfrak{a}:\mathfrak{b} = \mathfrak{a}$ iff $\mathfrak{b} \not\subseteq \mathfrak{p}_i$ for all i .
- (15) Let R be a ring and \mathfrak{a} an ideal. Show that if $\mathfrak{m}' \subseteq \mathfrak{a} \subseteq \mathfrak{m}$ for some maximal ideal \mathfrak{m} and some $r \geq 1$, then \mathfrak{a} is \mathfrak{m} -primary.

9.9 Dimension

The dimension of an algebraic variety may be defined as the maximum length of a chain of subvarieties. Since varieties correspond to prime ideals in the coordinate ring (Prop. 7.4), this suggests defining the dimension of a commutative ring in terms of chains of prime ideals.

In any commutative ring R , consider a chain of prime ideals

$$\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_r. \quad (1)$$

This chain is said to have *length r*; note that r is the number of links in the chain. Now the *Krull dimension*, or simply *dimension*, of R , written $\dim R$, is defined as the supremum of the lengths of chains of prime ideals in R (possibly infinite). For example, any field has dimension 0, $\dim \mathbb{Z} = 1$; more generally, any PID has dimension 1, and as we shall see in 9.10, $\dim k[x_1, \dots, x_n] = n$.

Our first aim is the observation that an integral extension does not raise the dimension; this will follow from the going-up theorem.

Let R be a ring, R' a ring containing R as a subring and such that R' is integral over R . We shall say more briefly: R' is an *integral extension* of R ; here neither R nor R' need be an integral domain. We first examine the behaviour of integral extensions under formation of quotient rings and rings of fractions.

PROPOSITION 9.1 *Let R' be an integral extension of R . Then (i) if \mathfrak{A} is an ideal in R' and $\mathfrak{a} = \mathfrak{A} \cap R$, then R'/\mathfrak{A} is an integral extension of R/\mathfrak{a} , (ii) if S is a multiplicative subset of R , then R'_S is an integral extension of R_S .*

Proof. (i) The inclusion $R \rightarrow R'$ induces a homomorphism $R/\mathfrak{a} \rightarrow R'/\mathfrak{A}$, given by $a \mapsto \bar{a}$, with kernel $(\mathfrak{A} \cap R)/\mathfrak{a} = 0$, hence an injection. Let $\xi \in R'/\mathfrak{A}$, say $\xi = \bar{x}$, where $x \in R'$ satisfies the monic equation $x^n + a_1x^{n-1} + \dots + a_n = 0$ ($a_i \in R$); on reducing mod \mathfrak{A} this becomes $\xi^n + \bar{a}_1\xi^{n-1} + \dots + \bar{a}_n = 0$, a monic equation for ξ , which is therefore integral over R/\mathfrak{a} .

(ii) It is clear that the induced mapping $R_S \rightarrow R'_S$ is injective, for any element of R_S has the form a/s , $a \in R$, $s \in S$, and $a/s = 0$ iff $at = 0$ for some $t \in S$. Now consider an element of R'_S , say x/s , where $x \in R'$, $s \in S$. If x satisfies the equation $x^n + a_1x^{n-1} + \dots + a_n = 0$, then $x/s = y$ satisfies $y^n + a_1/s \cdot y^{n-1} + \dots + a_n/s^n = 0$, and $a_i/s^i \in R_S$, so x/s is integral over R_S . ■

The key lemma for what follows is the next result:

LEMMA 9.2 *Let R' be an integral extension of R and assume that R' is an integral domain. Then R' is a field if and only if R is a field.*

Proof. Assume that R is a field and let $y \in R'$, $y \neq 0$; then y satisfies an equation $y^n + a_1y^{n-1} + \dots + a_n = 0$, where $a_i \in R$, and since R' is an integral domain, we may take $a_n \neq 0$. It follows that $a_n^{-1} \in R$, and so $y^{-1} = -a_n^{-1}(y^{n-1} + a_1y^{n-2} + \dots + a_{n-1}) \in R'$.

Conversely, if R' is a field, take $x \in R$, $x \neq 0$; then $x^{-1} \in R'$ and so we have an equation $x^{-m} + c_1x^{1-m} + \dots + c_m = 0$, where $c_i \in R$. Therefore $x^{-1} = -(c_1 + c_2x + \dots + c_mx^{m-1}) \in R$. ■

For arbitrary integral extensions this yields

COROLLARY 9.3 *Let R' be an integral extension of R , \mathfrak{P} a prime ideal in R' and $\mathfrak{p} = \mathfrak{P} \cap R$. Then \mathfrak{P} is maximal in R' if and only if \mathfrak{p} is maximal in R .*

For R'/\mathfrak{P} is an integral domain, integral over R/\mathfrak{p} by Prop. 9.1, so we can apply the lemma. ■

Given $R \subset R'$ and ideals $\mathfrak{p}, \mathfrak{P}$ in R, R' respectively, we shall say that \mathfrak{P} lies over \mathfrak{p} if $\mathfrak{P} \cap R = \mathfrak{p}$.

COROLLARY 9.4 *Let R' be an integral extension of R and let \mathfrak{p} be a prime ideal in R . If $\mathfrak{P} \subseteq \mathfrak{P}'$ are two prime ideals in R' both lying over \mathfrak{p} , then $\mathfrak{P}' = \mathfrak{P}$; thus the prime ideals lying over a given prime ideal of R are pairwise incomparable.*

Proof. Put $S = R \setminus \mathfrak{p}$; then R'_S is integral over $R_S (= R_{\mathfrak{p}})$ and the latter is a local ring with maximal ideal \mathfrak{p}_S . If $\mathfrak{m}, \mathfrak{m}'$ are the extensions of $\mathfrak{P}, \mathfrak{P}'$ to R'_S , then both contract to \mathfrak{p}_S in R_S , so by Cor. 9.3 both are maximal. But $\mathfrak{m} \subseteq \mathfrak{m}'$, hence they are equal and $\mathfrak{P} = \mathfrak{m} \cap R' = \mathfrak{m}' \cap R' = \mathfrak{P}'$. ■

The essential idea of this proof already occurred in the proof of Th. 5.5 (cf. the proof of Lemma 9.2 above).

The existence of prime ideals lying over a given prime ideal is assured by

LEMMA 9.5 *Let R' be an integral extension of R and let \mathfrak{p} be a prime ideal in R . Then there exists a prime ideal \mathfrak{P} in R' lying over \mathfrak{p} .*

Proof. Put $S = R \setminus \mathfrak{p}$ and form R_S, R'_S with canonical mappings $\lambda: R \rightarrow R_S, \lambda': R' \rightarrow R'_S$. If \mathfrak{m} is any maximal ideal in R'_S , then $\mathfrak{m} \cap R_S$ is maximal in R_S , by Cor. 9.3, so $\mathfrak{m} \cap R_S = \mathfrak{p}_S$. Now the inverse image of \mathfrak{m} in R' is a prime ideal: $\mathfrak{P} = \mathfrak{m}\lambda'^{-1}$, and by the commutativity of the diagram shown, we have $\mathfrak{P} \cap R = \mathfrak{p}$, so \mathfrak{P} is the desired prime ideal. ■

$$\begin{array}{ccc} R' & \xrightarrow{\lambda'} & R'_S \\ \uparrow & & \uparrow \\ R & \xrightarrow{\lambda} & R_S \end{array}$$

THEOREM 9.6 (Going-up theorem) *Let R' be an integral extension of R . Given a chain of prime ideals in R :*

$$\mathfrak{p}_1 \subset \mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_n, \quad (2)$$

and a chain

$$\mathfrak{P}_1 \subset \mathfrak{P}_2 \subset \cdots \subset \mathfrak{P}_m \quad (m \leq n), \quad (3)$$

in R' , where $\mathfrak{P}_i \cap R = \mathfrak{p}_i$ ($i = 1, \dots, m$), we can find prime ideals $\mathfrak{P}_{m+1}, \dots, \mathfrak{P}_n$ to extend the chain (3).

Proof. When $m = 0$, the second chain is absent and we can find \mathfrak{P}_1 lying over \mathfrak{p}_1 by Lemma 9.5. Now let $m \geq 1$, put $\bar{R} = R/\mathfrak{p}_m, \bar{R}' = R'/\mathfrak{P}_m$, so that \bar{R}' is an integral domain, integral over \bar{R} , and $\bar{\mathfrak{p}}_{m+1} = \mathfrak{p}_{m+1}/\mathfrak{p}_m$ is a prime ideal of \bar{R} . Then by Lemma 9.5 there is a prime ideal $\bar{\mathfrak{P}}_{m+1}$ in \bar{R}' lying over $\bar{\mathfrak{p}}_{m+1}$. The inverse image of $\bar{\mathfrak{P}}_{m+1}$ in R' is a prime ideal \mathfrak{P}_{m+1} containing \mathfrak{P}_m such that $\mathfrak{P}_{m+1} \cap R = \mathfrak{p}_{m+1}$. Now the result follows by induction on m . ■

With the help of this result it is easy to show that integral extensions preserve the dimension.

COROLLARY 9.7 *If R' is an integral extension of R , then $\dim R' = \dim R$.*

Proof. Let $\mathfrak{P}_0 \subset \mathfrak{P}_1 \subset \cdots \subset \mathfrak{P}_r$ be a prime ideal chain of length r in R' and put $\mathfrak{p}_i = \mathfrak{P}_i \cap R$; then by Cor. 9.4, the \mathfrak{p}_i are distinct and we get a chain

$$\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_r \quad (4)$$

in R ; this shows that $\dim R' \geq \dim R$. Conversely, over any chain (4) in R we can by Th. 9.6 find a chain in R' and so $\dim R' = \dim R$. ■

A natural question at this point concerns the relation between $\dim R$ and $\dim R[x]$, where x is an indeterminate. Given a prime ideal chain of length r in R , say (4), let \mathfrak{P}_i be the ideal of $R[x]$ generated by \mathfrak{p}_i ; then $R[x]/\mathfrak{P}_i \cong (R/\mathfrak{p}_i)[x]$; since the latter is an integral domain, \mathfrak{P}_i is a prime ideal. Likewise $\mathfrak{P}_i + (x)$ is prime, because $R[x]/(\mathfrak{P}_i + (x)) \cong R/\mathfrak{p}_i$. Thus

$$\mathfrak{P}_0 \subset \mathfrak{P}_1 \subset \cdots \subset \mathfrak{P}_r \subset \mathfrak{P}_r + (x) \quad (5)$$

is a prime ideal chain of length $r + 1$ in $R[x]$, for the inclusions in (5) are clearly proper. It follows that

$$\dim R[x] \geq \dim R + 1. \quad (6)$$

For Noetherian rings equality holds in (6) (cf. Matsumura 1985, Nagata 1962), but not in general (Seidenberg).

It is easy to determine the zero-dimensional rings completely. For this task we shall need

LEMMA 9.8 *Let R be a ring in which 0 can be written as a product of maximal ideals. Then any R -module is Artinian if and only if it is Noetherian.*

Proof. Assume that $\mathfrak{m}_1 \cdots \mathfrak{m}_r = 0$, where the \mathfrak{m}_i are maximal ideals of R , not necessarily distinct, and for any R -module M consider the chain

$$M \supseteq M\mathfrak{m}_1 \supseteq M\mathfrak{m}_1\mathfrak{m}_2 \supseteq \cdots \supseteq M\mathfrak{m}_1 \cdots \mathfrak{m}_r = 0. \quad (7)$$

Write $k_i = R/\mathfrak{m}_i$; since \mathfrak{m}_i is maximal in R , k_i is a field, and the quotient $M/M\mathfrak{m}_1$ can be considered as a k_1 -module, i.e. a vector space over k_1 ; for two elements of R have the same effect on $M/M\mathfrak{m}_1$ if they are congruent mod \mathfrak{m}_1 . Such a vector space is finite-dimensional iff it satisfies the maximum condition on subspaces, or equivalently the minimum condition, so we have a composition series from M to $M\mathfrak{m}_1$ iff the modules between M and $M\mathfrak{m}_1$ satisfy the maximum or minimum condition. The other links in the chain (7) can be treated similarly, hence M has a composition series (of R -modules) iff it satisfies the maximum or the minimum condition on submodules. ■

THEOREM 9.9 *Let R be a commutative ring. Then R is Artinian if and only if R is Noetherian and zero-dimensional.*

Proof. Suppose that R is Artinian. We first note that an Artinian domain is a field, for if $a \in R$, $a \neq 0$, then $(a) \supseteq (a^2) \supseteq \cdots$ is a descending chain, which must terminate, say $(a^r) = (a^{r+1})$, hence $a^r = a^{r+1}b$, and so $ab = 1$. It follows that in an Artinian ring any prime ideal is maximal, hence $\dim R = 0$. Further, R is Noetherian by Hopkins' theorem (Cor. 5.4.9).

Conversely, when R satisfies the conditions, take a primary decomposition of $0 = q_1 \cap \cdots \cap q_r$. This shows that R has finitely many minimal prime ideals $p_i = \sqrt{q_i}$; but each p_i is also maximal because $\dim R = 0$. So $\mathfrak{N} = p_1 \cap \cdots \cap p_r$ is the nilradical, and since R is Noetherian, $\mathfrak{N}^k = 0$ for some k , (cf. 9.4), so again $(p_1 \cdots p_r)^k = 0$, and by Lemma 9.8, we find that R is Artinian. ■

To study one-dimensional rings we first look at the form taken by the primary decomposition. We recall from 5.2 that for a family of pairwise comaximal ideals, their intersection equals their product. Moreover, if \sqrt{a}, \sqrt{b} are comaximal, then so are a, b . For if $a + b$ is proper, it is contained in some maximal ideal m , say, thus $a \subseteq m$, and so $\sqrt{a} \subseteq \sqrt{m} = m$, and similarly $\sqrt{b} \subseteq m$, but this contradicts the comaximality of \sqrt{a} and \sqrt{b} .

PROPOSITION 9.10 *In a Noetherian domain of dimension 1, every non-zero ideal can be written uniquely as a product of primary ideals with distinct radicals.*

Proof. Any non-zero ideal a has a primary decomposition $a = q_1 \cap \cdots \cap q_r$, which may be taken irredundant, so the $p_i = \sqrt{q_i}$ are all distinct. The p_i are non-zero, hence maximal, and so coprime. By the remark preceding Prop. 9.10 the q_i are pairwise comaximal, hence $a = q_1 \cdots q_r$. If we also have $a = q'_1 \cdots q'_s$, where the $\sqrt{q'_i}$ are distinct, then this is $q'_1 \cap \cdots \cap q'_s$ and here the q'_i are uniquely determined as the isolated components of a , hence $s = r$ and the q'_i and q_i coincide except for order. ■

To get a more precise description we want to be able to say that the q_i are actually powers of prime ideals; for this to hold the ring must be integrally closed and we thus again reach the class of Dedekind domains (9.5).

THEOREM 9.11 *A Noetherian domain of dimension 1 is Dedekind if and only if every primary ideal is a prime power.*

Proof. Let R be a one-dimensional Noetherian domain in which every primary ideal is a prime power. By Prop. 9.10 every non-zero ideal is then uniquely expressible as a product of powers of distinct prime ideals, hence R is then a Dedekind domain, by Th. 5.6. Conversely, let R be a Dedekind domain and $a \neq 0$ an ideal. By Th. 5.6 we can write $a = p_1^{n_1} \cdots p_r^{n_r}$, where the p_i are maximal and $n_i > 0$. Clearly $\sqrt{a} = p_1 \cdots p_r$, so a is primary iff $r = 1$, and then a is a power of a prime ideal. ■

Exercises

- (1) Let S be an integral domain and R a subring. If $a \neq 0$ is a finitely generated ideal in R and $c \in S$ is such that $ca \subseteq a$, then c is integral over R .

- (2) Let R' be an integral domain and R a subring and define the *conductor* of R in R' as $\mathfrak{f} = \{x \in R \mid xR' \subseteq R\}$. Show that \mathfrak{f} is the largest ideal in R which is also an ideal in R' . Let \mathfrak{f} be the conductor of a ring in its integral closure; show that for any multiplicative set S in R , if $\mathfrak{f} \cap S \neq \emptyset$, then R_S is integrally closed. Is the converse true?
- (3) Give an example of a local ring of infinite dimension.
- (4) Show that the ring $k[x_1, x_2, \dots] | x_i x_j = 0 \text{ for } i \neq j$ is zero-dimensional but not Artinian (so not Noetherian).
- (5) Show that an Artinian ring has only finitely many prime ideals. (*Hint.* Use Lemma 8.10.) Deduce that any Artinian (commutative) ring is a finite direct product of local rings.
- (6) Let K be a field and k a subfield. Show that if K is finitely generated as k -algebra, then K is algebraic over k .
- (7) Let R be a Noetherian ring and \mathfrak{a} an ideal of R . Show that R/\mathfrak{a} has finite length as R -module iff the associated prime ideals are all maximal. Show that the same holds for a finitely generated R -module M iff every ideal in $\text{Supp}(M)$ is maximal.
- (8) Give a direct proof that $\dim k[x_1, \dots, x_n] \geq n$.

9.10 The Hilbert Nullstellensatz

We now return to the subject of 9.7; our object will be to establish a bijection between algebraic sets and radical ideals. We begin with a lemma on finitely generated extensions of fields.

LEMMA 10.1 (Noether normalization lemma) *Let R be an integral domain, finitely generated as a ring over a field k . Then there exist $x_1, \dots, x_n \in R$, algebraically independent over k , such that R is integral over $k[x_1, \dots, x_n]$.*

Proof. Let R be generated by y_1, \dots, y_m over k ; if the y 's are algebraically independent, there is nothing to prove. Otherwise there will be a relation $f(y_1, \dots, y_m) = 0$. We write

$$z_i = y_i - y_1^{r^{i-1}} \quad (i = 2, \dots, m),$$

where the integer r is to be determined. Then our relation becomes

$$f(y_1, z_2 + y_1^r, \dots, z_m + y_1^{r^{m-1}}) = 0. \quad (1)$$

Each monomial $\prod y_i^{b_i}$ in f gives rise to a term y_1^c , where

$$c = b_1 + r b_2 + \dots + r^{m-1} b_m, \quad (2)$$

plus terms in z_i but of lower degree in y_1 . We now choose r so that all sums (2)

corresponding to the different terms in f are distinct; this can be done by avoiding the positive integers t satisfying the equations $\sum t^{i-1}(b_i - b'_i) = 0$, corresponding to terms $\prod y_i^{b_i}$, $\prod y_i^{b'_i}$ in f . Then (1) is an equation for y_1 over $k[z_2, \dots, z_m]$ which after multiplying by an element of k becomes monic, because by construction only a single term in f contributes to the leading term in y_1 ; hence y_1 is integral over $k[z_2, \dots, z_m]$. Now the result follows by induction on m and the transitivity of integral dependence. ■

COROLLARY 10.2 *For any field k and indeterminates x_1, \dots, x_n ,*

$$\dim k[x_1, \dots, x_n] = n.$$

Proof. Write $R = k[x_1, \dots, x_n]$; we know from (6) of 9.9 that $\dim R \geq n$. Now let

$$0 = p_0 \subset p_1 \subset \dots \subset p_r \quad (3)$$

be a prime ideal chain in R and consider $\bar{R} = R/p_1$. If $a \mapsto \bar{a}$ is the residue class map, then \bar{R} is generated by $\bar{x}_1, \dots, \bar{x}_n$ over k and by Lemma 10.1, there exist $y_1, \dots, y_m \in \bar{R}$, algebraically independent over k , such that \bar{R} is integral over $S = k[y_1, \dots, y_m]$. Moreover, the \bar{x}_i are algebraically dependent, because $p_1 \neq 0$, hence $m < n$. Using induction on n , Cor. 9.7 and the chain (3), we obtain

$$n > m = \dim S = \dim \bar{R} \geq r - 1,$$

hence $r \leq n$ and it follows that $\dim R = n$. ■

The Nullstellensatz (literally: zero-point theorem or solution-set theorem) shows that a family of polynomials over an algebraically closed field has a common solution whenever the ideal they generate in the polynomial ring is proper.

THEOREM 10.3 (Nullstellensatz, weak form) *Let k be an algebraically closed field. Then each maximal ideal m in $R = k[x_1, \dots, x_n]$ has the form*

$$m = (x_1 - \alpha_1, \dots, x_n - \alpha_n), \quad \text{where } \alpha_i \in k. \quad (4)$$

In particular, $V(a) \neq \emptyset$ for any proper ideal a of R .

Proof. The ideal m given by (4) is maximal, because it is the kernel of the surjective mapping $x_i \mapsto \alpha_i$ from R to k .

Conversely, let m be a maximal ideal in R . Then $K = R/m$ is a field, finitely generated as k -algebra; hence by Lemma 10.1, K is integral over $k[y_1, \dots, y_r]$, for suitable y_1, \dots, y_r algebraically independent over k . By Lemma 9.2, $k[y_1, \dots, y_r]$ is a field, which means that $r = 0$, so K is integral over k , i.e. algebraic; but k is algebraically closed, hence $K = k$. If $x_i \mapsto \alpha_i \in k$ in the natural homomorphism from R to $R/m = k$, then $m \supseteq (x_1 - \alpha_1, \dots, x_n - \alpha_n)$. Here the right-hand side is maximal, so equality holds and m has the form (4).

Now let \mathfrak{a} be any proper ideal in R . Then $\mathfrak{a} \subseteq \mathfrak{m}$ for some maximal ideal \mathfrak{m} , hence $V(\mathfrak{a}) \supset V(\mathfrak{m}) = (\alpha) \neq \emptyset$. ■

It is now an easy matter to deduce that every radical ideal is of the form $I(S)$ for some subset S of k^n :

THEOREM 10.4 (Nullstellensatz, full form) *Let $R = k[x_1, \dots, x_n]$ be a polynomial ring, where k is an algebraically closed field. Then for any ideal \mathfrak{a} in R ,*

$$IV(\mathfrak{a}) = \sqrt{\mathfrak{a}}.$$

Proof. ('Rabinowitsch trick') Clearly $\sqrt{\mathfrak{a}} \subseteq IV(\mathfrak{a})$, for if $V(\mathfrak{a}) = S$ and $f \in \sqrt{\mathfrak{a}}$, then $f^r \in \mathfrak{a}$ say, and so $f(\alpha)^r = 0$ for all $\alpha \in S$, hence $f(\alpha) = 0$, i.e. $f \in I(S)$.

Conversely, assume that $f(\alpha) = 0$ for all $\alpha \in V(\mathfrak{a})$; we have to show that $f \in \sqrt{\mathfrak{a}}$. We form the ring $R' = R[x_0]$ in a further indeterminate x_0 and in it consider the ideal $\mathfrak{b} = \mathfrak{a}R' + (1 - x_0f)$. If \mathfrak{b} were proper, it would have a zero α , by Th. 10.3. This means in particular, that all elements of \mathfrak{a} vanish at α , so $f(\alpha) = 0$; but also $1 - x_0f(\alpha) = 0$, a contradiction. Hence \mathfrak{b} is improper, i.e. if $\mathfrak{a} = (g_1, \dots, g_r)$,

$$1 = \sum g_i h_i + (1 - x_0f)h, \quad \text{where } h, h_i \in R'.$$

Let us replace x_0 by $1/f$; then the second term on the right reduces to 0, while each h_i is now a polynomial in x_1, \dots, x_n and $1/f$. If we clear the denominators, we obtain an equation

$$f^r = \sum g_i h'_i, \quad \text{where } h'_i \in R.$$

Thus $f^r \in \mathfrak{a}$ and so $f \in \sqrt{\mathfrak{a}}$, as we wished to show. ■

We thus have an order-reversing bijection between the closed sets on k^n and radical ideals in $k[x_1, \dots, x_n]$. Moreover, the closed sets satisfy the following relations:

- (i) $\bigcap V(\mathfrak{a}_i) = V(\sum \mathfrak{a}_i)$,
- (ii) $V(\mathfrak{a}_1) \cup V(\mathfrak{a}_2) = V(\mathfrak{a}_1 \cap \mathfrak{a}_2) = V(\mathfrak{a}_1 \mathfrak{a}_2)$,
- (iii) $V(0) = k^n$, $V(1) = \emptyset$.

These relations are easily verified, and they show that the sets of the form $V(\mathfrak{a})$ may be regarded as the closed sets of a topology on k^n . This is called the *Zariski topology*. Although very different from the classical topologies (it is non-Hausdorff), it is useful in discussing relations in k^n . For example, a subset S of k^n is said to be *dense* in k^n if its closure (in the Zariski topology) is the whole of k^n ; this means that any polynomial vanishing on S vanishes on all of k^n . If f is a non-zero polynomial over an infinite field k , then the points where f does not vanish form a dense subset of k^n . This is just a restatement of the density property for algebraic inequalities (cf. Vol. 1, p. 163).

Exercises

- (1) Show that the Nullstellensatz holds for any function ring of an algebraic set in the following form: Let $k[X]$ be the function ring of a subset X of k^n , where k is algebraically closed. If $f \in k[X]$ vanishes at all the points of X where g_1, \dots, g_r vanish, then $f \in \sqrt{(g_1, \dots, g_r)}$.
- (2) Let $f, g \in k[X]$, the function ring of an algebraic set X over an algebraically closed field k . Show that if f/g is defined at each point of X , then there exists $h \in k[X]$ such that $f = gh$. (Hint. Write $f/g = f_x/g_x$, where f_x, g_x are defined at the point $x \in X$, and use the fact that the g_x are elements of $k[X]$ which have no common zero.)
- (3) Show that the Nullstellensatz is false over any field that is not algebraically closed.
- (4) Let R be an integral domain generated by x_1, \dots, x_n over a ring A . Show that there are polynomials u_1, \dots, u_r in the x_i with integer coefficients which are algebraically independent, and a non-zero element $a \in A$ such that $R[a^{-1}]$ is integral over $A[a^{-1}, u_1, \dots, u_r]$.
- (5) Show that if a ring R is finitely generated (over \mathbf{Z} or a field), then its Jacobson radical is nilpotent.
- (6) Given a direct proof of the Nullstellensatz for primary ideals and deduce the general form by applying Th. 9.9 and the fact that for any p -primary ideal q over a Noetherian ring, $p^r \subseteq q \subseteq p$ for some r .

Further exercises on Chapter 9

- (1) Let m, n be positive integers. When is it true that $(m):(n) = (m/n)$? When is $(m):(n) = (m)$? What other cases can arise?
- (2) Show that the set of all zerodivisors and 0 in any (commutative) ring is a union of prime ideals.
- (3) Let R be a reduced ring. In $R[x]$, if $fg = 0$, where $f = \sum a_i x^i$, $g = \sum b_j x^j$, show that $a_i b_j = 0$ for all i, j .
- (4) (M. Zafrullah) Show that the set of all polynomials in an indeterminate x over the real numbers, with rational constant term, is an atomic integral domain, but not a UFD.
- (5) Let us call a polynomial *full* if the ideal generated by its coefficients is the whole ring. Show that for any ring R and any $f, g \in R[x]$, fg is full iff f and g are both full.
- (6) Let S be the set of all full polynomials in $R[x]$. Show that S is a multiplicative set which contains no zerodivisors. (Hint. Recall from Vol. 1, p. 183, that if $f \in R[x]$ is a zerodivisor, then it annihilates an element of R .) Deduce that if R is a local ring, then so is $R[x]_S$.
- (7) (M. Nagata) Let $q(x_1, \dots, x_n)$ be a non-singular quadratic form in $n \geq 5$ variables over

- C. Show that the ring $R = \mathbf{C}[x_1, \dots, x_n]/(q)$ is a UFD. (Hint. Write q as $x_1x_2 + q_1(x_3, \dots, x_n)$, where q_1 is irreducible; now use Th. 3.7.)
- (8) Show that an infinite UFD with a finite group of units has infinitely many primes.
- (9) Let R be a ring and S a multiplicative set. Given a finitely presented R -module M and any homomorphism $\varphi: M_S \rightarrow N_S$, where N is any R -module, find a homomorphism $f: M \rightarrow N$ and $s \in S$ such that $(x/1)\varphi = xf/s$ for all $x \in M$. Deduce that a finitely presented R -module is M -projective iff $M_{\mathfrak{p}}$ is free for all maximal ideals \mathfrak{p} .
- (10) Let R be a ring and M a finitely generated R -module. Show that if \mathfrak{a} is an ideal such that $M\mathfrak{a} = M$, then there exists $a \in \mathfrak{a}$ such that $1 + a$ annihilates M . (Hint. If $u = (u_1, \dots, u_n)$ is a row of elements generating M , find a matrix C over \mathfrak{a} such that $uC = u$ and hence obtain $M.\det(I - C) = 0$.) Deduce a proof of Nakayama's lemma for commutative rings (Lemma 5.4.6).
- (11) Let R be a Noetherian ring, \mathfrak{a} an ideal in R and M a finitely generated R -module. Defining $M_{\omega} = \bigcap M\mathfrak{a}^n$, show that $M_{\omega} = \{x \in M \mid x(1 - a) = 0 \text{ for some } a \in \mathfrak{a}\}$. Deduce that if $M_{\omega} \subseteq N \subseteq M$, then $M_{\omega} = N_{\omega}$. (Hint. Use Ex. (10).)
- (12) (Krull's intersection theorem) Let R be a Noetherian ring and \mathfrak{a} an ideal. Show that $\bigcap \mathfrak{a}^n = 0$ iff there is no zerodivisor $\equiv 1 \pmod{\mathfrak{a}}$. Deduce that (i) in a Noetherian domain every proper ideal \mathfrak{a} satisfies $\bigcap \mathfrak{a}^n = 0$, and (ii) in a Noetherian local ring the maximal ideal \mathfrak{m} satisfies $\bigcap \mathfrak{m}^n = 0$.
- (13) Let \mathfrak{o} be a Dedekind domain with field of fractions K , let L be a finite separable extension of K , and \mathfrak{D} the integral closure of \mathfrak{o} in L . For any subset X of L define the complementary set X' as $X' = \{y \in L \mid T(xy) \in \mathfrak{o} \text{ for all } x \in X\}$, where T is the trace $T_{L/K}$. Show: (i) X' is an \mathfrak{o} -submodule of L , (ii) $X \subseteq Y \Rightarrow X' \supseteq Y'$, (iii) $zX \subseteq X \Rightarrow zX' \subseteq X'$, (iv) $(zX)' = z^{-1}X'$. Deduce that the complementary set of a fractional ideal is again a fractional ideal, and if $\mathfrak{D} = \mathfrak{o}[\alpha]$ and f is the minimal polynomial of α over \mathfrak{o} , then $(\mathfrak{D}')^{-1} = (f'(\alpha))$, where f' is the usual derivative. (Here $(\mathfrak{D}')^{-1}$ is called the *different*.) (Hint. $1, \alpha, \dots, \alpha^{n-1}$ is an \mathfrak{o} -basis of \mathfrak{D} . If $f(x)/(x - \alpha) = b_0 + b_1x + \dots + b_{n-1}x^{n-1}$, use the Lagrange interpolation formula to show that the $b_i/f'(\alpha)$ form a dual basis.)
- (14) Show that an ideal maximal among the non-finitely generated ideals in a ring is prime. Deduce I. S. Cohen's theorem, that if all the prime ideals of a ring are finitely generated, then the ring is Noetherian. Similarly, prove that if all the prime ideals of a ring are principal, then the ring is a principal ideal ring. (Hint. Recall the proof of Th. 5.4(d) \Rightarrow (a)).
- (15) Prove the identity $(\mathfrak{a} + \mathfrak{b} + \mathfrak{c})(\mathfrak{bc} + \mathfrak{ca} + \mathfrak{ab}) = (\mathfrak{b} + \mathfrak{c})(\mathfrak{c} + \mathfrak{a})(\mathfrak{a} + \mathfrak{b})$ between ideals in any (commutative) ring. Deduce that R is a Dedekind domain if it is a Noetherian domain in which every 2-generator ideal $\neq 0$ is invertible (Dedekind).
- (16) Show that in a UFD an ideal is projective iff it is principal.
- (17) Let \mathfrak{o} be a Dedekind domain distinct from its field of fractions K . Show that a non-

zero \mathfrak{o} -submodule of K is a fractional ideal iff it is finitely generated. Give an example of a non-zero \mathfrak{o} -submodule of K which is not a fractional ideal.

- (18) Show that if \mathfrak{o} is a Dedekind domain and \mathfrak{a} an integral ideal of \mathfrak{o} , then $\mathfrak{o}/\mathfrak{a}$ is Artinian.
- (19) Let \mathfrak{o} be a ring and $f:\mathfrak{a} \rightarrow M$ a homomorphism from an invertible ideal \mathfrak{a} into an \mathfrak{o} -module M . Show that if $\sum a_i b_i = 1$, where $a_i \in \mathfrak{a}$, $b_i \in \mathfrak{a}^{-1}$, then the map $x \mapsto \sum a_i f.b_i x$ is a homomorphism from \mathfrak{o} to M . Deduce that every divisible module over a Dedekind domain is injective.
- (20) Show that for a Dedekind domain \mathfrak{o} , the monoid of projectives under direct sums has cancellation and so has a group of fractions $K_0(\mathfrak{o})$. Show that $K_0(\mathfrak{o})$ becomes a ring under the tensor product and that $K_0(\mathfrak{o}) \cong \mathbb{Z} \oplus C(\mathfrak{o})$, where $C(\mathfrak{o})$ is the projective class group, i.e. the quotient of $K_0(\mathfrak{o})$ by the submonoid of free modules, with zero-multiplication. (*Hint. Use (8) of 9.6 to verify the isomorphism.*)
- (21) Let R be the function ring over k of an algebraic set X . Show that the functions which vanish at a given point p of X form a prime ideal \mathfrak{p} in R , and that $R_{\mathfrak{p}}$ (the ‘local ring at p ’) may be interpreted as the ring of functions defined at p .
- (22) A mapping $f:X \rightarrow Y$ between algebraic sets X, Y is said to be *regular* if for any polynomial function p on Y , the map $x \mapsto p(f(x))$ is a polynomial function on X . Verify that the correspondence which assigns to each algebraic set its function ring is a contravariant functor from algebraic sets and regular mappings to k -algebras and homomorphisms. Show that injective homomorphisms correspond to regular mappings with dense image (in the sense of the Zariski topology). Verify that the projection of the hyperbola $xy = 1$ on the x -axis has a dense image but is not surjective.
- (23) Let R be a ring and S a multiplicative subset. For any R -module M show that $\mathfrak{p} \mapsto \mathfrak{p}_S$ defines an injection from the subset of $\text{Ass}(M)$ consisting of primes disjoint from S to $\text{Ass}(M_S)$. Moreover, this is a bijection when R is Noetherian.
- (24) Show that $\text{Ass}(M) \subseteq \text{Supp}(M)$ for any module M . Further show that if R is Noetherian, then $\text{Ass}(M)$ and $\text{Supp}(M)$ have the same minimal members.
- (25) Let $\text{Ass}(M) = S' \cup S''$, where $S' \cap S'' = \emptyset$. Show that the family \mathcal{F} of submodules M' such that $\text{Ass}(M') \subseteq S'$ is inductive. Verify that for a maximal M' in \mathcal{F} , $\text{Ass}(M/M') \subseteq S''$. Deduce that then $\text{Ass}(M') = S'$, $\text{Ass}(M/M') = S''$.
- (26) Let f_1, \dots, f_m be any polynomials in x_1, \dots, x_n over an algebraically closed field K . Show that the equations $f_1 = \dots = f_m = 0$ may have no solution in any extension field of K , but when they do have a solution in some extension, then they already have a solution in K .
- (27) Show that if \mathfrak{a} is any ideal in $k[x_1, \dots, x_n]$, where k is an algebraically closed field, then $\mathfrak{a} = \bigcap \{\mathfrak{m} \mid \mathfrak{m} \text{ maximal ideal and } \mathfrak{m} \supseteq \mathfrak{a}\}$. (Geometrically this states that an algebraic variety is the union of its points.)

10

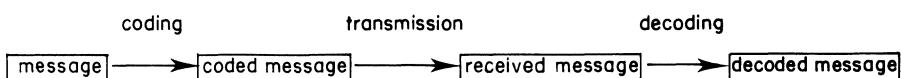
Coding theory

The theory of error-correcting codes deals with the design of codes which will detect, and if possible correct, any errors that occur in transmission. The subject dates from Shannon's classic paper on information theory (Shannon 1948) and we begin with a sketch of the background, leading up to a statement (but no proof) of Shannon's theorem. The remaining sections describe some of the more important codes; much of this is an application of the theory of finite fields (cf. 3.8).

10.1 The transmission of information

Coding theory is concerned with the transmission of information. For example, when a spacecraft wishes to send pictures back to Earth, these pictures are converted into electrical impulses, essentially representing strings of 0's and 1's, which are transmitted back to Earth. But the message sent may be distorted by 'noise' in space, and one has to build in some redundancy to overcome these errors (provided they are not too numerous).

To transmit our messages we need an *alphabet* Q consisting of q symbols, where $q \geq 2$; we also speak of a *q -ary code*. A finite string of letters from Q is called a *word*. The information to be transmitted is *encoded* by words of Q ; during transmission the coded message may be slightly changed (due to a 'noisy' channel) and, as a result, a slightly different message is received. However, if the code is appropriately chosen, it is nevertheless possible to decode the message so as to recover the original message.



Many of our codes are binary: $q = 2$. For example, in the game of 'twenty questions' an object has to be guessed by asking 20 questions which can be answered 'yes' or 'no'. This allows one to pick out one object in a million (since $2^{20} \sim 10^6$). Usually a binary code will have the alphabet {0, 1}; our coded message will then be a string of 0's and 1's. As a simple check we can add 1 when the

number of 1's in the coded message is odd and 0 when it is even. If the received message contains seven 1's we know that a mistake has occurred and we can ask for the message to be repeated (if this is possible). This is the *parity check*; it will show us when an odd number of errors occurs, but it does not enable us to correct errors, as is possible by means of more elaborate checks. Before describing ways of doing this we shall briefly discuss the question of information content, although strictly speaking this falls outside our topic. The rest of this section will not be used in the sequel and can be omitted without loss of continuity.

It is intuitively clear that the probability of error can be made arbitrarily small by adding sufficiently many checks to our message, and one might expect that this will make the transmission rate also quite small. However, a remarkable theorem due to Shannon asserts that every transmission channel has a capacity C , usually a positive number, and for any transmission rate less than C the probability of error can be made arbitrarily small. Let us briefly explain these terms.

The information content of a message is determined by the likelihood of the event it describes. Thus a message describing a highly probable event (e.g. 'the cat is on the mat') has a low information content, while for an unlikely message ('the cow jumped over the moon') the information content is large. If the probability of the event described by the message is p , where $0 \leq p \leq 1$, we shall assign as a measure of the information

$$-\log_2 p. \quad (1)$$

Here the minus sign is included to make the information positive and the logarithm is chosen to ensure that the resulting function is additive. If independent messages occur with probabilities p_1, p_2 then the probability that both occur is $p_1 p_2$ and here the information content is

$$-\log p_1 p_2 = -\log p_1 - \log p_2.$$

All logs are taken to the base 2 and the unit of information is the bit (= binary digit). Thus if we use a binary code and 0, 1 are equally likely, then each digit carries the information $-\log(1/2) = 1$, i.e. one bit of information.

Suppose we have a channel transmitting our binary code in which a given message, consisting of blocks of k bits, is encoded into blocks of n bits; the *information rate* of this system is said to be

$$R = k/n. \quad (2)$$

We assume further that the probability of error, x say, is the same for each digit; this is the *binary symmetric channel*. When an error occurs, the amount of information lost is $-\log x$, so on average the information lost per digit is $-x \cdot \log x$. But when no error occurs, there is also some loss of information (because we do not know that no error occurred); this is $(x - 1) \cdot \log(1 - x)$. The total amount of information lost per digit is therefore

$$H(x) = -x \cdot \log x - (1 - x) \cdot \log(1 - x).$$

This is also called the *entropy*, e.g. $H(0.1) = 0.469$, $H(0.01) = 0.0808$. The *channel capacity*, in bits per digit, is the amount of information passed, i.e.

$$C(x) = 1 - H(x).$$

Thus $C(0.1) = 0.531$, $C(0.01) = 0.9192$. We note that $C(0) = 1$; this means that for $x = 0$ there is no loss of information. By contrast, $C(1/2) = 0$, thus when there is an even chance of error, no information can be sent. The fundamental theorem of coding theory, proved by Shannon in 1948, states that for any $\delta, \varepsilon > 0$, there exist codes with information rate R greater than $C(x) - \varepsilon$, for which the probability of error is less than δ . In other words, information flows through the channel at nearly the rate $C(x)$ with a probability of error that can be made arbitrarily small. Here the information rate of the code is represented by (2). More generally, if there are M different code words, all of length n , then $R = (\log M)/n$. For a binary code there are 2^k words of length k , so $\log M = k$ and the rate reduces to k/n .

Exercises

- (1) How many questions need to be asked to determine one object in 10^9 if the reply is one of three alternatives?
- (2) In the binary symmetric channel with probability of error 1, no information is lost, i.e. $C(1) = 1$. How is this to be interpreted?
- (3) If n symbols are transmitted and the probability of error in one of them is x , show that the probability of k errors is $\binom{n}{k}x^k(1-x)^{n-k}$.

10.2 Block codes

Most of our codes in this chapter will be *block codes*, that is codes in which all code words have the same number of letters. This number, n say, is called the *length* of the code. Thus a block code of length n in an alphabet Q may be thought of as a subset of Q^n . For any $x, y \in Q^n$ we define the *distance* (also called the *Hamming distance*) between x and y , written $d(x, y)$, as the number of positions in which x and y differ, e.g. $d(pea, pod) = 2$. We note that this function satisfies the usual axioms of a metric space:

- M.1** $d(x, y) \geq 0$, with equality iff $x = y$,
- M.2** $d(x, y) = d(y, x)$,
- M.3** $d(x, y) + d(y, z) \geq d(x, z)$ (triangle inequality).

M.1–2 are clear, and **M.3** follows because if y differs in r places from x and in s places from z , then x and z can differ in at most $r + s$ places.

Let C be a block code which is q -ary of length n . The least distance $d = d(C)$ between code words of C is called the *minimum distance* of C . If the number of code words in C is M , then C is called a q -ary (n, M, d) -code, or an $(n, *, d)$ -code, if we do not wish to specify M . For successful decoding we have to ensure that the code words are not too close together: if $d(x, y)$ is large, this means that x and y differ in many of the n places, and x is unlikely to suffer so many changes in transmission that y is received. Thus our aim will be to find codes for which d is large. Our first result tells us how a large value of d allows us to detect and correct errors. A code is said to be *r-error-detecting (correcting)* if for any word differing from a code word u in at most r places we can tell that an error has occurred (resp. find the correct code word u).

We shall define the *r-sphere* about $x \in Q^n$ as the sphere of radius r with centre x :

$$B_r(x) = \{y \in Q^n \mid d(x, y) \leq r\}. \quad (1)$$

Clearly it represents the set of all words differing from x in at most r places.

PROPOSITION 2.1 *A code with minimum distance d can (i) detect up to $d - 1$ errors, and (ii) correct up to $\lceil (d - 1)/2 \rceil$ errors.*

Here $\lceil \xi \rceil$ denotes the greatest integer $\leq \xi$.

Proof. (i) If x is a code word and s errors occur in transmission, then the received word x' will be such that $d(x, x') = s$. Hence if $0 < s < d$, x' cannot be a code word and it will be noticed that an error has occurred.

(ii) If $e \leq \lceil (d - 1)/2 \rceil$, then $2e + 1 \leq d$ and it follows that the e -spheres about the different code words are disjoint. For if x, y are code words and $u \in B_e(x) \cap B_e(y)$, then

$$2e \geq d(x, u) + d(u, y) \geq d(x, y) \geq d \geq 2e + 1,$$

which is a contradiction. Thus for any word differing from a code word x in at most e places there is a unique nearest code word, namely x . ■

For example, the parity check code mentioned in 10.1 has minimum distance 2 and it will detect single errors, but will not correct errors.

Proposition 2.1 puts limits on the number of code words in an error-correcting code. Given n, d , we denote by $A(n, d)$ or $A_q(n, d)$ the largest number of code words in a q -ary code of length n and minimum distance d ; thus $A_q(n, d)$ is the largest M for which a q -ary (n, M, d) -code exists. A code for which this maximum is attained is also called an *optimal code*; any optimal (n, M, d) -code is necessarily *maximal*, i.e. it is not contained in an $(n, M + 1, d)$ -code.

To obtain estimates for $A_q(n, d)$ we need a formula for the number of elements in $B_r(x)$. This number depends on q, n, r but not on x ; it is usually denoted by $V_q(n, r)$. To find its value let us count the number of words at distance i from x .

These words differ from x in i places, and the values at these places can be any one of $q - 1$ letters, so there are $\binom{n}{i}(q - 1)^i$ ways of forming such words. If we do this for $i = 0, 1, \dots, r$ and add the results, we obtain

$$V(n, r) = V_q(n, r) = 1 + \binom{n}{1}(q - 1) + \binom{n}{2}(q - 1)^2 + \dots + \binom{n}{r}(q - 1)^r. \quad (2)$$

A set of spheres is said to *cover* Q^n or form a *covering* if every point of Q^n lies in at least one sphere. It is called a *packing* if every point of Q^n lies in at most one sphere (i.e. the spheres are non-overlapping). We note that

$$q^r = V_q(r, r) \leq V_q(n, r) \leq V_q(n, n) = q^n.$$

THEOREM 2.2 *Given integers $q \geq 2, n, d$, put $e = [(d - 1)/2]$. Then the number $A_q(n, d)$ of code words in an optimal $(n, *, d)$ -code satisfies*

$$\frac{q^n}{V_q(n, d - 1)} \leq A_q(n, d) \leq \frac{q^n}{V_q(n, e)}. \quad (3)$$

Proof. Let C be an optimal (n, M, d) -code; then C is maximal, and it follows that no word in Q^n has distance $\geq d$ from all the words of C , for such a word would allow us to enlarge the code, and so increase M . Hence every word of Q^n is within distance at most $d - 1$ of some word of C ; thus the $(d - 1)$ -spheres about the code words as centres cover Q^n and so $M \cdot V(N, d - 1) \geq q^n$, which gives the first inequality in (3).

On the other hand, we have $2e + 1 \leq d$, so the spheres $B_e(x)$, as x runs over an (n, M, d) -code, are disjoint and hence form a packing of Q^n . Therefore $M \cdot V(n, e) \leq q^n$, and the second inequality in (3) follows. ■

The above proof actually shows that a code with $q^n/V(n, d - 1)$ code words and minimum distance d can always be constructed; we shall not carry out the construction yet, since we shall see in 10.3 that it can always be realized by a linear code.

The first inequality in (3) is called the *Gilbert–Varshamov bound*, and the second is the *sphere-packing* or *Hamming bound*. A code is said to be *perfect* if, for some $e \geq 1$, the e -spheres with centres at the code words form both a packing and a covering of Q^n . Such a code is an $(n, M, 2e + 1)$ -code for which $M \cdot V(n, e) = q^n$, so it is certainly optimal. It is characterized by the property that every word of Q^n is nearer to one code word than to any of the others. To give an example, any code consisting of a single code word, or of the whole of Q^n is perfect. For $q = 2$ and odd n (with alphabet $\{0, 1\}$) the *binary repetition code* $\{0^n, 1^n\}$ is also perfect. These are the trivial examples; we shall soon meet non-trivial ones.

There are several ways of modifying a code to produce others, possibly with better properties. Methods of extending codes will be discussed in 10.3, when we

come to linear codes. For the moment we observe that from any (n, M, d) -code C we obtain an $(n - 1, M, d')$ -code, where $d' = d$ or $d - 1$, by deleting the last symbol of each word. This is called *puncturing* the code C . If we consider all words of C ending in a given symbol and take this set of words with the last symbol omitted, we obtain an $(n - 1, M', d')$ -code, where $M' \leq M$ and $d' \geq d$. This is called *shortening* the code C . Of course we can also puncture or shorten a given code by operating on any position other than the last one.

Exercises

- (1) Show that $A_q(n, 1) = q^n$, $A_q(n, n) = q$.
- (2) Use the proof of Th. 2.2 to construct an $(n, q^n/V_q(n, d - 1), d)$ -code, for any $q \geq 2$, n and d .
- (3) Show that there is a binary $(8, 4, 5)$ -code, and that this is optimal.
- (4) (The Singleton bound) Prove that $A_q(n, d) \leq q^{n-d+1}$. (*Hint.* Take an optimal (n, M, d) -code and puncture it repeatedly.)

10.3 Linear codes

By a *linear* code one understands a code with the finite field \mathbf{F}_q as alphabet, such that the set of code words forms a subspace. A linear code which is a k -dimensional subspace of an n -dimensional space over \mathbf{F}_q will be called an $[n, k]$ -code over \mathbf{F}_q . Thus any $[n, k]$ -code over \mathbf{F}_q is a linear (n, q^k, d) -code for some d . The number of non-zero entries of a vector x is called its *weight* $w(x)$; hence for a linear code we can write

$$d(x, y) = w(x - y). \quad (1)$$

A linear (n, M, d) -code has the advantage that to find d we need not check all the $\frac{1}{2}M(M - 1)$ distances between code words but only the M weights. Moreover, to describe an $[n, k]$ -code C we need not list all q^k code words but just give a basis of the subspace defining C . The $k \times n$ matrix whose rows form the basis vectors is called a *generator matrix* of C ; clearly its rank is k .

A matrix over a field is called *left full* if its rows are linearly independent, *right full* if its columns are linearly independent. For a square matrix these concepts are of course equivalent (Prop. 2 of 4.7, p. 98, Vol. 1) and we then speak of a *full* matrix. Thus a generator matrix of a linear code is left full, and any left full $k \times n$ matrix over \mathbf{F}_q forms a generator matrix of an $[n, k]$ -code.

Let C be an $[n, k]$ -code with generator matrix G . Any code word is a linear combination of the rows of G , since these rows form a basis for the code. Thus to encode a message $u \in \mathbf{F}_q^k$ we multiply it by G :

$$u \mapsto uG.$$

For example, for the simple parity check code with generator matrix $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ this takes the form

$$(u_1, u_2) \mapsto (u_1, u_2, u_1 + u_2).$$

For any $[n, k]$ -code we define the *dual code* C^\perp as

$$C^\perp = \{y \in \mathbf{F}_q^n \mid xy^T = 0 \text{ for all } x \in C\}.$$

Since any $x \in C$ has the form $x = uG$ ($u \in \mathbf{F}_q^k$), the vectors y of C^\perp are obtained as the solutions of the system $Gy^T = 0$. Here G has rank k , therefore C^\perp is an $(n - k)$ -dimensional subspace of \mathbf{F}_q^n ; thus C^\perp is an $[n, n - k]$ -code. It is clear from the definition that $C^{\perp\perp} = C$. When n is even, it may happen that $C^\perp = C$; in that case C is said to be *self-dual*. For example, the binary code with generator matrix $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$ is self-dual.

Generally, if G, H are generator matrices for codes C, C^\perp that are mutually dual, we have

$$GH^T = 0; \quad (2)$$

since G, H are both left full, it follows that the sum of their ranks is n and by the theory of linear equations (cf. Ch. 5, Vol. 1),

$$x = uG \text{ for some } u \in \mathbf{F}_q^k \Leftrightarrow xH^T = 0, \quad (3)$$

$$y = vH \text{ for some } v \in \mathbf{F}_q^{n-k} \Leftrightarrow yG^T = 0. \quad (4)$$

A generator matrix H for C^\perp is called a *parity check matrix* for C . For example, when $G = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, then $H = (1 \ 1 \ 1)$. For any code word x we form $xH^T = x_1 + x_2 + x_3$; if this is non-zero, an error has occurred. Our code detects one error, but no more (as we have already seen). Before introducing more elaborate codes we shall describe a normal form for the generator and parity check matrices.

Two block codes are said to be *equivalent* if one can be obtained from the other by permuting the n places of the code symbols and (in the case of a linear code) by multiplying the symbols in a given place by a non-zero scalar. For a generator matrix of a linear code these operations amount to (i) permuting the columns and (ii) multiplying a column by a non-zero scalar. We can of course also change the basis and this will not affect the code. This amounts to performing elementary operations on the rows of the generator matrix (and so may affect the encoding rules). We recall that any matrix over a field may be reduced to the form

$$\begin{pmatrix} I & P \\ 0 & 0 \end{pmatrix} \quad (5)$$

by elementary row operations and column permutations (Vol. 1, p. 117), and for a left full matrix the zero rows are clearly absent. Hence we obtain

THEOREM 3.1 *Any $[n, k]$ -code is equivalent to a code with a generator matrix of the form*

$$G = (I \quad P), \quad (6)$$

where P is a $k \times (n - k)$ matrix. ■

It should be emphasized that, whereas the row operations change merely the basis, the column permutations may change the code (to an equivalent code). More precisely, an $[n, k]$ -code has a generator matrix of the form (6) iff the first k columns of its generator matrix are linearly independent. This condition is satisfied in most practical cases. If we use a generator matrix G in the standard form (6) for encoding, the code word uG will consist of the message symbols u_1, \dots, u_k followed by $n - k$ check symbols.

The standard form (6) for the generator matrix of C makes it easy to write down the parity check matrix; its standard form is

$$H = (-P \quad I_{n-k})^T. \quad (7)$$

For we have $GH^T = P - P = 0$, and since H is left full, H^T is right full, thus of rank $n - k$, and it follows that the rows of H form a basis for the dual code C^\perp .

The process of decoding just consists in finding the code word nearest to the received word. Let us see how the parity check matrix may be used here. We have an $[n, k]$ -code C with parity check matrix H . For any vector $x \in \mathbf{F}_q^n$ the vector $xH^T \in \mathbf{F}_q^{n-k}$ is called the *syndrome* of x . By (3), the syndrome of x is 0 precisely when $x \in C$. More generally, two vectors $x, x' \in \mathbf{F}_q^n$ are in the same coset of C iff $xH^T = x'H^T$. Thus the syndrome determines the coset: if a vector x in C is transmitted and the received word y is $x + e$, then y and e have the same syndrome. To ‘decode’ y , i.e. to find the nearest code word, we choose a vector f of minimum weight in the coset of C containing y and then replace y by $y - f$. Such a vector f of minimum weight in its coset need not be unique; we choose such an f in each coset and call it the *coset leader*. The process described is called *syndrome decoding*.

For example consider a binary $[4, 3]$ -code with generator matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

To encode a vector we have

$$(u_1, u_2, u_3) \mapsto (u_1, u_2, u_3, u_1 + u_2 + u_3).$$

The parity check matrix is $H = (1 \quad 1 \quad 1 \quad 1)$. The possible syndromes are 0 and

1. We arrange the 16 vectors of \mathbf{F}_2^4 as a 2×8 array with cosets as rows, headed by the syndrome and coset leaders:

0	0000	1001	0101	0011	1100	1010	0110	1111
1	1000	0001	1101	1011	0100	0010	1110	0111

This is called a *standard array*. To decode x we form its syndrome $x_1 + x_2 + x_3 + x_4$. If this is 0, we can take the first three coordinates as our answer. If it is 1, we subtract the coset leader 1000 before taking the first three coordinates. We note that in this case there are four possible coset leaders for the syndrome 1; this suggests that the code is not very effective; in fact $d = 2$, so the code is 1-error-detecting.

Next take the $[4, 2]$ -code with generator and parity check matrices

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

A standard array is:

00	0000	1011	0101	1110
01	0100	1111	0001	1010
10	0010	1001	0111	1100
11	1000	0011	1101	0110

To decode $x = (1101)$ we form $xH^T = (11)$ and then subtract from x the coset leader for the syndrome (11), giving $(1101) - (1000) = (0101)$. The minimum distance is 2, so this code again detects single errors.

It is not necessary for decoding to write down the complete standard array, but merely the first column, consisting of the coset leaders. We note that this method of decoding assumes that all errors are equally likely, i.e. that we have a symmetric channel. In more general cases one has to modify the weight function by taking the probability of error into account; we shall not enter into the details.

We now turn to the construction of linear codes with a large value of M for given n and d . By the Gilbert–Varshamov bound in Th. 2.2 we have $A_q(n, d) \geq q^k$, provided that $q^{n-k} \geq V_q(n, d-1)$. However, this result does not guarantee the construction of linear codes. In fact we can construct linear codes with a rather better bound, as the next result shows.

THEOREM 3.2 *There exists an $[n, k]$ -code over \mathbf{F}_q with minimum distance at least d , provided that*

$$V_q(n-1, d-2) < q^{n-k}. \quad (8)$$

For comparison we note that Th. 2.2 gives $V_q(n, d-1) \leq q^{n-k}$, so (8) is a weaker condition.

Proof. Let C be any $[n, k]$ -code over \mathbf{F}_q with parity check matrix H . Each vector x in C satisfies $xH^T = 0$; this equation means that the entries of x define a linear dependence between the rows of H^T , i.e. the columns of H . We require a code for which the minimum distance is at least d , i.e. no vector in C has weight less than d ; this will follow if no $d - 1$ columns of H are linearly dependent.

To construct such a matrix H we need only choose successively n vectors in \mathbf{F}_q^{n-k} such that none is a linear combination of $d - 2$ of the preceding ones. In choosing the r th column we have to avoid the vectors that are linear combinations of at most $d - 2$ of the preceding $r - 1$ columns. We count the vectors to be avoided by picking $\delta \leq d - 2$ columns in $\binom{r-1}{\delta}$ ways and choosing the coefficients in $(q-1)^\delta$ ways. Hence the number of vectors to be avoided is

$$1 + \binom{r-1}{1}(q-1) + \binom{r-1}{2}(q-1)^2 + \cdots + \binom{r-1}{d-2}(q-1)^{d-2} \\ = V_q(r-1, d-2).$$

Thus we can adjoin an r th column, provided that $V_q(r-1, d-2) < q^{n-k}$. By (8) this holds for $r = 0, 1, \dots, n$, so we can form the required parity check matrix H . This proves the existence of a code with the required properties since it is completely determined by H . ■

Let us examine the case $d = 3$. If n, q are such that

$$V_q(n, 1) = q^{n-k}, \quad (9)$$

then (8) holds for $d = 3$, because V_q is an increasing function of its arguments; so when (9) holds, we can construct an $[n, k]$ -code C with minimum distance 3. This means that the 1-spheres about the code words are disjoint, and since by (9), $q^k \cdot V_q(n, 1) = q^n$, it follows that our code is perfect. These codes are known as *Hamming codes*. The equation (9) in this case reduces to $1 + n(q-1) = q^{n-k}$, i.e.

$$n = \frac{q^{n-k} - 1}{q - 1}. \quad (10)$$

Thus a Hamming code has the property that any two columns of its parity check matrix are linearly independent. In any code with odd minimum distance $d = 2e + 1$ every error pattern of weight at most e is the unique coset leader in its coset, because two vectors of weight $\leq e$ have distance $\leq 2e$ and so are in different cosets. For a perfect code all coset leaders are of this form. Thus in a Hamming code each vector of weight 1 is the unique vector of least weight in its coset. Now the number of cosets is $q^n/q^k = q^{n-k}$. Omitting the zero coset we see from (10) that we have just $n(q-1)$ non-zero cosets and these are represented by taking as coset leaders the $n(q-1)$ vectors of weight 1. This makes a Hamming code particularly easy to decode: given $x \in \mathbf{F}_q^n$, we calculate xH^T . If x has a single error (which is all

we can detect), then $xH^T = \gamma H_j^T$, where H_j is the j th column of H and $\gamma \in \mathbf{F}_q$. Now the error is corrected by subtracting γ from the j th coordinate of x .

The simplest non-trivial case is the binary [3, 1]-code with generator and parity check matrices

$$G = (1 \ 1 \ 1), \quad H = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

It consists in repeating each code word three times. The information rate is $\frac{1}{3} = 0.33$.

The next case of the Hamming code, the binary [7, 4]-code, is one of the best-known codes and one of the first to be discovered (in 1947). Its generator and parity check matrices are

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Here the information rate is $4/7 = 0.57$.

From any q -ary $[n, k]$ -code C we can form another code

$$\bar{C} = \{(x_1, x_2, \dots, x_n, -\sum x_i) | x \in C\},$$

called the *extension* of C by parity check. If C has odd minimum distance d , then \bar{C} has minimum distance $d + 1$ and its parity check matrix \bar{H} is obtained from that of C by bordering it first with a zero column and then a row of 1's. From an (n, M, d) -code we thus obtain an $(n + 1, M, d + 1)$ -code, and we can get C back by puncturing \bar{C} (in the last column).

Theorem 3.2 implicitly gives a lower bound for d in terms of q, n, k , but it does not seem easy to make this explicit. However, we do have the following upper bound for d .

PROPOSITION 3.3 (Plotkin bound) *Let C be a linear $[n, k]$ -code over \mathbf{F}_q . Then the minimum distance d of C satisfies*

$$d \leq \frac{n(q-1)q^{k-1}}{q^k - 1}. \quad (11)$$

Proof. C contains $q^k - 1$ non-zero vectors; their minimum weight is d , hence the sum of their weights is at least $d(q^k - 1)$. Consider the contribution made by the different components to this sum. If all vectors in C have zero first component, this contribution is 0; otherwise write C_1 for the subspace of vectors in C whose first component is zero. Then $C/C_1 \cong \mathbf{F}_q$ and so $|C_1| = q^{k-1}$. Thus there are $q^k - q^{k-1}$ vectors with non-zero first component. In all there are n components, and their total contribution to the sum of the weights is at most $n(q^k - q^{k-1})$. Hence $d(q^k - 1) \leq n(q^k - q^{k-1})$ and (11) follows. ■

Sometimes a more precise measure than the minimum distance is needed. This is provided by the *weight enumerator* of a code C , defined in terms of the weights of the code words as

$$A(z) = \sum_{a \in C} z^{w(a)} = \sum A_i z^i,$$

where A_i is the number of code words of weight i in C . A basic result, the *MacWilliams identity*, relates the weight enumerator of a code to that of its dual. This is useful for finding the weight enumerator of an $[n, k]$ -code when k is close to n , so that $n - k$ is small. We begin with a lemma on characters in fields. Here we understand by a character on a field F a homomorphism from the additive group of F to the multiplicative group of complex numbers, non-trivial if it takes values other than 1. By the duality of abelian groups (4.6) every finite field has non-trivial characters χ and $\sum \chi(a) = 0$ by orthogonality to the trivial character (Th. 4.6.3).

LEMMA 3.4 *Let χ be a non-trivial character of $(\mathbf{F}_q, +)$ and define*

$$f(u) = \sum_{v \in \mathbf{F}_q^n} \chi(uv^T) z^{w(v)}, \quad \text{where } u \in \mathbf{F}_q^n. \quad (12)$$

Then for any code C ,

$$\sum_{u \in C} f(u) = |C| \cdot B(z), \quad (13)$$

where $B(z)$ is the weight enumerator of the dual code C^\perp and $|C|$ is the number of code words in C .

Proof. We have

$$\sum_{u \in C} f(u) = \sum_{u \in C} \sum_{v \in \mathbf{F}_q^n} \chi(uv^T) z^{w(v)} = \sum_{v \in \mathbf{F}_q^n} z^{w(v)} \sum_{u \in C} \chi(uv^T).$$

For $v \in C^\perp$ the second sum on the right is $|C|$. If $v \notin C^\perp$, then uv^T takes every value in \mathbf{F}_q the same number of times, and we have $\sum \chi(uv^T) = N \cdot \sum \chi(a) = 0$, because χ is non-trivial. Hence the right-hand side reduces to $|C| \cdot B(z)$. ■

We can now derive the formula for the weight enumerator of the dual code,

THEOREM 3.5 (MacWilliams identity) *Let C be an $[n, k]$ -code over \mathbf{F}_q with weight enumerator $A(z)$ and let $B(z)$ be the weight enumerator of the dual code C^\perp . Then*

$$B(z) = q^{-k} [1 + (q - 1)z]^n \cdot A\left(\frac{1 - z}{1 + (q - 1)z}\right). \quad (14)$$

Proof. Let us extend the weight to \mathbf{F}_q by treating it as a one-dimensional vector

space; thus $w(a) = 1$ for $a \in \mathbf{F}_q^\times$ and $w(0) = 0$. Next, defining $f(u)$ as in (12), we have

$$\begin{aligned} f(u) &= \sum_{v \in \mathbf{F}_q^n} z^{\sum w(v_i)} \chi(\sum u_i v_i) \\ &= \sum_{v \in \mathbf{F}_q^n} \prod_{i=1}^n z^{w(v_i)} \chi(u_i v_i) \\ &= \sum_{t \in \mathbf{F}_q^\times} z^{w(t)} \chi(u_i t). \end{aligned}$$

If $u_i = 0$, the sum in this expression is $1 + (q - 1)z$, while for $u_i \neq 0$ it is

$$1 + z(\sum \chi(a)) = 1 - z.$$

Hence we obtain

$$f(u) = (1 - z)^{w(u)} [1 + (q - 1)z]^{n - w(u)}.$$

Substituting into (13) and remembering that $|C| = q^k$, we obtain (14). ■

Sometimes it is more convenient to use $A(z)$ in its homogeneous form, defined by $A(x, y) = A(yx^{-1})x^n = \sum A_i x^{n-i} y^i$. Then (14) takes the form

$$B(x, y) = q^{-k} A(x + (q - 1)y, x - y). \quad (15)$$

To illustrate Th. 3.5, consider the binary Hamming code C of length $n = 2^k - 1$ and dimension $n - k$ over \mathbf{F}_2 . Its dual code has as generator matrix the parity check matrix H of C , whose columns are all the non-zero vectors in \mathbf{F}_2^k . Hence any non-zero linear combination of the rows of $H = (h_{ij})$ has the i th coordinate

$$a_1 h_{1i} + a_2 h_{2i} + \cdots + a_k h_{ki}.$$

This vanishes for $2^{k-1} - 1$ columns (forming with 0 a $(k - 1)$ -dimensional subspace), and so is non-zero for the remaining 2^{k-1} columns. Hence every non-zero vector in the dual code C^\perp has exactly 2^{k-1} non-zero components, and so $B(z) = 1 + nz^{(n+1)/2}$. By Th. 3.5, the weight enumerator of the Hamming code is

$$A(z) = 2^{-k} \{ [1 + (q - 1)z]^n + n[1 + (q - 1)z]^{(n-1)/2} (1 - z)^{(n+1)/2} \}.$$

The weight enumerator is used in computing probabilities of transmission error. In any code C , an error, changing a code word x to $y = x + e$, will be detected provided that y is not a code word. Thus the error will go undetected if $y \in C$ or, equivalently, if $e \in C$. If the channel is binary symmetric, with probability p of symbol error, then the probability of the error vector e being of weight i is $p^i(1 - p)^{n-i}$. Thus the probability $P_{\text{un}}(C)$ that an incorrect code word will be received is independent of the code word sent and is

$$P_{\text{un}}(C) = \sum A_i p^i (1 - p)^{n-i} = A \left(\frac{p}{1-p} \right) (1 - p)^n.$$

Exercises

- (1) Show that a code C can correct t errors and detect a further s errors if $d(C) > 2t + s + 1$.
- (2) Let C be a block code of length n over an alphabet Q and let $\lambda \in Q$. Show that C is equivalent to a code which includes the word λ^n .
- (3) Show that for odd d , $A_2(n, d) = A_2(n+1, d+1)$. (*Hint.* Use extension by parity check and puncturing.)
- (4) Construct a table of syndromes and coset leaders for the ternary $[4, 2]$ -Hamming code.
- (5) Show that the binary $[7, 4]$ -Hamming code extended by parity check is self-dual.
- (6) Show that for linear codes $A_q(n, d) \geq q^k$, where k is the largest integer satisfying $q^k \cdot V_q(n-1, d-2) < q^n$.
- (7) Show that for an $[n, k]$ -code the MacWilliams identity can be written
$$\sum_{i=0}^n \binom{i}{r} A_i = q^{k-r} \cdot \sum_{i=0}^n (-1)^i \binom{n-i}{n-r} B_i, \quad \text{for } 0 \leq r \leq n.$$
- (8) Verify that the formula (14) is consistent with the corresponding formula for the dual code. (*Hint.* Use the form (15).)

10.4 Cyclic codes

A code C is said to be *cyclic* if the set of code words is unchanged by permuting the coordinates cyclically: if $c = (c_0, c_1, \dots, c_{n-1}) \in C$, then $(c_{n-1}, c_0, c_1, \dots, c_{n-2}) \in C$. We shall here assume all our cyclic codes to be linear. For cyclic codes it is convenient to number the coordinates from 0 to $n - 1$. We shall identify any code word c with the corresponding polynomial

$$c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1},$$

in the ring $A_n = \mathbf{F}_q[x]/(x^n - 1)$. In this sense we can interpret any linear code as a subset of A_n and the cyclic permutation corresponds to multiplication by x . Clearly a subspace of A_n admits multiplication by x iff it is an ideal, and this proves

THEOREM 4.1 *A linear code in \mathbf{F}_q^n is cyclic if and only if it is an ideal in $A_n = \mathbf{F}_q[x]/(x^n - 1)$.* ■

We note that the ring A_n has q^n elements; as a homomorphic image of $\mathbf{F}_q[x]$ it is a principal ideal ring, but it is not an integral domain, since $x^n - 1$ is reducible for $n > 1$.

Henceforth we shall assume that $(n, q) = 1$. Then $x^n - 1$ splits into distinct irreducible factors over \mathbf{F}_q and hence A_n is a direct product of extension fields of

\mathbf{F}_q (Cor. 5.7.4). By Th. 4.1 every cyclic code over \mathbf{F}_q can be generated by a polynomial g . Here g can be taken to be a factor of $x^n - 1$, since we can replace it by $ug - v(x^n - 1)$ without affecting the code. As a monic factor of $x^n - 1$ the generator of a cyclic code is uniquely determined. We record an expression for the generator matrix in terms of the generator polynomial:

THEOREM 4.2 *Let C be a cyclic code of length n with generator polynomial*

$$g = g_0 + g_1x + \cdots + g_rx^r \quad (g_r = 1).$$

Then $\dim(C) = n - r$ and a generator matrix for C is given by

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & \cdots & g_{r-1} & g_r & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & g_0 & g_1 & g_2 & \cdots & g_r \end{pmatrix}.$$

Proof. We have $g_r = 1$ and by considering the last r columns we see that G is left full. The $n - r$ rows represent the code words $g, xg, \dots, x^{n-r-1}g$ and we have to show that the linear combinations are just the code words. This is clear since the code words are of the form ag , where a is a polynomial of degree $< n - r$. ■

We next derive a parity check matrix for the cyclic code C . This is done most easily in terms of an appropriate polynomial. Let C be a cyclic $[n, k]$ -code with generator polynomial g . Then g is a divisor of $x^n - 1$, so there is a unique polynomial h satisfying

$$g(x)h(x) = x^n - 1. \quad (1)$$

h is called the *check polynomial* of C . Clearly it is monic, and its degree is $n - \deg g = n - (n - k) = k$. A cyclic code is called *maximal* if its generator polynomial is irreducible; if its dual is a maximal cyclic code, it is called *minimal* or *irreducible*. It is now an easy matter to describe a parity check matrix; to simplify the notation we shall use \equiv to indicate equality in A_n .

THEOREM 4.3 *Let C be a cyclic $[n, k]$ -code with check polynomial*

$$h = h_0 + h_1x + \cdots + h_kx^k.$$

Then

- (i) $a \in C$ if and only if $ah \equiv 0$,
- (ii) a parity check matrix for C is

$$H = \begin{pmatrix} h_k & h_{k-1} & \cdots & h_0 & 0 & 0 & \cdots & 0 \\ 0 & h_k & \cdots & h_1 & h_0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_k & h_{k-1} & h_{k-2} & \cdots & h_0 \end{pmatrix},$$

(iii) the dual code C^\perp is cyclic, generated by the reciprocal of the check polynomial for C :

$$\bar{h} = h_k + h_{k-1}x + \cdots + h_0x^k.$$

Proof. (i) By definition, $c \in C$ iff $c \equiv ag$. By (1) $gh \equiv 0$, hence if $c \in C$, then $ch \equiv agh \equiv 0$. Conversely, if $ch \equiv 0$, then ch is divisible by $x^n - 1 = gh$, hence c is divisible by g .

(ii) On multiplying the i th row of G by the j th row of H , we obtain

$$g_0h_{k-i+j} + g_1h_{k-i+j-1} + \cdots + g_{k-i+j}h_0, \quad (2)$$

and this vanishes, as the coefficient of x^{k-i+j} in gh . Thus $GH^T = 0$, and since H is a left full $r \times n$ matrix, it is indeed a parity check matrix for C .

(iii) By comparing the form of H with that for G , we see that \bar{h} is a generator polynomial for C^\perp . ■

We go on to describe how generator and check polynomials are used for coding and decoding a cyclic code. Let C be a cyclic $[n, k]$ -code with generator polynomial g of degree $r = n - k$ and check polynomial h of degree k . Given a message $a = a_0a_1 \cdots a_{k-1} \in \mathbf{F}_q^k$, we regard this as a polynomial $a = \sum a_i x^i$ of degree $< k$ over \mathbf{F}_q . We encode a by multiplying it by g : we obtain the polynomial $u = ag$ of degree $< n$. We note that any code word $\neq 0$ has degree at least $r = \deg g$.

For any polynomial f of degree $< n$ we calculate its syndrome $S(f)$ by multiplying the coefficients of f by the rows of the parity check matrix H . The result is

$$(fh)_k, (fh)_{k+1}, \dots, (fh)_{n-1}, \quad (3)$$

where for any polynomial φ in x , φ_i denotes the coefficient of x^i . To represent (3) as a polynomial, we take the polynomial part of $x^{-k}(fh)$, ignoring powers beyond x^{r-1} . This can also be achieved by reducing $fh \pmod{x^n - 1}$ to a polynomial of degree $< n$ and then taking the quotient of the division by x^k :

$$fh = x^k S(f) + p, \quad \text{where } \deg p < k. \quad (4)$$

Since $\deg f < n$, the highest possible power in fh is x^{n+k-1} . When reduced this becomes x^{k-1} and so does not affect the quotient in (4). Therefore $S(f)$ is indeed the syndrome of f , and as before, $S(f) = 0$ precisely when f has the form ag . By reducing $fh \pmod{x^n - 1}$, we obtain a representative of degree $< n$; hence $S(f)$ is of degree $< n - k = r$, as one would expect.

Now we choose for each possible syndrome u a coset leader $L(u)$ of least weight. To decode a word f we compute its syndrome $S(f)$ and subtract the corresponding coset leader: $f - LS(f)$ is a code word, so we have $(f - LS(f))h = a(x^n - 1)$, for some a , and this a is the required decoding of f . For example, $x^7 - 1 = (x - 1)(x^3 + x^2 + 1)(x^3 + x + 1)$ is a complete factorization over \mathbf{F}_2 . Let us take $g = x^3 + x + 1$, $h = x^4 + x^2 + x + 1$, $r = 3$, $k = 4$. Suppose we encode $x^2 + x$,

obtaining the code word $(x^2 + x)(x^3 + x + 1) = x^5 + x^4 + x^3 + x$. Owing to errors in transmission this is received as $x^5 + x^4 + x$. We have $x^5 + x^4 + x = (x^2 + x)(x^3 + x + 1) + x^3$, so the coset leader is x^3 , and adding this to the received word we have $x^5 + x^4 + x^3 + x$. Now $(x^5 + x^4 + x^3 + x)(x^4 + x^2 + x + 1) = (x^2 + x)(x^3 + x + 1)$, so our message is decoded as $x^2 + x$.

Sometimes it is useful to choose the generator for a cyclic code in a different way. We saw in Cor. 5.7.4 that $A_n = \mathbf{F}_q[x]/(x^n - 1)$ is a direct product of fields, say

$$A_n = K_1 \times \cdots \times K_s,$$

where K_i corresponds to the irreducible factor f_i of $x^n - 1$. The generator polynomial of a cyclic code C is the product of certain of the f_i , say (in suitable numbering) f_1, \dots, f_s . If e_i denotes the unit element in K_i , then $e = e_1 + \cdots + e_s$ is an element of A_n which is idempotent and which can also be used for the code. For the polynomials corresponding to code words are the elements of $K_1 \times \cdots \times K_s$ and these are the elements $c \in A_n$ such that $c = ce$. Thus every cyclic code has an idempotent generator; of course this will in general no longer be a factor of $x^n - 1$. To find the idempotent generator, suppose that C is a cyclic code with generator polynomial g and check polynomial h , so that $gh = x^n - 1$. Since g, h are coprime, there exist polynomials u, v such that $ug + vh = 1$. It follows that ug is the idempotent generator, for we have $(ug)^2 = ug(1 - vh) \equiv ug$. For example, in the above example the idempotent generator is $xg = x^4 + x^2 + x$.

To give examples of cyclic codes, let us again consider the binary $[n, n - k]$ -Hamming code. Its parity check matrix is $k \times n$ and its columns are all the non-zero vectors of \mathbf{F}_2^k , for they are distinct; hence any two are linearly independent, and there are $n = 2^k - 1$ of them. Let us write $q = 2^k$ and consider \mathbf{F}_q as a vector space over \mathbf{F}_2 ; this is a k -dimensional space, so the columns of the above parity check matrix are represented by the non-zero elements of \mathbf{F}_q . If these elements are $\alpha_1, \dots, \alpha_{q-1}$, we can regard

$$(\alpha_1, \dots, \alpha_{q-1}) \quad (5)$$

as a parity check matrix for the Hamming code. This will correct one error, and it seems plausible that we can correct more errors by including further rows, independent of (5).

To obtain such rows we recall that for any n distinct elements c_1, \dots, c_n over a field, the Vandermonde matrix

$$V(c_1, c_2, \dots, c_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ c_1 & c_2 & \cdots & c_n \\ c_1^2 & c_2^2 & \cdots & c_n^2 \\ \vdots & \vdots & & \vdots \\ c_1^{n-1} & c_2^{n-1} & \cdots & c_n^{n-1} \end{pmatrix}$$

is non-singular; in fact its determinant is $\prod_{i>j} (c_i - c_j)$ (Vol. 1, p. 192).

THEOREM 4.4 Let $q = 2^m$ and denote by $\alpha_1, \dots, \alpha_{q-1}$ the non-zero elements of \mathbf{F}_q . Then for any integers $t < q/2$ and $k \geq q - mt$ there is a $[q, k]$ -code over \mathbf{F}_2 with minimum distance at least $2t + 1$ and with parity check matrix

$$H = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{q-1} \\ \alpha_1^3 & \alpha_2^3 & \cdots & \alpha_{q-1}^3 \\ \alpha_1^5 & \alpha_2^5 & \cdots & \alpha_{q-1}^5 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{2t-1} & \alpha_2^{2t-1} & \cdots & \alpha_{q-1}^{2t-1} \end{pmatrix}. \quad (6)$$

Proof. A vector $c \in \mathbf{F}_2^q$ is a code word iff $cH^T = 0$, i.e.

$$\sum c_i \alpha_i^j = 0 \quad \text{for } j = 1, 3, 5, \dots, 2t-1. \quad (7)$$

On squaring the j th equation we find $(\sum c_i \alpha_i^j)^2 = \sum c_i \alpha_i^{2j}$, because $\alpha_i \in \mathbf{F}_q$ and $c_i \in \mathbf{F}_2$. Hence (7) holds for all $j = 1, 2, \dots, 2t$; if we insert the corresponding rows in the matrix (6), we see that the square matrix formed by the first $2t$ columns has determinant $\alpha_1 \cdots \alpha_{2t} V(\alpha_1, \dots, \alpha_{2t})$, and this is non-zero since $0, \alpha_1, \dots, \alpha_{2t}$ are distinct. Thus the first $2t$ columns of the new matrix are linearly independent, and similarly for any other set of $2t$ columns. This means that none of the vectors c in (7) can have weight $\leq 2t$, so the minimum distance of our code is at least $2t+1$. ■

The binary code with parity check matrix (6) is called a *BCH-code*, after its discoverers R. C. Bose, D. K. Ray-Chaudhuri and A. Hocquenghem.

Exercises

- (1) The *zero code* of length n is the subspace 0 of \mathbf{F}_q^n . Find its generator polynomial (as cyclic code) and describe its dual, the *universal code*.
- (2) The *repetition code* is the $[n, 1]$ -code consisting of all code words $(\gamma, \gamma, \dots, \gamma)$, $\gamma \in \mathbf{F}_q$. Find its generator polynomial and describe its dual, the *zero-sum code*.
- (3) Describe the cyclic code with generator polynomial $x + 1$ and its dual.
- (4) Verify that the $[7, 4]$ -Hamming code is cyclic and find its generator polynomial.
- (5) Show that the $[8, 4]$ -code obtained by extending the $[7, 4]$ -Hamming code by parity check is self-dual and has weight enumerator $z^8 + 14z^4 + 1$.
- (6) Show that a binary cyclic code contains a vector of odd weight iff $x - 1$ does not divide the generator polynomial. Deduce that such a code contains the repetition code.
- (7) The *weight* of a polynomial f is defined as the number $w(f)$ of its non-zero coefficients. Show that for polynomials over \mathbf{F}_2 , $w(fg) \leq w(f)w(g)$.

10.5 Other codes

There are many other codes adapted to various purposes, and it is neither possible nor appropriate to include the details here. But it may be of interest to make a brief mention of some of them.

(i) Goppa codes. In (7) of **10.4** we can take the elements of \mathbf{F}_q to be $\alpha_1^{-1}, \dots, \alpha_{q-1}^{-1}$. Then the defining equations for the BCH-code take the form $\sum_i c_i \alpha_i^{-j} = 0$ ($j = 1, 2, \dots, 2t$), or equivalently, $\sum_i c_i x^j \alpha_i^{-j} = 0$. This can also be written

$$\sum \frac{c_i}{x - \alpha_i} \equiv 0 \pmod{x^{2t}}, \quad (1)$$

and it leads to the following generalization.

DEFINITION Let g be a polynomial of degree t over \mathbf{F}_{q^m} and let $\alpha_0, \dots, \alpha_{n-1} \in \mathbf{F}_{q^m}$ be such that $g(\alpha_i) \neq 0$ ($i = 0, \dots, n-1$). The *Goppa code* with *Goppa polynomial* g is defined as the set of all vectors $c = (c_0, c_1, \dots, c_{n-1})$ satisfying

$$\sum \frac{c_i}{x - \alpha_i} \equiv 0 \pmod{g(x)}. \quad (2)$$

We see that Goppa codes are linear but not necessarily cyclic. In order to find a parity check matrix for the Goppa code we recall that in the special case of the BCH-code this was obtained by writing

$$(x - \alpha_i)^{-1} = - \sum x^j \alpha_i^{-j-1} (1 - x^{2t} \alpha_i^{-2t}).$$

The coefficients of x^j (taken mod x^{2t}) form the entries of the $(j+1)$ th row in the parity check matrix. We have

$$\frac{1}{x - \alpha} \equiv -g(\alpha)^{-1} \cdot \frac{g(x) - g(\alpha)}{x - \alpha} \pmod{g(x)}, \quad (3)$$

and here the right-hand side is a polynomial in x ; it is therefore the unique polynomial congruent (mod g) to $(x - \alpha)^{-1}$. So in order to express (2) as a polynomial in x we can proceed as follows: if $g(x) = \sum g_i x^i$, then

$$\frac{g(x) - g(y)}{x - y} = \sum g_{i+j+1} x^i y^j;$$

further write $g(\alpha_i)^{-1} = h_i$. Then (2) becomes $\sum c_i h_{ij} = 0$, where $h_{ij} = h_i \sum g_{v+j+1} x^j \alpha_i^v$. Thus the matrix (h_{ij}) is

$$\left(\begin{array}{cccc} h_0 g_t & h_1 g_t & \cdots & h_{n-1} g_t \\ h_0(g_{t-1} + g_t \alpha_0) & \cdots & \cdots & h_{n-1}(g_{t-1} + g_t \alpha_{n-1}) \\ \vdots & & & \vdots \\ h_0(g_1 + g_2 \alpha_0 + \cdots + g_t \alpha_0^{t-1}) & \cdots & \cdots & h_{n-1}(g_1 + g_2 \alpha_{n-1} + \cdots + g_t \alpha_{n-1}^{t-1}) \end{array} \right).$$

By elementary row transformations (remembering that $g_t \neq 0$) we find

$$H = \begin{pmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_0\alpha_0 & h_1\alpha_1 & \cdots & h_{n-1}\alpha_{n-1} \\ \vdots & \vdots & & \vdots \\ h_0\alpha_0^{t-1} & h_1\alpha_1^{t-1} & \cdots & h_{n-1}\alpha_{n-1}^{t-1} \end{pmatrix}.$$

We see again that any t columns are linearly independent, hence the Goppa code has minimum distance $> t$ and its dimension is $\geq n - mt$. There are several methods of decoding Goppa codes, based on the Euclidean algorithm (cf. McEliece 1977) and on work of Ramanujan (cf. Hill 1985).

(ii) Let C be any block code of length n . We obtain another code, possibly the same, by permuting the n places in any way. The permutations which do not change C form a subgroup of Sym_n , the group of C , which may be denoted by $G(C)$. For example, for a cyclic code the group contains all translations $i \mapsto i + r(\bmod n)$. If s is prime to n , we have the permutation

$$\mu_s : i \mapsto si \pmod{n}. \quad (4)$$

We remark that μ_s is an automorphism of $A_n = \mathbf{F}_q[x]/(x^n - 1)$. For if $a(x) = \sum a_i x^i$, then $a\mu_s = \sum a_i x^{si} = a(x^s)$, and the operation $a(x) \mapsto a(x^s)$ is clearly an endomorphism; since s is prime to n , μ_s has finite order dividing $\varphi(n)$, and so is an automorphism.

A QR-code is a cyclic code of length n , an odd prime, which admits the permutation (4) of its places, where s is a quadratic residue mod n . We shall examine a particular case of QR-codes, following essentially van Lint (1982). Let n again be an odd prime and q a prime power such that q is a non-zero quadratic residue mod n . We write Q for the set of all non-zero quadratic residues mod n , N for the set of all quadratic non-residues and let α be a primitive n th root of 1 in an extension of \mathbf{F}_q . Write $E = \mathbf{F}_q(\alpha)$ and put

$$g_0(x) = \prod_{r \in Q} (x - \alpha^r), \quad g_1(x) = \prod_{r \in N} (x - \alpha^r).$$

Then

$$x^n - 1 = (x - 1)g_0(x)g_1(x).$$

The Galois group of E/\mathbf{F}_q is generated by $x \mapsto x^q$. Since $q \in Q$, this operation permutes the zeros of g_0 , as well as those of g_1 ; therefore g_0 and g_1 have their coefficients in \mathbf{F}_q . We note that μ_s interchanges g_0 and g_1 if $s \in N$, hence the codes with generators g_0 and g_1 are equivalent. We shall be particularly interested in the QR-code generated by g_0 ; our aim will be to find restrictions on the maximum distance d . We recall from Ex. (24) of Ch. 3 that 2 is a quadratic residue mod n iff $n \equiv \pm 1 \pmod{8}$, and that -1 is a quadratic residue mod n iff $n \equiv 1 \pmod{4}$. We shall also need a lemma on weights; for a polynomial f (regarded as a code word) the weight $w(f)$ is of course just the number of non-zero coefficients.

LEMMA 5.1 Let f be a polynomial over \mathbf{F}_q such that $f(1) \neq 0$. Then $(1 + x + \cdots + x^{n-1})f$ has weight at least n . If $q = 2$ and $\deg f < n$, then $(1 + x + \cdots + x^{n-1})f$ has weight exactly n .

Proof. By the division algorithm, $f = (x - 1)u + c$, where $c = f(1) \neq 0$. It follows that

$$(1 + x + \cdots + x^{n-1})f = (x^n - 1)u + (1 + x + \cdots + x^{n-1})c. \quad (5)$$

Suppose that $w(u) = r$; then the right-hand side has at least r terms of degree $\geq n$, while the terms in u can cancel at most r terms in $(1 + x + \cdots + x^{n-1})c$. So the total weight is $\geq r + (n - r) = n$.

If $q = 2$ and $\deg f < n$, we again have (5), where now $\deg u < n - 1$. Each non-zero term in u will cancel a term in $1 + x + \cdots + x^{n-1}$, and this is exactly compensated by the corresponding term in $x^n u$. Hence there are exactly n terms on the right of (5). ■

PROPOSITION 5.2 Let C be a QR-code with generator g_0 and let $c = c(x)$ be a code word in C such that $c(1) \neq 0$. Then

$$w(c)^2 \geq n. \quad (6)$$

Moreover, if $n \equiv -1 \pmod{4}$, then

$$w(c)^2 - w(c) + 1 \geq n. \quad (7)$$

If further, $q = 2$ and $n \equiv -1 \pmod{8}$, then

$$w(c) \equiv -1 \pmod{4}. \quad (8)$$

Proof. The polynomial $c(x)$ is divisible by g_0 but not by $x - 1$, because $c(1) \neq 0$. For suitable s, μ_s will transform $c(x)$ into a polynomial $c^*(x)$ divisible by g_1 and again not by $x - 1$. This means that cc^* is a multiple of $g_0 g_1 = 1 + x + \cdots + x^{n-1}$, so by the lemma, $w(cc^*) \geq n$. Now (6) follows because $w(cc^*) \leq w(c)w(c^*) = w(c)^2$.

If $n \equiv -1 \pmod{4}$, then $-1 \in N$ and so the operation $x \mapsto -x$ transforms g_0 into g_1 ; thus $c(x)c(x^{-1})$ is divisible by $g_0 g_1$. Now corresponding terms in $c(x)$ and $c(x^{-1})$ give rise to a term of degree 0 in $c(x)c(x^{-1})$, so there are at most $w(c)^2 - w(c) + 1$ terms in all and (7) follows.

Finally assume that $n \equiv -1 \pmod{8}$ and $q = 2$; then (7) applies. Further, any code word c has degree $< n$; hence on writing $r = \deg c$, we have $x^r c(x^{-1})c(x) = f g_0 g_1$, where f is a polynomial of degree $< n$. By the lemma, this product has weight exactly n , and writing $d = w(c)$, we have $d^2 - d + 1 \geq n$. Now consider how terms in $c(x)c(x^{-1})$ can cancel. We have $c = \sum x^{r_i} c(x^{-1}) = \sum x^{-r_i}$ and a pair of terms in the product will cancel if $r_i - r_j = r_k - r_l$. But in this case $r_j - r_i = r_l - r_k$ and another pair will cancel, so that the terms cancel in fours. Hence we have $d^2 - d + 1 - 4t = n$; therefore $d^2 - d \equiv 2 \pmod{4}$, hence $d \equiv -1 \pmod{4}$. ■

Let us now take $q = 2$. Then the condition that q is a quadratic residue mod n gives $n \equiv \pm 1 \pmod{8}$. Consider the polynomial

$$\theta(x) = \sum_{r \in Q} x^r.$$

Since $\theta(x)^2 = \sum x^{2r} = \theta(x)$, it follows that θ is idempotent. In particular, for the primitive element α of E we have $\theta(\alpha)^2 = \theta(\alpha)$, hence $\theta(\alpha)$ is 0 or 1. For any $r \in Q$ we have $\theta(\alpha^r) = \theta(\alpha)$, while for $r \in N$, $\theta(\alpha^r) + \theta(\alpha) = \sum_1^{n-1} \alpha^r = 1$. If $\theta(\alpha) = 1$, replace α by α^s , where $s \in N$; since $(s, n) = 1$, α^s is again a primitive n th root of 1 and $\theta(\alpha^s) = 0$. Thus for suitable choice of α we have $\theta(\alpha) = 0$. It follows that

$$\theta(\alpha^i) = \begin{cases} 0 & \text{if } i \in Q, \\ 1 & \text{if } i \in N, \\ (n-1)/2 & \text{if } i = 0. \end{cases}$$

If $n \equiv 1 \pmod{8}$, then $\theta(\alpha^i)$ vanishes exactly when $i \in Q \cup \{0\}$, so the code is then generated by $(x-1)g_0$. Similarly, if $n \equiv 1 \pmod{8}$, then $\theta(\alpha^i)$ vanishes when $i \in Q$, so in this case the generator is g_0 .

Let $C(n)$ be the binary code defined in this way, and $\bar{C}(n)$ its extension by parity check. It can be shown that the group of $\bar{C}(n)$ is transitive on the $n+1$ places. (These places may be interpreted as the points on the projective line over \mathbf{F}_n , and the group is then the projective special linear group, cf. van Lint (1982, p. 88).) Consider a word $c \in C$ of minimum weight d . Since the group is transitive, we may assume that the last coordinate in \bar{c} (the parity check) is 1. This means that c has odd weight d , say, and so $c(1) = 1$. Hence by Prop. 5.2, $d \equiv -1 \pmod{4}$ and $d^2 - d + 1 \geq n$.

For example, for $n = 7$ we obtain the [7, 4]-Hamming code; here $d = 3$. A second (and important) example is the case $n = 23$. Here

$$g_0 = x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1.$$

Since $23 \equiv -1 \pmod{8}$, g_0 is a generator for this code, which is known as the [23, 12]-Golay code. Since $d^2 - d + 1 \geq 23$ and $d \equiv -1 \pmod{4}$, it follows that $d \geq 7$, so the 3-spheres about the code words form a packing. On checking their size, we note the remarkable fact that

$$V_2(23, 3) = 1 + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} = 2^{11}.$$

This shows that $C(23)$ is a perfect code, with minimum distance $d = 7$. The extended code \bar{C} is of length 24, with minimum distance 8, giving rise to the Leech lattice (a particularly close sphere packing in 24 dimensions). The symmetry group of a point in this lattice in \mathbf{R}^{24} is the first Conway group $\cdot 0$ ("dotto") of order $2^{22} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23 \sim 8.3 \times 10^{18}$. The quotient by its centre (which has order 2) is the sporadic simple group known as $\cdot 1$, discovered in 1968 by J. H. Conway (cf. Conway and Sloane 1988).

Exercises

- (1) Construct the ternary $[11, 6]$ -Golay code and verify that it is perfect. Find its weight enumerator.
- (2) Construct the extended ternary $[12, 6]$ -Golay code and find its weight enumerator. Is it self-dual?
- (3) A binary self-dual code is called *doubly-even* if all weights of code words are divisible by 4. Show that the extended $[8, 4]$ -Hamming code is doubly even. Show that if there is a $[2k, k]$ -code which is doubly even, then $k \equiv 0 \pmod{4}$.
- (4) Show that the extended binary $[24, 12]$ -Golay code is doubly-even. Find its weight enumerator.

Further exercises on Chapter 10

- (1) (The Plotkin bound) Prove that for any q -ary block code, if $d > \theta n$, where $\theta = 1 - q^{-1}$, then $A_q(n, d) \leq d/(d - \theta n)$. (*Hint.* Write the words of a maximal (n, M, d) -code as an $M \times n$ matrix and compute the sums of the distances between distinct words in two ways, along rows and along columns.)
- (2) Deduce the Plotkin bound of Prop. 3.3 from the general Plotkin bound of Ex. (1).
- (3) Show that the binary $[n, k]$ -Hamming code has a parity check matrix whose columns are the numbers 1 to $2^n - 1$ in binary notation.
- (4) Show that the weight enumerator $A(z)$ of the Hamming code satisfies the differential equation $(1 - z^2)A'' + (1 + nz)A = (1 + z)^n$.
- (5) Examine linear q -ary codes with the property that for any code word $c = (c_0, c_1, \dots, c_{n-1})$, $L(c) = (\lambda c_{n-1}, c_0, \dots, c_{n-2})$ is again a code word, where λ is a fixed element of \mathbf{F}_q . In particular consider the case $\lambda^n = 1$.
- (6) Show that for a self-dual code the homogeneous weight enumerator is invariant under the transformations

$$(x, y) \mapsto ([x + (q - 1)y]/\sqrt{q}, (x - y)/\sqrt{q}), \quad (x, y) \mapsto (x, \omega y),$$
 where $\omega^q = 1$. Show that for $q = 2$ the group generated by these transformations has order 16. What is the order for general q ?
- (7) Show that the weight enumerator of any binary self-dual code is a combination of $g_1 = z^2 + 1$ and $g_2 = z^8 + 14z^4 + 1$ (the Gleason polynomials). (*Hint.* Apply Molien's theorem (Ex. (8) of 7.4) to Ex. (6).)
- (8) In any ISBN book number $\alpha_1 \dots \alpha_{10}$ the final digit is chosen so that $\sum k\alpha_k \equiv 0 \pmod{11}$ (using digits 0, 1, ..., 9, X). Show that this allows single errors and transpositions to be detected, but not necessarily a total reversal (writing the number back to front).

11

Languages and automata

Many problems in mathematics consist in the calculation of a number or function, and our task may be to classify the different types of calculation that can arise. This can be done very effectively by describing simple machines which could carry out these calculations. Of course the discussion is entirely theoretical (we are not concerned with building the machines), but it is no accident that this way of thinking became current in the age of computers; A. M. Turing, one of the pioneers of digital computers, used just this method in 1936 to attack decision problems in logic, by introducing the class of ‘computable functions’, i.e. functions that could be computed on a Turing machine. This development has had many consequences, most of them outside our scope. However, the simplest machines, automata, have an immediate algebraic interpretation. In the algebraic study of languages one uses simple sets of rules (‘grammars’) to derive certain types of languages, not mirroring all the complexity of natural languages, but more akin to programming languages. It turns out that these languages can also be described in terms of the machines needed to generate them, and in this chapter we give a brief introduction to algebraic languages and automata.

The natural mathematical concept to describe languages is the free monoid, and we begin with a discussion of monoids and their actions. The monoid ring of a free monoid is a free associative algebra, an object of independent interest, which in turn can be used to study languages; so a brief account of free algebras and their completions (free power series rings) is included. This is also the natural place to study variable-length codes, which in their turn have influenced the development of free monoids and free algebras.

11.1 Monoids and monoid actions

We recall from Vol. 1 (3.1) that a *monoid* is a set M with a binary operation $x, y \mapsto xy$ and a distinguished element 1, the *neutral element* or also *unit element*, such that

$$\begin{aligned} M.1 \quad & x(yz) = (xy)z \quad \text{for all } x, y, z \in M \quad (\text{associative law}), \\ M.2 \quad & x1 = 1x = x. \end{aligned}$$

Groups form the particular case where every element has an inverse. As an example of a monoid other than a group we may take, for any set A , the set $\text{Map}(A) = A^A$ of all mappings of A into itself, with composition of mappings as multiplication. Many of the concepts defined for groups have a natural analogue for monoids; e.g. a *submonoid* of a monoid M is a subset of M containing 1 and admitting multiplication. A *homomorphism* between monoids M, N is a mapping $f : M \rightarrow N$ such that $(xy)f = xf.yf$, $1_M f = 1_N$ for $x, y \in M$. Here we had to assume explicitly that the unit element is preserved by f ; for groups this followed from the other condition. A *generating set* of a monoid M is a subset X such that every element of M can be written as a product of a number of elements of X . For example, the set \mathbf{N} of natural numbers is a monoid under multiplication, with neutral element the number 1; here a generating set is given by the set of all prime numbers, for every positive integer can be written as a product of prime numbers, with 1 expressed as the empty product. Likewise the set $\mathbf{N}_0 = \mathbf{N} \cup \{0\}$ is a monoid under addition, with neutral element 0 and generating set $\{1\}$.

An example of particular importance for us in the sequel is the following monoid. Let X be any set, called the *alphabet*, and denote by X^* the set of all finite sequences of elements of X :

$$w = x_1 x_2 \cdots x_r, \quad x_i \in X, r \geq 0. \quad (1)$$

Here we include the empty sequence, written as 1. We define multiplication in X^* as juxtaposition:

$$(x_1 \cdots x_r)(y_1 \cdots y_s) = x_1 \cdots x_r y_1 \cdots y_s. \quad (2)$$

The associative law is easily verified, and it is also seen that the empty sequence 1 is the neutral element. X^* is called the *free monoid* on X . We remark that when $X = \emptyset$, X^* reduces to the trivial monoid consisting of 1 alone. This case will usually be excluded in what follows. Apart from this trivial case the simplest free monoid is that on a one-element set $\{x\}$ say. The elements are $1, x, x^2, \dots$, with the usual multiplication. We see that $\{x\}^*$ is isomorphic to \mathbf{N}_0 , the monoid of non-negative integers under addition, by the rule $n \leftrightarrow x^n$. Since the expression (1) for an element w of a free monoid is unique, the number r of factors on the right is an invariant of w , called its *length* and written $|w|$.

The name ‘free monoid’ is justified by the following result:

THEOREM 1.1 *Every monoid is a homomorphic image of a free monoid.*

Proof. Let M be any monoid and A a generating set. Take a set A' in bijective correspondence with A and write F for the free monoid on A' . We have a mapping $f : F \rightarrow M$ defined by

$$(a'_1 \cdots a'_r)f = a_1 \cdots a_r, \quad (3)$$

where $a' \leftrightarrow a$ is the given correspondence between A' and A . Since every element of

F can be written as a product $a'_1 \cdots a'_r$ in just one way, f is well-defined by (3). It is surjective, because A generates M , and f is easily seen to be a homomorphism, using (2). ■

Just as groups can be represented by permutations, so can monoids be realized by means of mappings. If M is any monoid, then by an M -set or a set with an M -action we understand a set S , with a mapping from $S \times M$ to S , written $(s, x) \mapsto sx$, such that

$$\text{S.1} \quad s(xy) = (sx)y \quad \text{for all } s \in S, x, y \in M,$$

$$\text{S.2} \quad s1 = s.$$

Writing for the moment R_x for the mapping $s \mapsto sx$ of S into itself, we can express S.1, 2 as

$$R_{xy} = R_x R_y, \quad R_1 = 1. \quad (4)$$

This just amounts to saying that the mapping $R: x \mapsto R_x$ is a monoid homomorphism of M into $\text{Map}(S)$. For example, M itself is an M -set, taking the multiplication in M as M -action. This is sometimes called the *regular representation* of M . We can use it to obtain the following analogue of Cayley's theorem for groups (Th. 5 of 3.5, p. 62, Vol. 1):

THEOREM 1.2 *Every monoid can be faithfully represented as a monoid of mappings.*

Proof. Given a monoid M , we take the regular representation of M . If this is $x \mapsto \rho_x$, then ρ is a homomorphism from M to $\text{Map}(M)$, by what has been said, and if $\rho_x = \rho_y$, then $x = 1 \cdot \rho_x = 1 \cdot \rho_y = y$, hence the homomorphism is injective. ■

Let us return to a general monoid M and an M -set S . By Th. 1.1 we can write M as a homomorphic image of a free monoid X^* , for some set X . Thus we have a homomorphism

$$X^* \rightarrow M \rightarrow \text{Map}(S);$$

this shows that any set S with an M -action can also be regarded as a set with an X^* -action, where X corresponds to a generating set of M .

A free monoid has several remarkable properties, which can also be used to characterize it. A monoid M is called *conical* if $xy = 1 \Rightarrow x = y = 1$; M is said to have *cancellation* or be a *cancellation monoid* if for all $x, y \in M$, $xu = yu$ or $ux = uy$ for some $u \in M$ implies $x = y$. Further, M is *rigid* if it has cancellation and whenever $ac = bd$ there exists $z \in M$ such that either $a = bz$ or $b = az$. We observe that any free monoid is conical and rigid; the first property is clear from (2),

because the product in (2) cannot be 1 unless $r = s = 0$. To prove cancellation we note that in any element $x_1 \cdots x_r \neq 1$ the leftmost factor x_1 is unique, as is the rightmost factor x_r . Thus $x_1 \cdots x_r = y_1 \cdots y_s$ can hold only if $r = s$ and $x_i = y_i$ ($i = 1, \dots, r$). It follows that when $xu = yu$, say

$$x_1 \cdots x_r u_1 \cdots u_t = y_1 \cdots y_s u_1 \cdots u_t,$$

then both sides have the same length and $x_i = y_i$ ($i = 1, \dots, r = s$); therefore $x_1 \cdots x_r = y_1 \cdots y_r$; a similar argument applies when $ux = uy$. To prove rigidity, let $ac = bd$, say $a = x_1 \cdots x_r$, $b = y_1 \cdots y_s$, $c = u_1 \cdots u_h$, $d = v_1 \cdots v_k$. Then we have

$$x_1 \cdots x_r u_1 \cdots u_h = y_1 \cdots y_s v_1 \cdots v_k.$$

By symmetry we may assume that $r \leq s$; then $x_1 = y_1, \dots, x_r = y_r$, and hence $b = x_1 \cdots x_r y_{r+1} \cdots y_s = az$, where $z = y_{r+1} \cdots y_s$. This shows a free monoid to be rigid. We remark that when $ac = bd$, then $a = bz$ or $b = az$ according as $|a|$ is \geq or $\leq |b|$.

By a *unit* in a monoid M we understand an element u such that v exists in M satisfying $uv = 1$, $vu = 1$. For example, in a conical monoid the only unit is 1. When M has cancellation, it is enough to assume one of these equations, say $uv = 1$; for then $(vu)v = v(uv) = v1 = 1v$, hence $vu = 1$ by cancellation, and similarly if $vu = 1$. Let us define an *atom* as a non-unit which cannot be expressed as a product of two non-units (as in rings). For example, in a free monoid the atoms are just the elements of length 1. This shows incidentally that in a free monoid the free generating set is uniquely determined as the set of all atoms. We now have the following characterization of free monoids:

THEOREM 1.3 *Let F be a monoid and X the set of all its atoms. Then F is free, on X as free generating set, if and only if F is conical and rigid, and is generated by X .*

Proof. We have seen that in a free monoid these conditions are satisfied. Conversely, assume that they hold; we shall show that every element of F can be written in just one way as a product of elements of X . Any $a \in F$ can be expressed as such a product in at least one way, because X generates F . If we have

$$a = x_1 x_2 \cdots x_r = y_1 \cdots y_s, \quad x_i, y_j \in X,$$

then by rigidity, $x_1 = y_1 b$ or $y_1 = x_1 b$ for some $b \in F$, say the former holds. Since x_1, y_1 are atoms, b must be a unit and so $b = 1$ because F is conical. Thus $x_1 = y_1$ and we can cancel this factor and obtain $x_2 \cdots x_r = y_2 \cdots y_s$. By induction on $\max(r, s)$ we find $r - 1 = s - 1$, i.e. $r = s$ and $x_2 = y_2, \dots, x_r = y_r$. Thus F is indeed free on X , as we had to show. ■

Exercises

- (1) Show that every finite cancellation monoid is a group.

- (2) Let a, b be any elements of a monoid. Show that if ab and ba are invertible, then so are a and b , but this does not follow if we only know that ab is invertible. What can we say if aba is invertible?
- (3) Show that the additive monoid of non-negative rational numbers is conical and rigid, but not free.
- (4) Show that every finitely generated monoid which is conical and rigid is free. (*Hint.* Use Th. 1.1 with a minimal generating set.)
- (5) Show that a submonoid of a free monoid is free iff it is rigid. Give examples of submonoids of free monoids that are not free. (*Hint.* Consider the 1-generator case first.)
- (6) A set with an associative multiplication is called a *semigroup*. Show that any semigroup Σ may be embedded in a monoid by defining $\Sigma^1 = \Sigma \cup \{1\}$ with multiplication $x1 = 1x = x$ for $x \in \Sigma$.
- (7) A *zero* in a monoid is an element 0 such that $0x = x0 = 0$. Verify that (i) a monoid has at most one zero, (ii) every monoid M can be embedded in a monoid M_0 with zero. If M already contains a zero, how can the presence of these two zeros be reconciled with (i)?

11.2 Languages and grammars

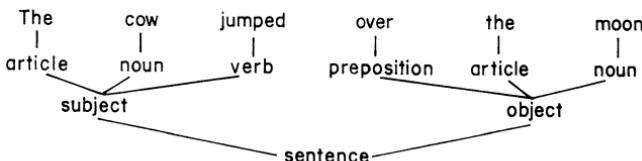
Algebraic language theory arose from the attempt by N. Chomsky to analyse and make precise the process of forming sentences in natural languages. The first point to notice is that whereas one may often test a sentence of a natural language like English for its grammatical correctness by checking its meaning, this is really irrelevant. This is well illustrated by Chomsky's example of a meaningless sentence which however is grammatically correct:

Colourless green ideas dream furiously.

To emphasize the point, he confronts it with a sentence which is not correct:

furiously ideas green colourless dream

In principle the analysis ('parsing') of a sentence consists in determining its constituents and checking that they have been put together according to prescribed rules:



The mathematical model consists of a set of rules of the form: $\text{sentence} \rightarrow \{\text{subject}, \text{object}\}$

verb}, noun → cow, etc., which will lead to all the sentences of the language and no others. This amounts to reading the above diagram from the bottom upwards.

In order to write our sentences we need a finite (non-empty) set X , our *alphabet*. As we have seen, the free monoid on X is the set X^* of all strings of letters from X , also called *words* in X (with a multiplication, which we ignore for the moment). By a *language* on X we understand any subset of X^* . Here we do not distinguish between words and sentences; we can think of an element of X^* as a message, with a particular symbol of X as a blank space, to separate the words of the message.

We single out a particular language by prescribing a set of rules according to which its sentences are to be formed. These rules constitute the grammar of the language and are formally defined as follows.

A *phrase structure grammar* or simply *grammar* G consists of three sets of data:

- (i) An *alphabet* X ; its letters are also called *terminal letters*.
- (ii) A set V of *clause-indicators* or *variables*, including a symbol σ for a complete sentence. We write $A = X \cup V$.
- (iii) A finite set of *rewriting rules*: $u \rightarrow v$, where $u, v \in A^*$ and u contains at least one variable, i.e. $u \notin X^*$.

A string of letters from A , i.e. a member of A^* is called *terminal* if it lies in X^* , *non-terminal* otherwise. The rewriting rule $u \rightarrow v$ is applied by replacing a string *fug* over A by *fvg*. To obtain a sentence in our language we start from σ and apply the rewriting rules until no variables (clause-indicators) are left. If the resulting terminal string is f , we write $\sigma \rightarrow \rightarrow f$ and call the sequence of rules applied a *derivation* of f , while f itself is a *sentence* of the language. In this way we obtain the language $L(G)$ generated by the given grammar; it consists of all the strings on X that are sentences of the language. A language is called *proper* if it does not include the empty word 1.

In giving examples, we shall use latin letters for the terminal letters and greek letters for the variables. With this convention it is not necessary to mention the alphabets X, V separately.

Examples

1. $L = \{x^n | n \geq 1\}$. Rules: $\sigma \rightarrow x$, $\sigma \rightarrow \sigma x$. We shall write this more briefly as $\sigma \rightarrow x$; σx . A typical derivation is $\sigma \rightarrow \sigma x \rightarrow \sigma x^2 \rightarrow \sigma x^3 \rightarrow x^4$. Similarly the language $\{x^{mr+ns} | m, n \geq 0\}$ is generated by the rules $\sigma \rightarrow \sigma x^r$; σx^s ; 1.
2. $L = \{xy^n | n \geq 0\}$, also written xy^* . Rules: $\sigma \rightarrow x$; σy .
3. $L = \{x^m y^n | m, n \geq 0\}$. Rules: $\sigma \rightarrow x\sigma$; σy ; 1.
4. $L = \{x^m y^n | 0 \leq m \leq n\}$. Rules: $\sigma \rightarrow x\sigma y$; σy ; 1.
5. $L = \{x^n z y^n | n \geq 0\}$. Rules: $\sigma \rightarrow x\sigma y$; z .
6. The *empty language* $L = \emptyset$ has the rule $\sigma \rightarrow \sigma$.
7. The *universal language* $L = X^*$ has the rules $\sigma \rightarrow \sigma x$; $1(x \in X)$.

This notion of language is of course too wide to be of use and one singles out certain classes of languages by imposing conditions on the generating grammar,

as follows. The classification below is known as the *Chomsky hierarchy*.

0. By a language of *type 0* or a *phrase structure language* we understand any language generated by a phrase structure grammar. By no means every language is of type 0; in fact, since the alphabet and the set of rewriting rules are finite, the set of all languages of type 0 is countable, whereas there are uncountably many languages on any alphabet, because an infinite set has uncountably many subsets. It can be shown that the languages of type 0 are precisely the recursively enumerable subsets of X^* (cf. e.g. M. Davis 1958).

1. A language is said to be of *type 1*, or *context-sensitive*, or a *CS-language* if it can be generated by a grammar in which all the rewriting rules are of the form

$$\alpha f g \rightarrow u f g, \quad \text{where } \alpha \in V, u \in A^+, f, g \in A^*, (A^+ = A^* \setminus \{1\}). \quad (1)$$

The grammar is then also called a *CS-grammar*. The rule (1) can be taken to mean: α is replaced by u in the context $f g$.

2. A language is said to be of *type 2*, or *context-free*, or a *CF-language* if it can be generated by a grammar with rewriting rules of the form

$$\alpha \rightarrow u, \quad \text{where } \alpha \in V, u \in A^*. \quad (2)$$

The grammar is also called a *CF-grammar*. The rule (2) means that α is replaced by u independently of the context in which it occurs.

3. A language is said to be of *type 3*, or *regular*, or *finite-state* if it can be generated by a grammar with rules of the form

$$\alpha \rightarrow x\beta, \quad \alpha \rightarrow 1, \quad \text{where } x \in X, \alpha, \beta \in V. \quad (3)$$

Again the grammar is called *regular*. Here α is replaced by a variable following a letter or by 1. Instead of writing the variable β on the right of the terminal letter we can also restrict the rules so as to have β on the left of the terminal letter throughout. It can be shown that this leads to the same class of languages (cf. Ex. (4), 11.3).

If \mathcal{L}_i ($i = 0, 1, 2, 3$) denotes the class of all proper languages of type i , then it is clear that

$$\mathcal{L}_0 \supseteq \mathcal{L}_1 \supseteq \mathcal{L}_2 \supseteq \mathcal{L}_3; \quad (4)$$

in fact the inclusions can all be shown to be strict, but in general it may not be easy to tell where a given language belongs, since there are usually many grammars generating it. Thus to show that a given language is context-free we need only find a CF-grammar generating it; but to show that a language is *not* context-free we must show that none of the grammars generating it is CF.

We note the following alternative definition of grammars of type 1, 2:

PROPOSITION 2.1 *Let G be a grammar with alphabets X, V . Then*

- (i) *If G is a CS-grammar, then for every rule $u \rightarrow v$ in G ,*

$$|u| \leq |v|. \quad (5)$$

(ii) If G is a CF-grammar, then for every rule $u \rightarrow v$ in G ,

$$|u| = 1. \quad (6)$$

Conversely, if G satisfies (5), (6) resp., then there is a CS-grammar, resp. a CF-grammar generating $L(G)$.

Proof. (i) Let G be a CS-grammar; any rule in G has the form $f\alpha g \rightarrow f\beta g$, where $u \neq 1$, hence $|u| \geq 1$ and so $|f\beta g| \geq |f| + 1 + |g| = |f\alpha g|$, and (5) follows. Conversely, when (5) holds for every rule, we can achieve the effect of $u \rightarrow v$ by replacing the letters in u one at a time by a letter in v , taking care to leave a (new) variable until last. To give a typical example, if $u = u_1\alpha u_2 u_3$, $v = v_1 \cdots v_5$, we replace $u \rightarrow v$ by the rules $u_1\alpha u_2 u_3 \rightarrow v_1\beta u_2 u_3 \rightarrow v_1\beta u_2 v_5 \rightarrow v_1\beta v_4 v_5 \rightarrow v$, where β does not occur elsewhere. (ii) It is clear from the definition of a CF-language that its rules are characterized by (6); the details may be left to the reader. ■

Sometimes one may wish to include the empty word in a proper language; this is most easily done by replacing any occurrence of σ on the right of a rule by a new variable λ , say, for each rule with σ on the left add the same rule with σ replaced by λ on the left and adding the rule $\sigma \rightarrow 1$. For example, to generate the language $\{x^n z y^n, 1 \mid n \geq 0\}$ we modify the example 5 above: $\sigma \rightarrow x\lambda y; 1, \lambda \rightarrow x\lambda y; z$. If we just add $\sigma \rightarrow 1$ to the rules of 5, we would also get xy .

From an improper CF-language L we can obtain a proper CF-language $L \setminus \{1\}$ by replacing in any CF-grammar for L , any rule $\alpha \rightarrow 1$ by $\beta \rightarrow \bar{u}$, where \bar{u} runs over all words obtained from derivations of the form, $\beta \rightarrow \bar{u}$, where u contains α , by replacing α in u by 1. For example, the language $\{x^m y^n \mid m + n > 0\}$ is given by $\sigma \rightarrow x\sigma; \sigma y; y; x$.

Looking at the examples given earlier, we see that 1, 2 and 3 are regular, as well as 6 and 7. Examples 4 and 5 are context-free, but not regular, as we shall see in 11.3. We conclude with an example of a CS-language which is not context-free, as Prop. 3.6 will show.

Example

8. $\{x^n z^n y^n \mid n \geq 0\}$ has the generating grammar $\sigma \rightarrow x\sigma\lambda\mu; xz\mu; 1, \mu\lambda \rightarrow \lambda\mu, z\lambda \rightarrow z^2, \mu \rightarrow y$. The first two rules generate all the words $x^n z \mu (\lambda\mu)^{n-1}$, the fourth moves the λ 's past the μ 's next to z and the next replaces each λ by z . Finally each μ is replaced by y . To obtain the same language without 1 we simply omit the rule $\sigma \rightarrow 1$.

Exercises

- (1) Find a regular grammar to generate the set of all words of even length in X .
- (2) Show that each finite language is regular.
- (3) Show that if L, L' are any languages of type i ($= 0, 1, 2$ or 3), then so is $L \cup L'$, $LL' = \{uv \mid u \in L, v \in L'\}$ and L^{op} obtained from L by writing each word in reverse order.

- (4) Show that if L is regular, then so is L^* , the language whose words are all the finite strings of words from L .
- (5) Show that regular languages form the smallest class containing all finite languages and closed under union, product and $*$.
- (6) Show that every context-free language can be generated by a CF-grammar G with the property: for each non-terminal variable α there is a derivation $\alpha \rightarrow u$ ($u \in X^*$) and for each terminal letter x there is a derivation $\alpha \rightarrow u$, where u includes x .
- (7) Show that for any CF-grammar $G = (X, V)$ there is a CF-grammar G' producing the same language as G such that (i) G' contains no rule $\alpha \rightarrow \beta$, where $\alpha, \beta \in V$, (ii) if $L(G)$ is improper, then G' contains the rule $\alpha \rightarrow 1$ but no other rules with 1 on the right-hand side and (iii) no rule has σ occurring on the right. Thus all rules of G' have the form $\sigma \rightarrow 1$, $\alpha \rightarrow x$ or $\alpha \rightarrow f$, where $f \in (X \cup V \setminus \{\sigma\})^+$, $|f| \geq 2$.
- (8) Show that for a given CF-grammar G there exists a CF-grammar G' producing the same language as G , with rules $\alpha \rightarrow xf$, $f \in V^*$ and possibly $\sigma \rightarrow 1$ (Greibach normal form).

11.3 Automata

Logical machines form a convenient means of studying recursive functions. In particular, Turing machines lead precisely to recursively enumerable sets, and so correspond to grammars of type 0, as mentioned earlier. These machines are outside the scope of this book and will not be discussed further, but we would expect the more restricted types 1–3 of grammars to correspond to more special machines. This is in fact the case and in this section we shall define the types of machines corresponding to these grammars and use them to derive some of their properties.

A *sequential machine* M is given by three sets and two functions describing its action. There is a set S of states as well as two finite alphabets: an *input* X and an *output* Y . The action is described by a *transition function* $\delta: S \times X \rightarrow S$ and an *output function* $\lambda: S \times X \rightarrow Y$. To operate the machine we start from a given state s and input x ; then the machine passes to the state $\delta(s, x)$ and produces the output $\lambda(s, x)$. In general the input will not just be a letter but a word w on X . The machine reads w letter by letter and for each letter x in w gives out $y \in Y$ according to the output function λ , while passing through the different states in accordance with the transition function δ . The output is thus a word on Y , of the same length as w , obtained as follows. Define mappings $\delta': S \times X^* \rightarrow S$, $\lambda': S \times X^* \rightarrow Y^*$ by the equations

$$\delta'(s, 1) = s, \quad \delta'(s, ux) = \delta(\delta'(s, u), x) \quad s \in S, x \in X, u \in X^*, \quad (1)$$

$$\lambda'(s, 1) = 1, \quad \lambda'(s, ux) = \lambda'(\delta'(s, u), \lambda(\delta'(s, u), x)). \quad (2)$$

These equations define δ' , λ' by induction on the length of words. It is clear that δ' , λ' extend δ , λ respectively and so we may without risk of confusion omit the

primes from δ' , λ' . From (1) it is clear that

$$\delta(s, 1) = s, \quad \delta(s, uv) = \delta(\delta(s, u), v) \quad s \in S, u, v \in X^*, \quad (3)$$

so the mapping δ just defines an action of the free monoid X^* on S . We note that this holds even though no conditions were imposed on δ .

It is clear from this definition that a machine is completely specified by the set of all quadruples of the form $(s, x, \lambda(s, x), \delta(s, x))$. Sometimes it is preferable to start from a more general notion. Let S, X, Y be as before and define an *automaton* A as a set of quadruples

$$P = P(A) \subseteq S \times X \times Y \times S. \quad (4)$$

The members of P are called its *edges*; each edge (s, x, y, s') has an initial state s , input x , output y and final state s' . For a sequential machine each pair $(s, x) \in S \times X$ determines a unique edge (s, x, y, s') and whenever our set P of edges is such that

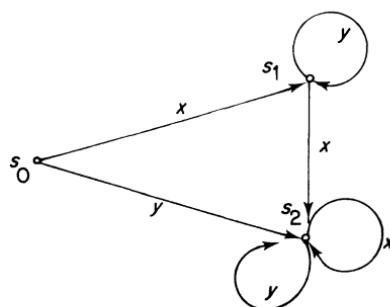
- C for each pair $(s, x) \in S \times X$ there exist a unique $y \in Y$ and $s' \in S$ such that $(s, x, y, s') \in P$,

then we can define λ, δ by writing $y = \lambda(s, x)$, $s' = \delta(s, x)$ and we have a sequential machine. Two edges are *consecutive* if the final state of the first edge is also the initial state of the second. By a *path* for A we understand a sequence $w = (u_1, \dots, u_n)$ of consecutive edges

$$u_i = (s_{i-1}, x_i, y_i, s_i).$$

Its *length* is n , s_0 is the initial and s_n the final state, $x_1 \dots x_n$ its *input label* and $y_1 \dots y_n$ its *output label*. It is clear how an automaton can be represented by a graph with the set S of states as vertex set, each edge being labelled by its input and output. Sometimes one singles out two subsets I, F of S ; a path is called *successful* if its initial state is in I and its final state in F . The set $L(A)$ of all input labels of successful paths is a subset of X^* , called the *behaviour* of A , or also the set *accepted* by A . We note that the output does not enter into the behaviour; when the output is absent (so that P now consists of triples (s, x, s')), A is called an *acceptor*. As an example consider the acceptor with states s_0, s_1, s_2 , input $X = \{x, y\}$ and transition function

δ	x	y
s_0	s_1	s_2
s_1	s_2	s_1
s_2	s_2	s_2



The graph is as shown. If $I = \{s_0\}$, $F = \{s_1\}$, then the behaviour is xy^* ; for $I = \{s_0\}$, $F = \{s_2\}$ the behaviour is $xy^*xX^* \cup yX^*$.

An automaton is said to be *complete* if it satisfies condition C above (so that we have a sequential machine), and the set I of initial states consists of a single state. To operate a complete acceptor we take any word in X^* and use it as input with the machine in state I ; this may or may not lead to a successful path, i.e. a path ending in F . We shall be interested in its behaviour, i.e. the set of input labels corresponding to successful paths. For example, the above example is a complete acceptor; we note how its graph makes it very easy to compute its behaviour.

In constructing an acceptor it is usually convenient not to demand completeness, although complete acceptors are easier to handle. Fortunately there is a reduction allowing us to pass from one to the other.

PROPOSITION 3.1 *For each acceptor A there is a complete acceptor B with the same behaviour. If A is finite (i.e. with finite set of states), then so is B.*

Proof. Let the set of states for A be S , with initial state set I and final state set F . We take B to be on the same alphabet as A , with state set the set of all subsets of S , initial state $\{I\}$ and final set of states all sets meeting F . The transition function for B is given by $\delta(U, x) = V$, where V consists of all states v such that (u, x, v) is an edge in A for some $u \in U$. It is clear that B has the same behaviour as A , and it is easily seen to be complete. ■

We remark that every subset Y of X^* is the behaviour of some acceptor; we take X^* as state set, 1 as initial state and Y as final set of states, with right multiplication as the transition function. Our aim is to describe the behaviour of finite acceptors; we shall find that this consists precisely of all regular languages. To prove this result we shall need to construct a ‘minimal’ acceptor for a given language.

We shall use the notation (S, i, F) for a complete acceptor, where S is the set of states, i is the initial state and F is the final set of states; the alphabet is usually denoted by X and so will not be mentioned explicitly, and the transition function is indicated by juxtaposition; thus instead of $\delta(s, x) = s'$ we write $sx = s'$. A state s in an acceptor $A = (S, i, F)$ is said to be *accessible* if there is a path from i to s , *coaccessible* if there is a path from s to a state in F . As far as the behaviour of A is concerned, we can clearly neglect any states that are not both accessible and coaccessible. If every state of A is both accessible and coaccessible, then A is called *trim*.

Given two acceptors A, A' with state sets S, S' , we define a *state homomorphism* from A to A' as a map $f : S \rightarrow S'$ such that $(s, x, t) \in A \Rightarrow (sf, x, tf) \in A'$. If f has an inverse which is also a state homomorphism, f is called an *isomorphism*. To give an example, let us put

$$L_s = \{v \in X^* | s.v \in F\};$$

thus L_s consists of the set of words which give a successful path with s as initial state. Two states s, t are called *separable* if $L_s \neq L_t$, *inseparable* otherwise. By identifying any pairs of inseparable states we obtain a state homomorphism to an acceptor with fewer states. If any two distinct states are separable, the acceptor is said to be *reduced*. Now it is clear that from any acceptor we can find a state homomorphism to a reduced acceptor by identifying all pairs of inseparable states. Thus every acceptor (finite or infinite) has a homomorphic image which is reduced; moreover, the behaviour of this image is the same as that of the original acceptor.

For every subset Y of X^* we can define a reduced acceptor $A(Y)$ whose behaviour is Y . The states of $A(Y)$ are the non-empty sets

$$u^{-1}Y = \{v \in X^* \mid uv \in Y\},$$

where u ranges over X^* . The initial state is $1^{-1}Y = Y$ and the final states are the states $u^{-1}Y$ containing 1. The transition function is defined by

$$Z.u = u^{-1}Z, \quad \text{where } u \in X. \quad (5)$$

This is a partial function (i.e. not everywhere defined) since $u^{-1}Z$ may be empty, but it is single-valued. If for u we take a word in X , we have, by induction on the length of u ,

$$w \in Z.ux \Leftrightarrow (Z.u)x \Leftrightarrow xw \in u^{-1}Z \Leftrightarrow uxw \in Z.$$

This shows that (5) holds for any $u \in X^*$. As a consequence we have

$$w \in L(A(Y)) \Leftrightarrow 1 \in Y.w \Leftrightarrow w \in Y,$$

showing that the behaviour of $A(Y)$ is indeed Y . We shall call $A(Y)$ the *minimal acceptor* for Y ; its properties follow from

THEOREM 3.2 *Let $A = (S, i, F)$ be a trim acceptor and put $Y = L(A)$, the behaviour of A . Then there is a state homomorphism $\varphi: A \rightarrow A(Y)$ to the minimal acceptor for Y which is surjective on states, given by*

$$\varphi: s \mapsto L_s = \{v \in X^* \mid s.v \in F\}. \quad (6)$$

Proof. Since A is trim, any state s in S is accessible, so $iu = s$ for some $u \in X^*$; further, s is coaccessible, so $sv \in F$ for some $v \in X^*$ and it follows that L_s defined by (6) is non-empty. Thus φ is well-defined. To show that it is a homomorphism we have to verify that when $s.x = t$, then $L_s.x = L_t$. But we have

$$w \in L_s.x \Leftrightarrow xw \in L_s \Leftrightarrow s.xw \in F \Leftrightarrow w \in L_t;$$

so φ is indeed a homomorphism. It is surjective, because if $u^{-1}Y \neq \emptyset$, then there is a successful path in A with label uv , where $v \in Y$. Now

$$v \in u^{-1}Y \Leftrightarrow uv \in Y \Leftrightarrow sv = iuv \in F \Leftrightarrow v \in L_s;$$

thus $u^{-1}Y = L_s$ and this shows (6) to be surjective. ■

As we have seen, for any subset Y of X^* there is a reduced acceptor with behaviour Y ; taking this to be A in Th. 3.2, we find φ in this case to be an isomorphism, by the definition of 'reduced'. It follows that $A(Y)$ must be reduced:

COROLLARY 3.3 *The minimal acceptor for any subset of X^* is reduced.* ■

Of course this is also not hard to verify directly.

We can now establish

THEOREM 3.4 *A language is regular if and only if it is the precise set accepted by a finite acceptor.*

Proof. Let $A = (S, i, F)$ be a finite acceptor with behaviour Y and write the transition function as δ for clarity. As our grammar $G = \{X, V, \rightarrow\}$, we take X to be the input of A and $V = S$, the set of states, with $\sigma = i$, the initial state. For each state α in S and $x \in X$ we include in G the rule

$$\alpha \rightarrow x\beta \quad \text{if } \delta(\alpha, x) = \beta, \tag{7}$$

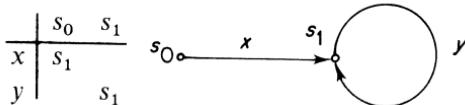
and for each final state ω we include the rule $\omega \rightarrow 1$. Given any word $w = x_1 \cdots x_r$, let us put $\delta(i, x_1) = s_1, \dots, \delta(s_{r-1}, x_r) = s_r$. Then the rules (7) include $i \rightarrow x_1 s_1, \dots, s_{r-1} \rightarrow x_r s_r$, hence $\sigma \rightarrow x_1 s_1 \rightarrow x_1 x_2 s_2 \rightarrow \cdots \rightarrow x_1 \cdots x_r s_r$. If $s_r \in F$, then $s_r \rightarrow 1$ and $x_1 \cdots x_r$ is included in $L(G)$. On the other hand, if $x_1 \cdots x_r \in L(G)$, consider the rules in G : they are all of the form $\alpha \rightarrow x\beta$ or $\alpha \rightarrow 1$ and the number of variables is constant in the application of the former rule and decreases by 1 when the latter is applied. Thus any derivation of $x_1 \cdots x_r$ must be of the form $i \rightarrow x_1 s_1, s_1 \rightarrow x_2 s_2, \dots, s_{r-1} \rightarrow x_r s_r, s_r \rightarrow 1$. This means that $\delta(s_{k-1}, x_k) = s_k$ ($k = 1, \dots, r$) and $s_r \in F$, so $x_1 \cdots x_r$ is accepted by A .

Conversely, let G be a regular grammar, with derived language $L(G)$. For our acceptor A we take the alphabet X of G as input and the set V of variables as state set, with σ as initial state and a triple (α, x, β) for each rule $\alpha \rightarrow x\beta$, while the final state set consists of all α such that $\alpha \rightarrow 1$. Then it is clear that the derivations of G correspond precisely to the successful paths in A ; the details may be left to the reader. Hence $L(G)$ is the set accepted by A . ■

The acceptor constructed in the proof may not be complete, but it is trim provided that any superfluous variables have been removed from G . It follows by Th. 3.2 that for a regular language the minimal acceptor is finite. This provides a practical way of determining whether a language is regular: Y is a regular language iff its minimal acceptor $A(Y)$ is finite. We illustrate this result by some examples.

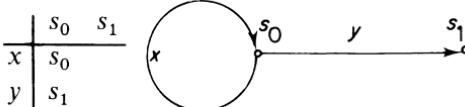
Examples.

1. $\{xy^n \mid n \geq 0\}$. The minimal acceptor has states $s_0 = xy^*$ and $s_1 = y^*$, and its operation is given by the table:

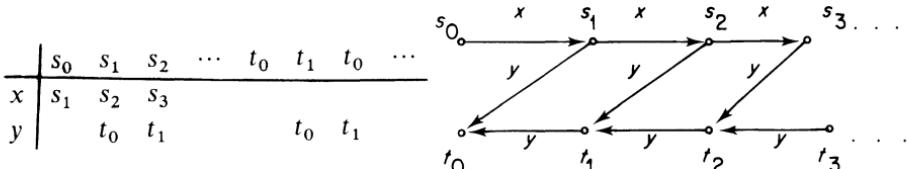


The initial state is s_0 and the final state is s_1 . The behaviour is easily read off from the graph.

2. $\{x^n y \mid n \geq 0\}$. Here the states are $s_0 = x^* y$ and $s_1 = \{1\}$, with initial state s_0 and final state s_1 :



3. $\{x^n y^n \mid n \geq 0\}$. We have the states $s_0 = \{x^n y^n \mid n \geq 0\}$, $s_1 = \{x^n y^{n+1} \mid n \geq 0\}$, $s_2 = \{x^n y^{n+2} \mid n \geq 0\}$, ..., $t_0 = \{1\}$, $t_1 = \{y\}$, $t_2 = \{y^2\}$, The initial state is s_0 and the final states are s_0, t_0 , while the operations are given by:



It should be clear from these examples how the behaviour of an acceptor may be read off from its graph. We note that in none of the cases is the acceptor complete; the transition function, though single-valued, is not everywhere defined. But this does not impair its usefulness; in any case we could replace it by a complete acceptor, using Prop. 3.1. Sometimes it is easier to test for regularity by means of the following necessary condition.

PROPOSITION 3.5 *Let L be an infinite regular language on X . Then there exist $w', y, w'' \in X^*$, $y \neq 1$, such that $w'y^n w'' \in L$ for $n = 1, 2, \dots$.*

Proof. Since L is regular, there is an acceptor A for L , with a finite number, n say, of states. The language L is infinite on a finite alphabet, so it contains a word w of length $> n$. The acceptor reads w letter by letter, moving from state to state as it reads. Since there are more than n steps, two of these states must be the same; say s_1 occurs twice. If y is the portion read between the first time and the second time in s_1 , then $w = w'yw''$ and $y \neq 1$. Clearly our machine will also accept $w'y^2w''$, and generally $w'y^n w''$; thus $w'y^n w'' \in L$. ■

For context-free languages there is a condition similar to that in Prop. 3.5; this is sometimes known as the *pumping lemma*:

PROPOSITION 3.6 *Let G be a CF-grammar. Then there exist integers p, q such that every word w of length greater than p in $L(G)$ can be written as $w = w'uzvw''$, where $uv \neq 1$, $|uv| \leq q$ and $w'u^n zv^n w'' \in L(G)$ for all $n \geq 1$.*

Proof. The rules of G are all of the form $\alpha \rightarrow t$, $t \in A^*$. Let the number of variables be k and choose p so large that every word $w \in L(G)$ of length $\geq p$ has more than k steps in its derivation. Each of these steps has the form $\alpha \rightarrow t$; hence some variable occurs twice, so the part of the derivation from the first to the second occurrence of α reads $\alpha \rightarrow \dots \rightarrow u\alpha v$, where $uv \neq 1$. It follows that $w = w'uzvw''$, where α occurs only once in the derivation $\alpha \rightarrow \dots \rightarrow z$, which therefore has at most k steps. Therefore uv has bounded length, and by repeating the steps between α and $u\alpha v$ we obtain $w'u^n zv^n w''$ for all $n \geq 1$. ■

Proposition 3.5 shows that the language $\{x^nzy^n\}$ which we saw to be context-free, is not regular, and Prop. 3.6 shows that $\{x^n y^n z^n\}$ is not context-free.

Let us return to the example $\{x^nzy^n\}$; we have just seen that it is not regular, and so cannot be obtained from a finite acceptor. Intuitively we can see that a finite acceptor does not have the means of comparing the exponents of x and y . To make such a comparison requires a memory of some kind, and we shall now describe a machine with a memory capable of accepting CF-languages. The memory to be described is of a rather simple sort, a ‘first-in, last-out’ store, where we only have access to the last item in the store.

A *pushdown acceptor* (PDA for short) is an acceptor which in addition to its set of states S and input alphabet X has a set Σ of *store symbols*, with initial symbol λ_0 and a transition function $\delta: S \times X \times \Sigma \rightarrow S \times \Sigma^*$; but for a given triple of arguments there may be several or no values. At any stage the machine is described by a triple (s_i, w, α) , where $s_i \in S$, $w \in X^*$, $\alpha \in \Sigma^*$. We apply δ to the triple consisting of s_i , the first letter of w and the last letter of α . If $w = xw'$, $\alpha = \alpha'\lambda$ say, and (s_j, β) is a value of $\delta(s_i, x, \lambda)$, then

$$(s_i, xw', \alpha'\lambda) \rightarrow (s_j, w', \alpha'\beta)$$

is a possible move. Thus the effect of δ is to move into state s_j , remove the initial factor x from w and replace the final letter λ of α by β . We say that a word w on X is *accepted* by the machine if, starting from (s_0, w, λ_0) , there is a series of moves to take us to $(s_f, 1, \gamma)$, where s_0 is the initial and s_f a final state. With this definition we have

THEOREM 3.7 *The context-free languages constitute the precise class of sets accepted by pushdown acceptors.* ■

We shall not give the proof here (cf. Arbib 1969), but as an example we describe a PDA for $\{x^n y^n \mid n \geq 0\}$. Its states are s_0, s_1, s_2 , where s_0 is initial and s_2 final. The store symbols are λ (initial symbol), μ, v . We give the values for $\delta(s_i, \dots)$ in the form of a table for each s_i :

s_0	λ	μ	v	s_1	λ	μ	v
x	$s_0\mu$	$s_0\mu v$	s_0v^2	x			
y		$s_2\mu$	s_1	y		$s_2\mu$	s_1

Blanks and the remaining values (for s_2) are undefined. To see how $x^n y^n$ is accepted, but no other strings, we note how the store acts as a memory, remembering how many factors x have been taken off. If we think of the store as arranged vertically, at each stage we remove the topmost symbol and add a number of symbols to the top, rather like a stack of plates in a cafeteria; this explains the name.

By a somewhat more elaborate process, with a tape on which the input is written (a ‘linear-bounded’ automaton) one can devise a class of machines which accept precisely all the CS-languages (cf. Landweber 1963). These machines are more special than Turing machines in that their tape length is bounded by a linear function of the length of the input word.

Finally we note the following connexion with monoids:

THEOREM 3.8 *A language L on X is regular if and only if there is a homomorphism $f:X^* \rightarrow M$ to a finite monoid M such that L is the complete inverse image of a subset N of M : $L = f^{-1}(N)$.*

Proof. Given a regular language L on X , we have a finite acceptor A which accepts precisely L . Now A defines an action of X^* on the set S of states of A . Thus we have a homomorphism $f:X^* \rightarrow \text{Map}(S)$. If P is the subset of $\text{Map}(S)$ of all mappings taking s_0 into the set F of final states, then $w \in L$ iff $wf \in P$; thus $L = f^{-1}(P)$, and by definition $\text{Map}(S)$ is a finite monoid. Thus the condition is satisfied.

Conversely, given $f:X^* \rightarrow M$ with $L = f^{-1}(N)$ for some $N \subseteq M$, we consider the acceptor A consisting of the alphabet X , state set M , action $\delta(a, x) = a.(xf)$, ($a \in M, x \in X$), with neutral element 1 as initial state and N as set of final states. The language accepted by A is just L . ■

Exercises

- (1) Find the languages generated by the following grammars: (i) $\sigma \rightarrow \sigma^2; x; y$, (ii) $\sigma \rightarrow \sigma^3; x; y$, (iii) $\sigma \rightarrow \sigma^2; x\sigma y; y\sigma x; 1$, (iv) $\sigma \rightarrow x\sigma x; xyx; x$.
- (2) Find a CF-grammar on x, y generating the set of all words in which x is immediately followed by y .

- (3) Find a grammar on x, y generating the set of all words in which each left factor has at least as many x 's as y 's. Is this a CF-language?
- (4) A grammar with rules of the form $\alpha \rightarrow x, \alpha \rightarrow \beta x$ is sometimes called *left regular*, while a grammar with rules of the form $\alpha \rightarrow x, \alpha \rightarrow x\beta$ is called *right regular*. Show that every language generated by a left regular grammar can also be generated by a right regular grammar. (*Hint.* Interpret the words of the language as circuits in the acceptor graph; a left and a right regular grammar correspond to the two senses of traversing these loops.)
- (5) Show that every CF-language in one letter is regular.
- (6) Show that a language on a one-letter alphabet $\{x^n | n \in I\}$ is regular iff the set I of exponents is ultimately periodic.
- (7) Construct a PDA for the set of palindromes with 'centre marker', i.e. $L(G)$, where $G: \sigma \rightarrow x\sigma x; y\sigma y; z$. (*Hint.* Put one half of the word in store and then match the other half.)
- (8) Find a PDA for the set of even palindromes $G: \sigma \rightarrow x\sigma x; y\sigma y; 1$. (*Hint.* Construct a PDA to 'guess' the centre.)
- (9) An automaton A is called *deterministic* (resp. *total*) if for each $s \in S, x \in X$ there exists at most (resp. at least) one pair $y \in Y, s' \in S$ such that $(s, x, y, s') \in P(A)$. For any A define its *reverse* A° as the automaton with state set S , input Y , output X and $P(A^\circ)$ as the set of all (s', y, x, s) such that $(s, x, y, s') \in A$. Show that A° is deterministic whenever A is reduced.
- (10) A complete automaton A with N states s_1, \dots, s_N can be described by a set of $N \times N$ matrices $P(x|y)$ ($x \in X, y \in Y$) where the (i, j) -entry of $P(x|y)$ is 1 if $\delta(s_i, x) = y$ and $\lambda(s_i, x) = s_j$, and 0 otherwise. Define $P(u|v)$ recursively for $u \in X^*, v \in Y^*$ by $P(ux|vy) = P(u|v)P(x|y)$, $P(u|v) = 0$ if $|u| \neq |v|$. Show that $P(uu'|vv') = P(u|v)P(u'|v')$. Further put $P(x) = \sum_y P(x|y)$ and write π for the row vector whose i th component is 1 if s_i is the initial state and 0 otherwise; further write f for the column vector with component 1 for a final and 0 for a non-final state. Show that for any $u \in X^*$, $\pi P(u)f$ is 1 if u is accepted and 0 otherwise.
- (11) With the notation of Ex. (10), put $T = \sum_x P(x)$; verify that the (i, j) -entry of T^n is the number of words of length n in X which give a path from s_i to s_j . Show that if $\lambda(n)$ denotes the number of words of length n accepted, then the length generating function $L(t) = \sum \lambda(n)t^n$ satisfies $L(t) = \pi(I - tT)^{-1}f$. Use the characteristic equation of T to find a recursion formula for $\lambda(n)$.

11.4 Variable-length codes

The codes studied in Ch. 10 were block codes, where each code word has the same length. However, in practice different letters occur with different frequencies and an efficient code will represent the more frequently occurring letters by the shorter code words; e.g. in Morse code, the letters e, t which are among the most commonly occurring are represented by a dot and a dash respectively. For this

reason it is of interest to have codes in which the code words have varying lengths. In this section we shall describe such codes; the main problem is to design the code so as to ensure that messages can be uniquely decoded. Of course we can only take the first steps in the subject, but these will include results which are of interest and importance in general coding theory, besides leading to a better understanding of free monoids and free algebras.

Let $X = \{x_1, \dots, x_r\}$ be our alphabet and X^* the free monoid on X . Any subset A of X^* generates a submonoid, which we write as $\langle A \rangle$. By a *code* on X we understand a subset A of X^* such that every element of $\langle A \rangle$ can be uniquely factorized into elements of A . Thus if $w \in \langle A \rangle$ and

$$w = a_1 \cdots a_m = b_1 \cdots b_n, \quad a_i, b_j \in A,$$

then $m = n$ and $a_i = b_i$, $i = 1, \dots, n$.

For example, X itself is a code; more generally, X^n , the set of all products of length n (for any given $n \geq 1$), is a code. Further, any subset of a code is a code. The set $\{x, xy, yx\}$ is *not* a code, because the word $xyx = xy \cdot x = x \cdot yx$ has two distinct factorizations.

Our first problem is how to recognize codes. If A is *not* a code, then we have an equality between two distinct words in A , and by cancellation we may take this to be of the form

$$au = bv, \quad \text{where } a, b \in A, u, v \in \langle A \rangle.$$

If we express everything in terms of X we find by the rigidity of X^* ,

$$\text{either } a = bz \quad \text{or} \quad b = az, \quad \text{for some } z \in X^*.$$

If $a = bz$ say, we say that b is a *prefix* of a . Let us define a *prefix set* as a subset of X^* in which no element is a prefix of another. For example, $\{1\}$ is a prefix set; any prefix set $A \neq \{1\}$ cannot contain 1, because 1 is a prefix of any other element of X^* .

What we have just found is that if A is not a code (and $A \neq \{1\}$), then A is not a prefix set; thus we have

PROPOSITION 4.1 *Every prefix set $\neq \{1\}$ is a code.* ■

By a *prefix code* we shall understand a prefix set $\neq \{1\}$. To give an example, $\{y, xy, x^2y, x^3y\}$ is a prefix set and hence a prefix code. We can think of this example as the alphabet 1, x , x^2 , x^3 with y as place marker.

By symmetry we define a *suffix set* as a subset of X^* in which no element is a suffix, i.e. right-hand factor, of another. Now the left-right symmetry of the notion of code shows that every suffix set $\neq \{1\}$ is a code; such a code will be called a *suffix code*. Since there exist suffix codes which are not prefix, e.g. $\{x, xy\}$, we see that the converse of Prop. 4.1 is false. In fact there exist procedures for determining whether a given subset of X^* is a code, but they are quite lengthy and

will not be given here (the Sardinas–Patterson algorithm; cf Lallement 1979, Berstel and Perrin 1985). In any case, prefix codes are of particular interest in coding theory, since any message in a prefix code can be deciphered reading letter-by-letter from left to right (it is a ‘zero-delay’ code). This property actually characterizes prefix codes, for if a code is not prefix, say $a = bz$, where a, b are both code words, then at any occurrence of b in a message we have to read past this point to find out whether a or b is intended.

On any monoid M we can define a preordering by left divisibility:

$$u \leq v \Leftrightarrow v = uz \quad \text{for some } z \in M. \quad (1)$$

Clearly this relation is reflexive and transitive; we claim that when M is conical, with cancellation, then ‘ \leq ’ is antisymmetric, so that we have a partial ordering. For if $u \leq v$, $v \leq u$, then $v = uz$, $u = vz'$, hence $v = uz = vz'z$, so $z'z = 1$ by cancellation, and since M is conical, we conclude that $z = z' = 1$.

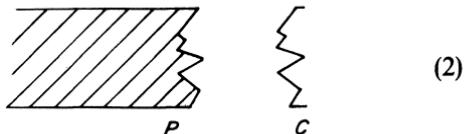
We shall be particularly interested in the ordering (1) on free monoids. In that case the set of left factors of any element u is totally ordered, by rigidity, and since the length of chains of factors is bounded by $|u|$, the ordering satisfies the minimum condition. In terms of the ordering (1) on a free monoid, a prefix set is just an anti-chain, and by Prop. 2.2.13 there is a natural bijection between anti-chains and lower segments. Here a ‘lower segment’ is a subset containing with any element all its left factors; such a set, if non-empty, is called a *Schreier set*; it is clear that every Schreier set contains 1.

Let us describe explicitly the correspondence of Prop. 2.2.13, as applied to prefix sets and Schreier sets: if C is a prefix set in X^* , then the corresponding Schreier set is the complement of CX^* in X^* ; for a Schreier set P the corresponding prefix set is the set of minimal elements in the complement of P .

A prefix set C is said to be *right large* if CX^* meets wX^* for every $w \in X^*$. By the rigidity of X^* this just amounts to saying that every element of X^* is comparable with some element of C . Hence the Schreier set corresponding to a right large prefix set C consists precisely of the proper left factors of elements of C . To sum up these relations we need one more definition. In any monoid a product AB of subsets A, B is said to be *unambiguous* if each element c of AB can be written in just one way as $c = ab$, where $a \in A$, $b \in B$.

PROPOSITION 4.2 *Let X^* be the free monoid on a finite alphabet X . Then there is a natural bijection between prefix sets and Schreier sets: to each prefix set C corresponds $P = X^* \setminus CX^*$ and to each Schreier set P corresponds the set C of minimal elements in $X^* \setminus P$, and we have the unambiguous product*

$$X^* = C^*P.$$



(2)

Moreover, P is finite if and only if C is finite and right large.

Proof. The description of each of P , C in terms of the other follows by Prop. 2.2.13. Thus C is the set of minimal elements in $X^* \setminus P$; since $1 \in P$, any such minimal element must have the form px ($p \in P$, $x \in X$), and moreover, $px \notin P$. Hence

$$C = PX \setminus P = \{px \mid p \in P, x \in X, px \notin P\}. \quad (3)$$

Now to establish (2) take $w \in X^*$; either $w \in P$ or w has a maximal proper left factor p in P . In the latter case $w = pxu$, where $x \in X$ and $px \notin P$; therefore $px \in C$ by (3). Further $|u| < |w|$, so by induction on the length, $u \in C^*P$, hence $w \in C^*P$ and (2) follows. Now if

$$w = u_1 \cdots u_r p = v_1 \cdots v_s q, \quad \text{where } p, q \in P, u_i, v_j \in C,$$

then since C is prefix, $u_1 = v_1$; hence we can cancel u_1 and conclude by induction on r that $r = s$, $u_i = v_i$, $i = 1, \dots, r$, $p = q$. This shows (2) to be unambiguous.

Suppose that P is finite; then so is C , by (3). Given $w \in X^*$, either $w \in P$; then some right multiple of w is not in P , because the length of elements in P is bounded. At least such element c is a right multiple of w in C , so $w < c$. Or $w \notin P$; then since $1 \in P$, there is a minimal left factor c of w not in P and this is again in C , so $c \leq w$. This shows C to be right large.

Conversely, suppose that C is finite right large and let P be the corresponding Schreier set. Given $w \in X^*$, either $w \geq c$ or $w < c$ for some $c \in C$, and the first alternative is excluded for members of P . Thus P consists of all proper left factors of elements of C and this is again a finite set. ■

For a closer study of codes it is useful to have a numerical measure for the elements of X^* . By a *measure* on X^* we understand a homomorphism μ of X^* into the multiplicative monoid of positive real numbers, such that

$$\sum_{x \in X} \mu(x) = 1. \quad (4)$$

Clearly $\mu(1) = 1$ and the values of μ on X can be assigned arbitrarily as positive real numbers, subject only to (4); once this is done, μ is completely determined on X^* by the homomorphism property. For example, writing $m(x) = r^{-1}$, we obtain the *uniform measure* on X^* :

$$m(w) = r^{-|w|}.$$

Any measure μ on X^* can be extended to subsets by putting

$$\mu(A) = \sum_{a \in A} \mu(a).$$

We note that $\mu(A)$ is a real number ≥ 0 or ∞ , and $\mu(X) = 1$ by (4). For a product of

subsets we have

$$\mu(AB) \leq \mu(A)\mu(B), \quad (5)$$

with equality if the product is unambiguous. To prove (5), let us first take A, B finite, say $A = \{a_1, \dots, a_m\}$, $B = \{b_1, \dots, b_n\}$. We have

$$\sum \mu(a_i b_j) = \sum \mu(a_i) \mu(b_j) = (\sum \mu(a_i))(\sum \mu(b_j)),$$

and here each member of AB occurs just once if AB is unambiguous, and otherwise more than once, so we obtain (5) in this case, with equality in the unambiguous case. In general (5) holds for any finite subsets A', B' of A, B by what has been shown. Therefore $\mu(A'B') \leq \mu(A)\mu(B)$, and now (5) follows by taking the limit.

In particular, for any code C the product CC is unambiguous by definition, hence on writing $C^2 = CC$, etc., we have

$$\mu(C^n) = \mu(C)^n, \quad n = 1, 2, \dots \quad (6)$$

Let us apply (6) to X ; we have $\mu(X) = 1$ by (4), and X is clearly a code, hence we obtain

$$\mu(X^n) = 1 \quad \text{for all } n \geq 1. \quad (7)$$

We shall need an estimate of $\mu(A)$ for finite sets A .

LEMMA 4.3 *If A is a finite subset of X^+ and μ is any measure on X^* , then*

$$\mu(A) \leq \max \{|a| \mid a \in A\}. \quad (8)$$

Proof. Since A is finite, the right-hand side of (8) is finite, say it equals d . Then $A \subseteq X \cup X^2 \cup \dots \cup X^d$, and so

$$\mu(A) \leq \mu(X) + \mu(X^2) + \dots + \mu(X^d) = d,$$

by (7). ■

From this lemma we can obtain a remarkable inequality satisfied by codes, which indicates that to be a code, a set must not be too large.

McMILLAN INEQUALITY *Let C be any code on X^* . Then for any measure μ on X we have*

$$\mu(C) \leq 1. \quad (9)$$

Proof. Consider first the case where C is finite and let $\max \{|c| \mid c \in C\} = d$. Then by (6), (8),

$$\mu(C^n) = \mu(C^n) \leq nd,$$

since the elements of C^n have length at most nd . Taking n th roots, we find that $\mu(C) \leq (nd)^{1/n}$. Here d is fixed; letting $n \rightarrow \infty$, we have $(nd)^{1/n} \rightarrow 1$, therefore $\mu(C) \leq 1$. In the general case, every finite subset of C is a code and so satisfies (9), hence this holds for C itself. ■

Of course the condition (9) is by no means sufficient for a code. For if $X = \{x, y\}$, $\mu(x) = p$, $\mu(y) = q$, consider e.g. $A = \{xy, yx, xy^2x\}$. Clearly A is not a code, but since $p + q = 1$, we have $pq \leq 1/4$, hence $\mu(A) = 2pq + p^2q^2 \leq 1/2 + 1/16 < 1$.

A code on X is said to be *maximal* if it is not a proper subset of a code on X . Maximal codes always exist by Zorn's lemma, since the property of being a code is of finite character. The above inequality provides a convenient test for maximality:

PROPOSITION 4.4 *Let C be a code on X . If $\mu(C) = 1$ for some measure μ , then C is a maximal code.*

Proof. Suppose that C is not a maximal code. Then we can find a code B containing C and another element b , say. We have $\mu(B) \geq \mu(C) + \mu(b) > 1$, and this contradicts (9). ■

For example, X and more generally X^n for any n is a maximal code.

Although the inequality (9) is not sufficient to guarantee that C is a code, there is a sense in which this inequality leads to a code. This is expressed in the next result, which gives a construction for codes with prescribed uniform measure.

THEOREM 4.5 *Let n_1, n_2, \dots be any sequence of positive integers. Then there exists a code $A = \{a_1, a_2, \dots\}$ with $|a_i| = n_i$ in an alphabet of r letters if and only if*

$$r^{-n_1} + r^{-n_2} + \dots \leq 1 \quad (\text{Kraft-McMillan inequality}). \quad (10)$$

We remark that the left of (10) is just the uniform measure of A .

Proof. The necessity of (10) is clear by (9) and the above remark. Conversely, assume that (10) holds, take the n_i to be ordered by size: $n_1 \leq n_2 \leq \dots$, and let $X = \{0, 1, \dots, r - 1\}$ be the alphabet. Define the partial sums of (10):

$$s_k = r^{-n_1} + \dots + r^{-n_{k-1}},$$

and for each s_k define an integer

$$p_k = r^{n_k} s_k = r^{n_k - n_1} + \dots + r^{n_k - n_{k-1}}.$$

Each p_k is an integer and since $s_k < 1$ by (10), we have

$$0 \leq p_k < r^{n_k}.$$

Now we take a_k to be the element of X^* formed by expressing p_k in the scale of r , with enough 0's prefixed to bring the length up to n_k :

$$a_k = \alpha_1 \alpha_2 \cdots \alpha_{n_k} = \alpha_1 r^{n_k-1} + \alpha_2 r^{n_k-2} + \cdots + \alpha_{n_k-1} r + \alpha_{n_k}, \quad \alpha_i \in X.$$

We claim that $A = \{a_1, a_2, \dots\}$ is a code of the required type. The lengths are right by construction, and we shall complete the proof by showing that A is a prefix code. If a_j is a prefix of a_i , $j < i$, then a_j is obtained from a_i by cutting off the last $n_i - n_j$ digits. Thus p_j is the greatest integer in the fraction

$$\frac{p_i}{r^{n_i-n_j}} = r^{n_j} s_i \geq r^{n_j} (s_j + r^{-n_j}) = p_j + 1,$$

but this is a contradiction. Thus A is indeed a prefix code. ■

For example, take $r = 3$ and consider the sequence 1, 1, 2, 2, 3, 3, 3. We have $\mu(A) = 1/3 + 1/3 + 1/9 + 1/9 + 1/27 + 1/27 + 1/27 = 1$, so we have a maximal code. It is given by the table:

k	1	2	3	4	5	6	7
n_k	1	1	2	2	3	3	3
p_k	0	1	6	7	24	25	26
a_k	0	1	20	21	220	221	222

The construction of a_k given in Th. 4.5 can be described by the rule: choose the least number in the ternary scale which is not a prefix of 222 and which has no a_i ($i < k$) as a prefix.

We have seen that codes are certain subsets of X^* that are not too large; we now introduce a class of subsets that are not ‘too small’, in order to study the interplay between these classes. A subset A of X^* is said to be *complete* if the submonoid generated by it meets every ideal of X^* , i.e. every word in X^* occurs as a factor in some word in $\langle A \rangle$:

$$X^* w X^* \cap \langle A \rangle \neq \emptyset \quad \text{for all } w \in X^*.$$

PROPOSITION 4.6 *Let A be a finite complete subset of X^* and let m be the uniform measure on X . Then $m(A) \geq 1$.*

Proof. If $1 \in A$, the result is clear, because $m(1) = 1$, so we may assume that $1 \notin A$. Let d be the maximal length of words in A and put

$$f_k = |\langle A \rangle \cap X^k|;$$

thus f_k is the number of words in $\langle A \rangle$ of length k . We claim that

$$\sum_{k=0}^{2d} f_{n+k} \geq \frac{r^n}{2d+1} \quad \text{for all } n \geq 1, \text{ where } r = |X|. \quad (11)$$

To prove (11), consider X^n ; by the completeness of A , each word in X^n is a factor of a word in $\langle A \rangle$, whose length can be taken to lie between n and $n + 2d$. In a word of length $n + k$ there are at most $k + 1$ different factors of length n , so the words in $\langle A \rangle$ with length in the interval $[n, n + 2d]$ have at most $\sum_0^{2d} (k + 1)f_{n+k}$ factors of length n . It follows that

$$(2d + 1) \sum_{k=0}^{2d} f_{n+k} \geq \sum (k + 1)f_{n+k} \geq |X|^n = |X^n| = r^n,$$

and rearranging this inequality we obtain (11).

Now let a_k be the number of products of elements in A that are of total length k ; thus a_k counts the number of occurrences in $\langle A \rangle$, and hence $a_k \geq f_k$, with equality iff A is a code. We have

$$\sum_{s=1}^{\infty} m(A)^s = \sum_{k=1}^{\infty} \frac{a_k}{r^k},$$

and so

$$\begin{aligned} \sum m(A)^s &\geq \sum_{k=1}^{\infty} \frac{f_k}{r^k} \\ &\geq r^{-2d-1} \sum_{k=0}^{2d} f_{k+1} + r^{-4d-2} \sum_{k=0}^{2d} f_{2d+2+k} + \dots \\ &\geq \frac{1}{2d+1} (r^{-2d} + r^{-2d} + \dots), \end{aligned}$$

by (11). This latter series diverges, hence $m(A) \geq 1$, as claimed. ■

We next establish a connexion with codes:

THEOREM 4.7 (Schützenberger) *Any maximal code is complete.*

Proof. Let A be a code which is not complete; we shall show how to enlarge it. If $|X| = 1$, any non-empty set is complete, so we have $A = \emptyset$ and then X is a larger code. When $|X| > 1$, we have to find $b \notin A$ such that $A \cup \{b\}$ is a code. Since A is not complete, there is a word $c \in X^*$ such that $X^*cX^* \cap \langle A \rangle = \emptyset$. One might be tempted at this point to adjoin c to A , but this leads to problems because c might intersect itself; we shall construct b to avoid this (see the figure below). Let $|c| = \gamma$ and put $c = xc'$, where $x \in X$. Choose $y \neq x$ in X and put

$$b = cxy^\gamma = xc'xy^\gamma. \quad (12)$$

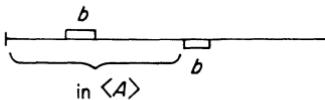
From the definition of c it is clear that

$$X^*bX^* \cap \langle A \rangle = \emptyset. \quad (13)$$

We claim that $A \cup \{b\}$ is a code. For if not, then we have an equation

$$a_1 \cdots a_r = a'_1 \cdots a'_s \quad \text{where } a_i, a'_j \in A \cup \{b\}, \quad (14)$$

and (14) is non-trivial. Since A is a code, b must occur in (14), and by (13) it must occur on both sides of (14), say $a_i = a'_j = b$ for some i, j . We take i, j minimal; if the two occurrences of b do not overlap, we have a contradiction to (13) (see diagram).



If there is an overlap, it must be at least γ letters, by (12), but that is impossible, because in b , letters 1 and $\gamma + 1$ are x , whereas the last γ letters of b are y . ■

Our final result clarifies the relation between complete sets and codes.

THEOREM 4.8 (Boë, de Luca and Restivo 1980) *Let A be a finite subset of X^+ . Then any two of the following imply the third:*

- (a) A is a code,
- (b) A is complete,
- (c) $m(A) = 1$.

Proof. (a, b) \Rightarrow (c). If A is a complete code, then $m(A) = 1$ by Prop. 4.6 and McMillan's inequality. (b, c) \Rightarrow (a). If $m(A) = 1$ but A is not a code, then for some n , $m(A^n) < m(A)^n = 1$, hence A^n is not complete, and so neither is A . (c, a) \Rightarrow (b). If $m(A) = 1$ and A is a code, then A is a maximal code and so it is complete, by Th. 4.7. ■

The situation is reminiscent of what happens for bases in a vector space. A basis is a linearly independent spanning set, and of the following conditions any two imply the third:

- (a) A is linearly independent,
- (b) A is a spanning set,
- (c) $|A| = \dim V$.

In the proof the exchange axiom plays a vital role: it has the consequence that any spanning set contains a linearly independent set which still spans. The analogue here is false: there are minimal complete sets that are not codes. For example, $X = \{x, y\}$, $A = \{x^3, x^2yx, x^2y, yx, y\}$. It is easily verified that A is minimal complete. The subset $A_1 = \{\dot{x}^3, x^2y, yx, y\}$ is not a code and $m(A_1) = 1$, while all

other proper subsets of A are codes. What can be shown is the following (cf. Boë, de Luca and Restivo 1980):

A minimal complete set is a code iff all its proper subsets are codes.

For finite codes the converse of Th. 4.7 holds: every finite complete code is maximal (this follows from Th. 4.8 and Prop. 4.4), but there are infinite complete codes which are not maximal (cf. Ex. (5)); this also shows that Th. 4.8 does not extend to infinite sets.

Exercises

- (1) Determine which of the following are codes: (i) $\{xy, xy^2, y^2\}$, (ii) $\{x^2, xy, x^2y, xy^2, y^2\}$, (iii) $\{x^2, xy^2, x^2y, xy^3, y^2, yx\}$.
- (2) Construct a code for $r = 2$ and the sequence $1, 2, 3, \dots$. Likewise for $r = 4$ and the sequence $1, 1, 1, 2, 2, 2, 3, 3, \dots$.
- (3) Let X be an alphabet, G a group and $f: X^* \rightarrow G$ a homomorphism. Show that for any subgroup H of G , Hf^{-1} is a free submonoid of X^* , and its generating set is a maximal code which is *bifix* (i.e. prefix and suffix). Such a code is called a *group code*.
- (4) Let X be an alphabet and $w_x(u)$ the x -length of $u \in X^*$. Show that for any integer m the mapping $f: u \mapsto w_x(u) \pmod{m}$ is a homomorphism from X^* to \mathbf{Z}/m and describe the group code $0f^{-1}$. Show that for $m = 2$, $X = \{x, y\}$, $0f^{-1} = \{y\} \cup \{xy^*x\}$; find $0g^{-1}$ where $g: u \mapsto w_x(u) \pmod{3}$.
- (5) Let $X = \{x, y\}$. Show that $\delta(u) = w_x(u) - w_y(u)$ is a homomorphism to \mathbf{Z} . Describe the corresponding group code D (this is known as the *Dyck code*). Show that D is complete and remains complete if any one element is omitted.
- (6) Let $f: X^* \rightarrow Y^*$ be an injective homomorphism of free monoids. Show that if A is a code in X^* , then Af is a code in Y^* ; if B is a code in Y^* then Bf^{-1} is a code in X^* .
- (7) Let A, B be any codes in X^* . Show that A^n is a code for all $n \geq 1$, but AB need not be a code.

11.5 Free algebras and formal power series rings

There is a further way of describing languages, namely as formal power series. This is in some respects the simplest and most natural method. Let X be a finite alphabet and k a commutative field. By a *formal power series* in X over k we understand a function f on X^* with values in k . The value of f at $u \in X^*$ is denoted by (f, u) , and f itself may be written as a series

$$f = \sum (f, u)u. \quad (1)$$

Here (f, u) is called the *coefficient* of u and, in particular, $(f, 1)$ is called the *constant term* of f .

Series are added and multiplied by the rules

$$(f + g, u) = (f, u) + (g, u), \quad (2)$$

$$(fg, u) = \sum_{yz=u} (f, y)(g, z). \quad (3)$$

Since each element of X^* has only a finite number of factors, the sum in (3) is finite, so fg is well-defined. The set of all these power series is denoted by $k \ll X \gg$; it is easily seen to form a k -algebra with respect to these operations. For each power series f its *support* is defined as

$$D(f) = \{u \in X^* \mid (f, u) \neq 0\}.$$

Thus u lies in the support of f precisely if it occurs in the expression (1) for f . The elements of finite support are called *polynomials* in X ; they are in fact just polynomials, i.e. k -linear combinations of products of elements of X , but care must be taken to preserve the order of the factors, since the elements of X do not commute. These polynomials form a subalgebra $k\langle X \rangle$, called the *free k -algebra* on X . We remark that $k\langle X \rangle$ can also be defined as the monoid algebra of X^* , in analogy to the group algebra.

For each power series f we define its *order* $o(f)$ as the minimum of the lengths of terms in its support. The order is a positive integer or zero, according as $(f, 1)$ is or is not zero. For a polynomial f we can also define the *degree* $d(f)$; it is the maximum of the lengths of terms in its support. If all the terms of f have the same length r , so that $o(f) = d(f) = r$, then f is said to be *homogeneous* of degree r .

We remark that if u is a series of positive order, we can form the series

$$u^* = 1 + u + u^2 + \dots$$

The infinite series on the right ‘converges’ because if $o(u) = r$, then for any given d , u^n contributes only if $rn \leq d$. Thus in calculating the terms of degree d in u^* we need only consider u^n for $n = 0, 1, \dots, [d/r]$. We also note that u^* satisfies the equations

$$u^*u = uu^* = u^* - 1.$$

Hence $(1 - u)u^* = u^*(1 - u) = 1$, so u^* is the inverse of $1 - u$:

$$(1 - u)^{-1} = u^* = \sum_0^\infty u^n. \quad (4)$$

It is easily verified that $k\langle X \rangle$ has the familiar universal property: every mapping $\varphi: X \rightarrow A$ into a k -algebra A can be extended in just one way to a homomorphism $\bar{\varphi}: k\langle X \rangle \rightarrow A$. As a consequence every k -algebra can be written as a homomorphic image of a free k -algebra, possibly on an infinite alphabet. Of course free algebras on an infinite alphabet are defined in exactly the same way;

for power series rings there are several different possible definitions, but this need not concern us here, as we shall only consider the case of a finite alphabet.

The free algebra $k\langle X \rangle$ may be regarded as a generalization of the polynomial ring $k[x]$, to which it reduces when X consists of a single element x . The polynomial ring is of course well-known and has been thoroughly studied. The main tool in its study is the Euclidean algorithm; this allows one to prove that $k[x]$ is a principal ideal domain (PID) and a unique factorization domain (UFD) (Vol. 1, p. 319ff.). The UF-property extends to polynomials in several (commuting) variables, as we saw in 9.3, but there is no analogue to the principal ideal property in this case. For the non-commutative polynomial ring $k\langle X \rangle$ the UF-property persists, albeit in a more complicated form, and we shall have no more to say about it here (cf. Cohn 1985, Ch. 3 and Ex. (3)). The principal ideal property generalizes as follows. Let us define a *right free ideal ring*, *right fir* for short, as a ring R with invariant bases number (IBN) in which every right ideal is free as right R -module; *left firs* are defined similarly, and a left and right fir is called a *fir*. In the commutative case a fir is just a PID and the fact that $k[x]$ is a PID generalizes to the assertion that $k\langle X \rangle$ is a fir. This is usually proved by means of the *weak algorithm*, a generalization of the Euclidean algorithm, to which it reduces in the commutative case. We shall not enter into the details (cf. Cohn 1985, Ch. 2) but confine ourselves below to giving a direct proof that $k\langle X \rangle$ is a fir. This method, similar to the technique used to prove that subgroups of free groups are free, is due to J. Lewin; our exposition follows essentially Berstel and Reutenauer (1984) (cf. also Cohn 1985, Ch. 6).

THEOREM 5.1 *Let $F = k\langle X \rangle$ be the free algebra on a finite set X over a field k and let α be any right ideal of F . Then there exists a Schreier set P in X^* which is maximal k -linearly independent $(\text{mod } \alpha)$. If C is the corresponding prefix set, determined as in Prop. 4.2, then for each $c \in C$ there is an element of α :*

$$f_c = c - \sum_{p \in P, \alpha_{c,p} \in k} \alpha_{c,p} p \quad (p \in P, \alpha_{c,p} \in k)$$

where the sum ranges over all $p \in P$, but has only finitely many non-zero terms for each $c \in C$, such that α is free as right F -module on the f_c ($c \in C$) as basis. Similarly for left ideals, and F is a fir.

Proof. The monoid X^* is a k -basis of F , hence its image in F/α is a spanning set and it therefore contains a basis of F/α as k -space. Moreover, we can choose this basis to be a Schreier set, by building it up according to length. Thus if P_n is a Schreier set which forms a basis for all elements of F of degree at most n (mod α), then the set $P_n X$ spans the space of elements of degree at most $n+1$ and by choosing a basis from it we obtain a Schreier set P_{n+1} containing P_n and forming a basis (mod α) for the elements of degree at most $n+1$. In this way we obtain a Schreier set $P = \bigcup P_n$ which is maximal k -linearly independent (mod α) and hence a k -basis. Let C be the corresponding prefix set. For each $c \in C$ the set $P \cup \{c\}$ is

still a Schreier set, but by the maximality of P it is linearly dependent mod \mathfrak{a} , say

$$f_c = c - \sum \alpha_{c,p} p \in \mathfrak{a}, \quad (5)$$

where the sum ranges over P and almost all the $\alpha_{c,p}$ vanish. We claim that every element $b \in F$ can be written as

$$b = \sum f_c g_c + \sum \beta_p p, \quad \text{where } g_c \in F, \beta_p \in k, \quad (6)$$

and the sums range over C and P respectively. By linearity it is enough to prove this when b is a monomial. When $b \in P$, this is clear; we need only take $\beta_p = 1$ for $p = b$ and the other coefficients zero. When $b \notin P$, we can by Prop. 4.2 write $b = cu$ for some $c \in C$. By (5) we have

$$b = f_c u + \sum \alpha_{c,p} pu. \quad (7)$$

For any $p \in P$, either $pu \in P$ or $pu = c_1 u_1$ where $c_1 \in C$ and hence $|p| < |c_1|$, $|u_1| < |u|$. In the first case we have achieved the form (6); in the second case we use induction on $|u|$ to express $c_1 u_1$ in the same form. Thus we can reduce all the terms on the right of (7) to the form (6) and the conclusion follows.

We claim that the elements (5) form the desired basis of \mathfrak{a} . To show that they generate \mathfrak{a} , let us take $b \in \mathfrak{a}$ and apply the natural homomorphism $F \rightarrow F/\mathfrak{a}$. Writing the image of r as \bar{r} , we have

$$0 = \bar{b} = \sum \beta_p \bar{p}.$$

Since the \bar{p} are linearly independent by construction, we have $\beta_p = 0$, so $b = \sum f_c g_c$ and it follows that the f_c generate \mathfrak{a} . To prove their linear independence, assume that $\sum f_c g_c = 0$, where not all the g_c vanish. Then by (5)

$$\sum c g_c = \sum \alpha_{c,p} p g_c. \quad (8)$$

Take a word w of maximal length occurring in some g_c , say in g_{c_0} . Since C is a prefix code, $c_0 w$ occurs with a non-zero coefficient λ on the left of (8). Hence

$$\lambda = \sum \alpha_{c,p} \mu_{c,p},$$

where $\mu_{c,p}$ is the coefficient of $c_0 w$ in $p g_c$. Now the relation $c_0 w = pu$ can hold only when p is a proper left factor of c_0 , hence $|p| < |c_0|$, $|u| > |w|$ and this contradicts the definition of w . This contradiction shows that the f_c are linearly independent, so they form a basis of \mathfrak{a} , which is therefore a free right ideal. By symmetry every left ideal is free, and F clearly has IBN, since we have a homomorphism $F \rightarrow k$, obtained by setting $X = 0$. This shows F to be a fir. ■

We recall the equation $X^* = C^* P$ obtained in Prop. 4.2. In the power series ring this may be rewritten as $(1 - X)^{-1} = (1 - C)^{-1} P$, where X, C, P are now the sums of the corresponding sets. On multiplying up, we obtain $1 - C = P(1 - X)$, or

$$C = P X - (P - 1). \quad (9)$$

This tells us again that the prefix set C consists of all products $px(p \in P, x \in X)$ which are not in P . In the case where P and hence C is finite, we obtain

COROLLARY 5.2 *Let $F = k\langle X \rangle$ be the free algebra as in Th. 5.1 and \mathfrak{a} a right ideal. If $|X| = d$ and \mathfrak{a} has finite codimension r in F , then its rank (as free F -module) is finite, say n , and is given by*

$$n - 1 = r(d - 1). \quad (10)$$

This follows from the formula $1 - C = P(1 - X)$ on replacing each element of X by 1. ■

We remark that (10) is analogous to Schreier's formula for the rank of a subgroup of a free group; it is known as the *Schreier–Lewin formula*.

We now turn to consider the power series ring. To describe the structure of $k\langle\langle X \rangle\rangle$ we recall that a *local ring* is a ring R in which the set of all non-units forms an ideal \mathfrak{m} . Clearly \mathfrak{m} is then the unique maximal ideal of R , and R/\mathfrak{m} is a skew field, called the *residue class field* of R .

PROPOSITION 5.3 *The power series ring $k\langle\langle X \rangle\rangle$ on any finite set X over a field k is a local ring with residue class field k . Its maximal ideal consists of all elements with zero constant term.*

Proof. The mapping $X \rightarrow 0$ defines a homomorphism of $k\langle\langle X \rangle\rangle$ onto k , hence the kernel \mathfrak{m} , consisting of all elements with zero constant term, is an ideal in $k\langle\langle X \rangle\rangle$. It follows that $k\langle\langle X \rangle\rangle/\mathfrak{m} \cong k$, and \mathfrak{m} contains no invertible element. To complete the proof we need only show that every element not in \mathfrak{m} has an inverse. Consider first an element with constant term 1, say $f = 1 - u$, where $o(u) > 0$. By (4) we have $f^{-1} = (1 - u)^{-1} = u^*$.

In general f has a constant term $\lambda \neq 0$. Then $f\lambda^{-1}$ has constant term 1 and so it has an inverse g ; it follows that f has the inverse $\lambda^{-1}g$, for $f\lambda^{-1}g = 1$ and similarly $\lambda^{-1}gf = gf\lambda^{-1} = 1$. ■

A power series f is called *rational* if it can be obtained from elements of $k\langle X \rangle$ by a finite number of operations of addition, multiplication and inversion of series with non-zero constant terms. The rational series form a subring of $k\langle\langle X \rangle\rangle$, denoted by $k_{\text{rat}}\langle X \rangle$, as we shall see in Prop. 5.4 below, and the method of Prop. 5.3 shows that $k_{\text{rat}}\langle X \rangle$ is again a local ring.

Let us note that any square matrix A over $k_{\text{rat}}\langle X \rangle$ is invertible provided that its constant term is invertible over k . To prove this fact, we write $A = A_0 + B$, where A_0 is over k and B has no constant term. By hypothesis A_0 is invertible over k , and on writing $A_0^{-1}A = I + A_0^{-1}B$ we reduce the problem to the case where $A_0 = I$. Now A is a matrix whose diagonal elements are units, while the other entries are non-units. We shall use induction on the size of A . By row operations

we can reduce the $(1, 1)$ -entry of A to 1, and subtracting suitable multiples of the first row from the others, we reduce the other elements of the first column to 0. We now have a matrix

$$\begin{pmatrix} 1 & a \\ 0 & A' \end{pmatrix}$$

and here A' again has units on the main diagonal and non-units elsewhere. By induction we can invert A' and so accomplish the task of inverting A . We shall use this result to obtain a criterion for the rationality of a power series.

PROPOSITION 5.4 (Schützenberger) *The set $k_{\text{rat}}\langle X \rangle$ of rational series is a subalgebra of $k\langle\langle X \rangle\rangle$. Moreover, for any $f \in k\langle\langle X \rangle\rangle$ the following conditions are equivalent:*

- (a) f is rational,
- (b) $f = u_1$ is the first component of the solution of a system

$$u = Bu + b, \quad (11)$$

where B is a matrix and b a column over $k\langle X \rangle$, B is homogeneous of degree 1 and b has degree at most 1,

- (c) $f = u_1$ is a component of the solution of a system

$$Fu = b, \quad (12)$$

where F is a matrix and b a column over $k\langle X \rangle$, and F has invertible constant term.

We remark that (11) can also be written as

$$(I - B)u = b; \quad (13)$$

thus it has the form (12), where F now has constant term I and has no term of degree higher than 1.

Proof. (a) \Rightarrow (b). We shall show that the set of elements satisfying (b) forms a subalgebra of $k\langle\langle X \rangle\rangle$ containing $k\langle X \rangle$, in which every series of order 0 is invertible. It then follows that this subalgebra contains $k_{\text{rat}}\langle X \rangle$.

It is clear that $a \in k \cup X$ is the solution of $u_1 = a$. Given f, g , suppose that $f = u_1$, where u is the solution of (11), and $g = v_1$, where v is the solution of $v = Cv + c$, where C satisfies the same conditions as B in (11). We shall rewrite these equations as $(I - B)u = b$, $(I - C)v = c$. Then $f - g$ is the first component of the solution of the system

$$\left(\begin{array}{c|cc} I - B & e_1 - B_1 & 0 \\ \hline 0 & I - C & \end{array} \right) w = \begin{pmatrix} b \\ c \end{pmatrix}, \quad (14)$$

where $e_1 = (1, 0, \dots, 0)^T$ and B_1 is the first column of B ; for (14) is satisfied by

$w = (u_1 - v_1, u_2, \dots, u_m, v_1, \dots, v_n)^T$. In (14) the matrix on the left is not of the required form, but it can be brought to the form of a matrix with constant term I (and no term of degree higher than 1) by subtracting row $m + 1$ from row 1 to get rid of the coefficient 1 in the $(1, m + 1)$ -position.

Similarly, fg is the first component of the solution of

$$\left(\begin{array}{c|cc} I - B & b & 0 \\ \hline 0 & & I - C \end{array} \right) w = \begin{pmatrix} 0 \\ c \end{pmatrix}, \quad (15)$$

for (15) is satisfied by $w = (u_1 v_1, u_2 v_1, \dots, u_m v_1, v_1, \dots, v_n)^T$. If b has a non-zero constant term, we can bring (15) to the required form by subtracting appropriate multiples of row $m + 1$ from row 1, ..., row m .

It remains to invert a series of order 0 or, what comes to the same thing (after what has been shown), a series with constant term 1. Let f have zero constant term and suppose that $f = u_1$, where u satisfies (11). We shall invert $1 + f$ by finding an equation for g , where $(1 - g)(1 + f) = 1$. We may assume that $b = (0, \dots, 0, 1)^T$ by writing (11) in the form

$$\left(\begin{array}{c|c} I - B & -b \\ \hline 0 & 1 \end{array} \right) \begin{pmatrix} u \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

subtracting multiples of column $2, \dots, n - 1$ from column n to reduce the constant term of b to 0, and adding corresponding multiples of 1 to the components of u . This leaves u_1 unchanged and it is sufficient: b_1 already has zero constant term because this is true of u_1 . Thus our system now has the form (after a change of notation)

$$(I - B)u = e_n,$$

and here $n > 1$, because u_1 has zero constant term. Let E_{n1} be the matrix with 1 in the $(n, 1)$ -position and 0's elsewhere. Then we have

$$(I - B + E_{n1})u = e_n(1 + u_1). \quad (16)$$

But we can also solve the system

$$(I - B + E_{n1})v = e_n, \quad (17)$$

and we can again bring the matrix on the left to the required form by subtracting the first row from the last, without affecting the solution. Comparing (16) and (17), we find that

$$u = v(1 + u_1).$$

In particular, $u_1 = v_1(1 + u_1)$, hence $(1 - v_1)(1 + u_1) = 1 + u_1 - u_1 = 1$, so we have found a left inverse for $1 + u_1$. By uniqueness it is a two-sided inverse, which is what we had to find. This then shows that the set of components of solutions of systems (11) is a ring containing $k\langle X \rangle$ and admitting inversion when possible.

(b) \Rightarrow (c) is clear, and to prove (c) \Rightarrow (a) we must show that the solution of (12) has rational components. We have seen that the matrix $I - B$ has an inverse whose entries are rational, hence the same is true of $u = (I - B)^{-1}b$.

We have now shown (a)–(c) to be equivalent, and the elements satisfying (11) form a subring containing $k\langle X \rangle$, hence a subalgebra. It follows that $k_{\text{rat}}\langle X \rangle$ is a subalgebra in which all series of order 0 have inverses. ■

With every language L on the alphabet X we associate a power series f_L , its *characteristic series*, defined by

$$(f_L, u) = \begin{cases} 1 & \text{if } u \in L, \\ 0 & \text{if } u \notin L. \end{cases}$$

In this way the language is described by a single element of $k \ll X \gg$. Our main object will be to characterize the regular and context-free languages.

THEOREM 5.5 (Schützenberger) *A language on X is regular if and only if its characteristic series is rational.*

Proof. Let L be a regular language. Then L is generated by a grammar with rules of the form

$$\alpha \rightarrow x\beta, \quad \alpha \rightarrow y; \tag{18}$$

moreover, each word of L has a single derivation. Let us number the variables of the grammar as u_1, \dots, u_n , where $u_1 = \sigma$, and number the terminal letters as x_1, \dots, x_r . The rules (18) may be written as

$$\alpha = \sum x\beta + \sum y, \tag{19}$$

where the summations are over all the right-hand sides of rules with α on the left. If we express (19) in terms of the u 's and x 's, we find

$$u_i = \sum b_{ij}u_j + b_i; \tag{20}$$

to generate the language we replace $\sigma = u_1$ by the right-hand side in (20) and continue replacing each u_j by the corresponding right-hand side. We thus obtain power series f_1, \dots, f_n such that $u_i = f_i$ satisfies (20), and f_1 is the characteristic series of L ; clearly f_1 is rational.

Conversely, if the characteristic series for L is rational, then it is given as a component of the solution of the system (20), where the b_{ij} are linear homogeneous in the x_i and the b_i are of degree at most 1. We take the grammar of L in the form $u_i \rightarrow x_k u_j$ if x_k occurs in b_{ij} and $u_i \rightarrow 1$ if b_i has a non-zero constant term. This ensures that the derivations give precisely the words of L , thus L is regular. ■

For example, consider the language $\{xy^n\}$, with the grammar $\sigma \rightarrow \sigma y; x$. Its characteristic series is obtained by solving the equation $u = uy + x$, i.e. $u = x(1 - y)^{-1} = \sum xy^n$.

Next we consider the problem of characterizing context-free languages. For this purpose we define another subalgebra of the power series ring. An element f of $k \ll X \gg$ is said to be *algebraic* if it is of the form $f = \alpha + u_1$, where $\alpha \in k$ and u_1 is the first component of the solution of a system of equations

$$u_i = \varphi_i(u, x) \quad i = 1, \dots, n, \quad (21)$$

where φ_i is a (non-commutative) polynomial in the u 's and x 's without constant term or linear term in the u 's. The set of all algebraic elements is denoted by $k_{\text{alg}} \langle X \rangle$; we shall now show that it is a subalgebra of $k \ll X \gg$.

PROPOSITION 5.6 *Any system (21), where each φ_i is a polynomial without constant term or linear term in the u 's, has a unique solution in $k \ll X \gg$ with components of positive order, and the set $k_{\text{alg}} \langle X \rangle$ is a subalgebra and a local ring, each element of order 0 being invertible.*

Proof. Writing $u_i^{(v)}$ for the component of degree v of u_i , we find, by equating homogeneous components in (21),

$$u_i^{(v)} = \varphi_i^{(v)}(u, x).$$

Here $\varphi_i^{(v)}$ is the sum of all terms of degree v in φ_i . By hypothesis, for any term $u_j^{(\mu)}$ occurring in $\varphi_i^{(v)}$ we have $\mu < v$, so the components $u_i^{(v)}$ are uniquely determined in terms of the $u_j^{(\mu)}$ with $\mu < v$, while $u_i^{(0)} = 0$, again by hypothesis. Thus (21) has a unique solution of positive order.

If $u_i = \varphi_i(u, x)$ ($i = 1, \dots, m$), $v_j = \psi_j(u, x)$ ($j = 1, \dots, n$) are two such systems, then to show that $u_1 - v_1, u_1 v_1, \sum u_1^v = (1 - u_1)^{-1}$ are algebraic, we combine the above systems of equations for u_i, v_j with the equation $w = \varphi_1 - \psi_1$, $w = \varphi_1 \psi_1$, $w = \varphi_1 + u_1 w$ respectively. This shows that we have a subalgebra. Moreover, the elements of order 0 are invertible, so we have a local ring. ■

It is clear that we have the inclusions

$$k \langle X \rangle \subset k_{\text{rat}} \langle X \rangle \subset k_{\text{alg}} \langle X \rangle \subset k \ll X \gg; \quad (22)$$

that the inclusions are strict is easily seen, by considering the case where X consists of a single letter.

We now come to the promised characterization of context-free languages.

THEOREM 5.7 (Schützenberger) *A language L in an alphabet X is context-free if and only if its characteristic series is algebraic.*

Proof. Let L be context-free, generated by a grammar with the rules $u_i \rightarrow w_{ik}$,

where w_{ik} is a word in the u 's and x 's, and $u_1 = \sigma$ is the sentence symbol. If there is a rule of the form $u_i \rightarrow u_j$, we replace it by the rules $u_i \rightarrow f$, where f runs over the right-hand sides of all the rules $u_j \rightarrow w_{jk}$. We now write

$$u_i = \varphi_i(u, x), \quad (23)$$

where $\varphi_i(u, x)$ is the sum of the right-hand sides of all the rules with u_i on the left. On solving the system (23) we obtain for u_1 the series f_L , hence f_L is algebraic.

Conversely, assume that f_L is algebraic, given by u_1 , where u is the solution of (23). Then the language L is obtained by applying all the rules $u_i \rightarrow w$, where w runs over all the words in the support of $\varphi_i(u, x)$; hence L is context-free, as we had to show. ■

To give an example, the language $\{x^n y^n\}$ has the characteristic series $u = \sum x^n y^n$, which is obtained by solving the equation

$$u = 1 + xuy.$$

Exercises

- (1) Adapt the proof of Th. 5.1 to the case of an infinite alphabet.
- (2) Verify that the prefix set associated with a right ideal of finite codimension in a free algebra is a maximal code.
- (3) Factorize the element $xyzyx + xyz + zyx + xyx + x + z$ of $F = k\langle x, y, z \rangle$ in all possible ways. (It can be shown that any two complete factorizations of an element of F have the same number of terms, and these terms can be paired off in such a way that corresponding terms are 'similar', where a, b are similar if $F/aF \cong F/bF$.)
- (4) Show that in $\mathbf{R}\langle x, y \rangle$ the element $xy^2x + xy + yx + x^2 + 1$ is an atom, but it does not remain one under extension of \mathbf{R} to \mathbf{C} .
- (5) Show that every CF-language in one letter is regular.
- (6) Show that the inclusions in (22) are strict.
- (7) Define the *Hankel matrix* of a power series f as the infinite matrix $H(f)$ indexed by X^* , whose (u, v) -entry is (f, uv) . Show that f is rational iff $H(f)$ has finite rank.

Further exercises on Chapter 11

- (1) Let A be an infinite set and $M(A)$ the set of all injective mappings of A into itself such that the complement of the image is infinite. Show that $M(A)$ is a semigroup admitting right cancellation and right division (writing mappings on the right), i.e. given $\alpha, \beta \in M(A)$, the equation $\alpha x = \beta$ has a unique solution.

- (2) Show that two elements of a free monoid commute iff they can be written as powers of the same element.
- (3) Let F be a free monoid. Show that if $uv = vw$, then there exist $a, b \in F$ such that $u = ab$, $w = ba$, $v = (ab)^r a = a(ba)^r$.
- (4) Let C be the monoid on a, b as generating set with defining relation $ba = 1$. Show that each element of C can be written uniquely as $a^r b^s$, $r, s \geq 0$. (*Hint.* Define C as set of mappings of \mathbf{N}^2 into itself by the rules: $(m, n)a = (m, n - 1)$ if $n \geq 1$, $(m, 0)a = (m + 1, 0)$, $(m, n)b = (m, n + 1)$, and verify that this is a faithful representation, i.e. distinct mappings have distinct images. C is called the *bicyclic monoid*.)
- (5) Show that any CF-language can be generated by a grammar whose rules are all of the form $\alpha \rightarrow \beta\gamma$ or $\alpha \rightarrow x(\alpha, \beta, \gamma \in V, x \in X)$. (*Hint.* Use induction on the lengths of the right-hand sides of all rules not of this form. This is called the *Chomsky normal form*.)
- (6) A grammar is called *self-embedding* if it includes a derivation $\alpha \rightarrow u\alpha v$, where $uv \neq 1$ and $\alpha \in V$. Show that a language L is regular iff there is a CF-grammar generating L which is not self-embedding.
- (7) Show that every language of type 0, 2, 3 is closed under substitution: if L has type v ($= 0, 2, 3$) with alphabet $X = \{x_1, \dots, x_r\}$ and x_1 is replaced by a language of type v with an alphabet Y disjoint from X , then the resulting language is of type v on $X \cup Y$. A CS-language is closed under substitution provided that the language substituted is proper.
- (8) Show that a CF-language satisfies the following strengthening of the pumping lemma (sometimes called the *iteration lemma*): If L is context-free, there exists an integer p such that for any word w of length $|w| \geq p$ and any partition (v_1, \dots, v_5) of $|w|$ such that $v_3 > 0$ and either $v_2 > 0$ or $v_4 > 0$, there is a factorization $w = v_1 \cdots v_5$ with $|v_i| = v_i$ and $v_1 v_2^n v_3 v_4^n v_5 \in L$ for all n . Use the result to verify that $L = \{a^*bc\} \cup \{a^pba^nca^n | p \text{ prime}, n \geq 0\}$ is not context-free but satisfies the conditions of the pumping lemma (cf. Berstel 1979).
- (9) Show that the generating set of a maximal free submonoid of a free monoid is a maximal code.
- (10) Show that a subset C of X^* is a code iff the submonoid generated by C has the characteristic series $(1 - C)^{-1}$.
- (11) Verify that the power series ring on X may be interpreted as the incidence algebra of X^* when the latter is partially ordered by right divisibility (cf. 2.4). (*Hint.* Define $f_{u,v} = (f, z)$ if $u = zv$ and 0 otherwise.) Hence deduce Prop. 5.3 from Prop. 2.4.1.

Bibliography

This is primarily a list of books where the topics are pursued further (and which were often used as sources); it also includes papers referred to in the text as well as a small selection of articles having a bearing on the text.

General

(Including items referred to in more than one chapter.)

Bourbaki, N., *Algèbre*, Chs. 1–10, 1961–80, Hermann, Paris, later Masson, Paris.

Bourbaki, N., *Éléments d'Histoire de Mathématiques*, 1984, Masson, Paris.

Cohn, P. M., *Free Rings and Their Relations* (2nd edn), LMS Monographs No. 19, 1985, Academic Press, New York.

Jacobson, N., *Basic Algebra I* (2nd edn), 1985, II, 1980, Freeman, San Francisco.

Lang, S., *Algebra* (2nd edn), 1984, Addison-Wesley, Reading, MA.

Lidl, R. and Pilz, G., *Applied Abstract Algebra*, 1984, Springer, Berlin.

Mac Lane, S. and Birkhoff, G., *Algebra* (3rd edn), 1988, Chelsea, New York.

v. d. Waerden, B. L., *Algebra I, II*, 1971, 1976, Springer, Berlin.

Weber, H., *Lehrbuch der Algebra I–III*, 1898–1908, Teubner, Leipzig, reprinted 1963, Chelsea, New York.

Chapters 1–2

Barwise, J. (ed.), *Handbook of Logic*, 1977, North-Holland, Amsterdam.

Birkhoff, G., *Lattice Theory* (3rd edn), 1967, AMS, Providence, RI.

Cohen, P. J., *Set Theory and the Continuum Hypothesis*, 1966, Benjamin, New York.

Cohn, P. M., *Universal Algebra* (2nd edn), 1981, Reidel, Dordrecht.

Hodges, W. A., Six impossible rings, *J. Algebra* 31 (1974), 218–44.

Kaplansky, I., *Set Theory and Metric Spaces*, 1972, Allyn and Bacon, Boston.

Mac Lane, S., *Categories for the Working Mathematician*, 1971, Springer, Berlin.

Rota, G.-C., On the foundations of combinatorial theory I. Möbius functions, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 2 (1964), 340–68.

Sierpiński, W., *Cardinal and Ordinal Numbers*, 1956, Pan. Wyd. Nauk. Warsaw.

Chapter 3

Artin, E., *Galois Theory*, Notre Dame Math. Lectures No. 2, 1948, Notre Dame, Indiana.

Chase, S. U., Harrison, D. K. and Rosenberg, A., *Galois Theory and Cohomology Theory of Commutative Rings*, Memoirs of the AMS No. 52, 1965, AMS, Providence, RI.

- Dedekind, R., *Über die Theorie der ganzen algebraischen Zahlen*, 1964, Vieweg, Braunschweig.
- Galois, E., *Oeuvres Mathématiques*, 1951, Gauthier-Villars, Paris.
- Ore, O., *Cardano the Gambling Scholar*, 1953, Princeton Univ. Press Princeton, NJ.
- Steinitz, E., Algebraische Theorie der Körper, *J. reine angew. Math.* **137** (1910), 167–309; reprinted 1930, Teubner, Leipzig 1950, Chelsea, New York.
- Witt, E., Über die Kommutativität endlicher Schiefkörper, *Hamb. Abh.* **8** (1931), 315–22.

Chapters 4–5

- Anderson, F. W. and Fuller, K. R., *Rings and Categories of Modules*, 1973, Springer, Berlin.
- Cohn, P. M., Some remarks on the invariant basis property, *Topology* **5** (1966), 215–228.
- Fuchs, L., *Abelian Groups I, II*, 1970, 1973, Academic Press, New York.
- Pierce, R. S., *Modules Over Commutative Regular Rings*, Memoirs of the AMS No. 70, 1967, AMS, Providence, RI.
- Rowen, L. H., *Ring Theory I, II*, 1988, Academic Press, New York.

Chapter 6

- Artin, E., *Geometric Algebra*, 1957, Interscience, New York.
- Lam, T. Y., *The Algebraic Theory of Quadratic Forms* (2nd printing, revised), 1980, Benjamin-Cummings, New York.
- Scharlau, W., *Quadratic and Hermitian Forms*, 1985, Springer, Berlin.

Chapter 7

- Curtis, C. W. and Reiner, I., *Methods of Representation Theory I, II*, 1981, 1987, John Wiley & Sons, New York.
- Feit, W., *The Representation Theory of Finite Groups*, 1982, North-Holland, Amsterdam.
- Huppert, B., *Endliche Gruppen I*, 1967, Springer, Berlin.
- Serre, J.-P., *Représentations Linéaires des Groupes Finis* (2nd edn), 1971, Hermann, Paris.
- Weyl, H., *The Classical Groups*, 1939, Princeton Univ. Press Princeton, NJ.

Chapters 8–9

- Endler, O., *Valuation Theory*, 1972, Springer, Berlin.
- Fossum, R. M., *The Divisor Class Group of a Krull Domain*, 1973, Springer, Berlin.
- Hilbert, D., Bericht über die Theorie der algebraischen Zahlkörper, *Jahrber. DMV* iv, 1897; reprinted in Vol. 1 of the collected works.
- Kaplansky, I., Projective modules, *Ann. Math.* **68** (1958), 372–7.
- Lang, S., *Algebraic Number Theory*, 1970, Addison-Wesley, Reading, MA.
- Mahler, K., *p-adic Numbers and their Functions*, Second Edn, 1981, Cambridge University Press, Cambridge.
- Matsumura, H., *Commutative Rings*, 1985, Cambridge Univ. Press Cambridge.
- Nagata, M., *Local Rings*, 1962, Interscience, New York.
- Rudin, W., *Real and Complex Analysis*, 1966, McGraw-Hill, New York.
- Steinitz, E., Rechteckige Systeme und Moduln in algebraischen Zahlkörpern I, II, *Math. Ann.* **71** (1911), 328–54, **72** (1912) 297–345.

Chapters 10–11

- Arbib, M. A., *Theories of Abstract Automata*, 1969, Prentice-Hall, Englewood Cliffs, NJ.
- Berstel, J., *Transductions and Context-Free Languages*, 1979, Teubner, Stuttgart.
- Berstel, J. and Perrin, D., *Theory of Codes*, 1985, Academic Press, New York.
- Berstel, J. and Reutenauer, C., *Les Séries Rationnelles et Leurs Langages*, 1984, Masson, Paris.
- Boë, J. M., de Luca, A. and Restivo, A., Minimal completable sets of words, *Theor. Comput. Sci.* **12** (1980), 325–32.
- Cohn, P. M., Algebraic language theory, *Bull. Lond. Math. Soc.* **7** (1975), 1–29.
- Conway, J. H. and Sloane, N. J. A., *Sphere Packings, Lattices and Groups*, 1988, Springer, Berlin.
- Davis, M., *Computability and Unsolvability*, 1958, McGraw-Hill, New York.
- Eilenberg, S., *Automata, Languages and Machines A–C*, 1974–78, Academic Press, New York.
- Herman, G. T. and Rozenberg, G., *Developmental Systems and Languages*, 1975, North-Holland, Amsterdam.
- Hill, R., *Introduction to Coding Theory*, 1985, Oxford University Press, Oxford.
- Lallement, G., *Semigroups and Combinatorial Applications*, 1979, John Wiley & Sons, New York.
- Landweber, P. S., Three theorems on phrase-structure grammars of type 1, *Inform. Control* **6** (1963), 131–6.
- van Lint, J. H., *Introduction to Coding Theory*, 1982, Springer, Berlin.
- Lothaire, M., *Combinatorics on Words*, Encycl. of Math. and its Applications Vol. 17, 1983, Addison-Wesley, Reading, MA.
- McEliece, R. J., *The Theory of Information and Coding*, Encycl. of Math. and its Applications Vol. 3, 1977, Addison-Wesley, Reading, MA.
- Shannon, C. E., A mathematical theory of communication, *Bell Syst. Tech. J.* **27** (1948), 379–423, 623–56; reprinted in Slepian, D. (ed.), *Key Papers in the Development of Information Theory*, 1974, IEEE Press, New York.

List of notations

The number indicates the page where the term is first used or defined. When no page number is given, the term is defined in Vol. 1.

Number systems

\mathbb{N}	the natural numbers
\mathbb{N}_0	the natural numbers with 0
\mathbb{Z}	the integers
\mathbb{Q}	the rational numbers
\mathbb{Q}_+	the non-negative rational numbers
\mathbb{R}	the real numbers
\mathbb{C}	the complex numbers
\mathbb{U}_m	the group of m th roots of unity 153
$\mathbb{Z}(p^\infty)$	the group of all p^n th roots of 1, for $n = 1, 2, \dots$
\mathbb{Z}/n	the integers mod n
$\mathbb{U}(n)$	the group of units mod n 95
\mathbb{F}_q	the field of q elements 98
\mathbb{Z}_p	the p -adic integers 272
$\mathbb{Q}_p = \mathbb{Z}_p[\,p^{-1}\,]$	the p -adic numbers 272

Set theory

\emptyset	the empty set 1
$ X $	the cardinal of the set X 2
$\mathcal{P}(X)$	the power set (set of all subsets) of X 7
$X \setminus Y$	the complement of Y in X
Y^X	the set of all mappings from X to Y 5
\aleph_0	aleph-null, the cardinal of \mathbb{N} 2

Number theory

$\max(a, b)$	the larger of a, b
$\min(a, b)$	the smaller of a, b

$a b$	a divides b 264, 305
(a, b)	highest common factor (HCF) of a and b 305
$[a, b]$	least common multiple (LCM) of a and b 305
$[x]$	greatest integer not exceeding x 359
δ_{ij}	Kronecker delta
$\mu(n)$	Möbius function 55
Φ_m	cyclotomic polynomial 94
$\varphi(m)$	Euler function 93

Group theory

Sym_n or S_n	the symmetric group of degree n
Alt_n or A_n	the alternating group of degree n
$\text{sgn } \sigma$	the sign of the permutation σ
C_n	the cyclic group of order n 153
D_m	the dihedral group of order $2m$ 239
G^+	the set of elements $\neq 1$ in G
G'	the derived group
$N \triangleleft G$	N is a normal subgroup of G
$\mathcal{N}(G)$	the set of normal subgroups of G
$(G:H)$	the index of H in G
$\text{GL}_n(R)$	the general linear group over a ring R 221
$\text{O}(V)$, $\text{SO}(V)$	the orthogonal (special orthogonal) group on an inner product space V 241
$\text{U}_n(\mathbb{C})$	the unitary group over \mathbb{C} 241
$\text{Aff}_n(k)$	the affine group over a field k

Rings and modules

${}^m V^n$	the space of all $m \times n$ matrices over V 136
${}^m V$	the space of m -component column vectors in V ($= {}^m V^1$) 136
V^n	the space of n -component row vectors in V ($= {}^1 V^n$) 136
$\mathfrak{M}_n(R)$ or R_n	the $n \times n$ matrix ring over R 136
$\text{Lat}(M)$	the lattice of all submodules of M 138
$\text{Hom}(U, V)$	the set of all homomorphisms from U to V 125
$\text{End}(U)$	the ring of all endomorphisms of U 124
$U \otimes V$	the tensor product of U and V 156
tM	the torsion submodule of M 331
R^0	the opposite ring 125
R^\times	the set of all non-zero elements in an integral domain
A^1	the augmented algebra 166
$\text{Ann}(X)$	the annihilator of X 129
$\text{Ass}(M)$	the assassinator of M 339

$\text{Supp}(M)$	the support of M 315
R_S or $R_{\mathfrak{p}}$	localization of R at S (or at the complement of \mathfrak{p}) 311f.
$\sqrt{\mathfrak{a}}$	the radical of an ideal \mathfrak{a} 309
$K[X]$	the polynomial ring on X over K 307
$K[\![X]\!]$	the formal power series ring on X over K 183
$K\langle X \rangle$	the free K -algebra on X 405
$K\langle\!\langle X \rangle\!\rangle$	the free power series algebra on X over K 405
$s\mathcal{M}_R$	the category of (S, R) -bimodules 127
A_E	the algebra obtained from A by extending the groundfield to E 185
X^*	the free monoid on X 380
$X^+ \setminus \{1\}$	$X^* \setminus \{1\}$ 385

Field theory

$[V:k]$ or $\dim_k V$	dimension of the k -space V 63
$k(\alpha)$	the field generated by α over k 65
$k[\alpha]$	the ring generated by α over k 65
$\text{Gal}(E/F)$	group of the Galois extension E/F 86
$T_{E/F}(x)$	trace of x 104
$N_{E/F}(x)$	norm of x 104
$U \perp V$	orthogonal sum of inner product spaces 200
U^\perp	orthogonal complement 201
$\langle a_1, \dots, a_n \rangle$	quadratic form (in diagonal form) 202

Index

- Abel, N. H., 62, 113
Absolute value, 212, 270
Absorptive law, 31
ACC (= ascending chain condition), 38
Acceptor, 388
Accessible, 389
Acquaintanceship graph, 22
Acyclic, 23
Additive category, functor, 146
Adjacent, adjacency matrix, 21, 29
Adjoint associativity, 157, 162
Affine group, 102, 117, 119, 263
Afford, 223
Aleph, 2
Algebra, 165
Algebraic, 65 f
Algebraic closure, 74 f
Algebraic integer, 168, 289
Algebraic power series, 412
Algebraic set, 336
Almost all, xiii
Alphabet, 356, 380, 384
Alternating form, 204
Anisotropic, 205, 210
Annihilator, 129
Anti-chain, 9
Antihomomorphism, 126
Approximation theorem, 274, 327
Archimedean absolute value, 270
Archimedean ordering, 217
(Archimedes, 287–212 BC)
Arrow, 23
Artin, E. (1898–1962), 41, 90, 176, 274,
 276
Artinian module, ring, 41
Artin's theorem, 84
Ascending chain condition, 38
Assassinator, 339
Associated prime ideals, $\text{Ass}(M)$, 339
Associativity, adjoint, 157
Atom, 52, 307, 382
Atomic Boolean algebra, 61
Atomic integral domain, 307
Augmentation ideal, map, 166
Augmented algebra, 166
Automaton, 388
Averaging lemma, 227
Axiom of choice, 10
Baer's criterion, 151
(Baer, R., 1902–79)
Balanced map, 161
Basis, 140
BCH-code, 373
Behaviour, 388
Bernstein, F. (1878–1956), 4
Berstel, J. (1941–), 406, 414
Bialternant, 252
Bicyclic monoid, 414
Bidual, 153
Bifix, 404
Bifunctor, 128
Bilinear form, 197
Bimodule, 125
Binary symmetric channel, 357
Bit, 357
Block code, 358
Boë, J. M., 403 f
Bolzano, B. (1781–1848), 2
Boolean algebra, ring, 45 ff
Bose, R. C., 372
Bound, bounded, 9
Box principle, 2
Burali-Forti paradox, 6
(Burali-Forti, C., 1861–1931)
Burnside's $p^\alpha q^\beta$ -theorem, 258 f
Cancellation law, 381

- Cantor, G., 2, 7
 Cantor's diagonal argument, 8
 Cardano, G. (1501–76), 113
 Cardinal, 1 f
 Cartan, E. (1869–1951), 206
Casus irreducibilis, 122
 Category, 16
 Cauchy sequence, 215, 271
 Centralizer, centre, 136, 165, 186
 CF-grammar, -language, 385
 Chain, 8
 Chain condition, 38
 Channel capacity, 358
 Character, 152, 233
 Character group, 153
 Characteristic of a field, ring, 62 f
 Characteristic function, 7, 49
 Characteristic series, 411
 Check polynomial, 370
 Chinese remainder theorem, 150, 171
 Chomsky, N. (1928–), 383
 Chomsky hierarchy, 385
 Chomsky normal form, 414
 Class, 17
 Clause-indicator, 384
 Clifford, W. K. (1845–79), 209
 Coaccessible, 389
 Code, 356, 396
 (see also under the particular type of code)
 Cofinite subset, 46
 Cohen's theorem, 354
 Cohen, I. S. (1917–55)
 Cohen, P. J. (1934–), 7, 11
 Coimage, cokernel, 126
 Comaximal, 171
 Comma category, 20
 Comparable, 8
 Complement(ary subgroup), 131, 260
 Complementary graph, 21
 Complete graph, 21
 Complete lattice, 33
 Complete ordered field, 215
 Complete valued field, 271
 Completely integrally closed, 303
 Completely primary ring, 174
 Completely reducible, 225
 Completion, 217, 272
 Composite of fields, 192
 Concrete category, 18
 Conductor, 350
 Cone, 213
 Congruent matrices, 198
 Conical monoid, 381
 Conjugate, 104, 252
 Conjunctive normal form, 49
 Connected graph, 24
 Conorm, 328
 Consistent system, 335
 Context-free, sensitive, 385
 Continuum hypothesis, 7
 Contracted ideal, 313
 Contragredient, 226
 Contravariant, 17
 Converge, 215
 Conway, J. H. (1932–), 377
 Coordinate ring, 335
 Coproduct, 128
 Coset leader, 363
 Countable, 2
 Covariant, 17
 Covering, 360
 CS-grammar, -language, 385
 Cubical norm, 275
 Cycle, 23, 247
 Cyclic code, 369
 Cyclic extension, 110
 Cyclotomic polynomial, 94
 DCC (= descending chain condition), 38
 Decomposition lemma, 43, 318, 322
 Dedekind, J. W. R. (1831–1916), 2, 33,
 80, 84, 90, 320, 354
 Dedekind domain, 150, 322
 Dedekind's lemma, 81
 Degree, 63, 66, 222, 405
 Degree of field extension, 63
 Delian problem, 67
 De Morgan laws, 45
 (De Morgan, A., 1806–71)
 Dense, 19, 214, 216
 Density property, 352
 Derivation of a sentence, 384
 Derivative of a polynomial, 78
 Determinant of inner product space, 200
 Diamond lattice, 36
 Dieudonné, J. (1906–), 206
 Difference operator, 286
 Different, 354
 Digraph, 23
 Dilworth's theorem, 23
 (Dilworth, R. P., 1914–)

- Dimension of a ring, 345
 Dimension of a module, 177
 Diophantos (*ca.* 250 AD), 113
 Direct product of rings, modules, 128, 168
 Direct sum, power, 128
 Directed graph, set, 23
 Dirichlet, P. G. L., 96, 327
 Dirichlet box principle, 2
 Discriminant, 106
 Disjunctive normal form, 48
 Distributive lattice, law, 36
 Distributive module, 135
 Divisible, 151
 Division algebra, ring, 167
 Dominate, 290
 Dual basis, 149
 Dual code, 362
 Dual homomorphism, 48
 Dual of a category, 19
 Duality, 8, 48, 153
 Duplication of the cube, 67
 Dyck code, 404

 Edge, 21, 388
 Einseinheiten, 284
 Eisenstein equation, 303
 Embedded component, 343
 Empty set, 1
 Endpoint, 21
 Entropy, 358
 Enumerable, 2
 Equipotent, 1
 Equivalence of absolute values, 272
 Equivalence of categories, 19
 Equivalence of representations, 222
 Equivalence of valuations, 268
 Erdős, P. (1913–), 27
 Error-correcting and -detecting, 359
 Essential extension, 163
 Euclid, 67
 Euclidean domain, 320
 Euler, L., 22, 117
 Euler function, 58, 93
 Eulerian graph, 28
 Euler's criterion, 123
 Exact functor, 146
 Expanded ideal, 312
 Exponent, 93
 Extension, of fields, 63
 Extension theorem of valuations, 293

 Faithful functor, 19
 Faithful module, representation, 168, 222
 Faltings, G. (1954–), 319
 Fermat primes, 117
 Fermat's last theorem, 319 f
 Ferrari, L. (1522–65), 113
 Ferro, S. del (1465–1526), 113
 Field of sets, 46
 Final object, 20
 Final vertex, 23
 Finite, 1, 292
 Finite character, 13
 Finite field, 97
 Finite (field) extension, 63
 Finite state language, 385
 Finitely presented, related, 142
 Fir (= free ideal ring), 406
 Fixed field, 85
 Flat module, 315
 Forgetful functor, 18
 Fractional ideal, 321
 Fractions, 310
 Free algebra, 405
 Free ideal ring, 406
 Free module, 140
 Free monoid, 380
 Frobenius, G., 229, 247, 259 f
 Frobenius endomorphism, 77
 Frobenius kernel, subgroup, 260
 Frobenius reciprocity formula, 255
 Frobenius' theorem, 260
 Full functor, 19
 Full matrix, 361
 Full subcategory, 17
 Full subgraph, 21
 Fully invariant, 133
 Function ring, 335
 Functionally complete, 50, 100
 Functor, 17
 Fundamental theorem of algebra, 70
 Fundamental theorem of Galois theory, 86

 Galois, E. (1811–32), 62, 80, 90
 Galois connexion, 86 f, 336 f
 Galois' criterion, 114
 Galois extension, group, 86 ff
 Gaussian extension of a valuation, 296
 Gaussian integer, 102
 Gaussian sum, 123

- Gauss's lemma, 91, 316
 General linear group, 221
 Generating set, 124, 380
 Generator matrix, 361
 Generic point, 338
 Gilbert–Varshamov bound, 360
 Gödel, K. (1906–78), 7, 11
 Going-up theorem, 347
 Golay code, 377
 Goodearl, K. R., 140
 Goppa code, 374
 Grammar, 384
 Graph, 21 ff, 135
 Greatest, 9
 Gregory, J. (1638–75), 113
 Greibach normal form, 387
 Ground field, 63
 Group algebra, 167
 Group code, 404
 Group of an equation, 107

 Hall subgroup, 262
 Hall's theorem, 23
 (Hall, P., 1904–82)
 Halmos, P. R. (1916–), 24
 Hamming bound, 360
 Hamming code, 365 f
 Hamming distance, 358
 (Hamming, R. W., 1915–)
 Hankel matrix, determinant, 413
 (Hankel, H., 1839–73)
 Hensel, K. (1861–1941), 280
 Henselization, 299
 Hensel's lemma, 298 f
 Hereditary ring, 322, 329
 Hilbert basis theorem, 318
 Hilbert Nullstellensatz, 351 f
 (Hilbert, D., 1862–1943)
 Hocquenghem, A., 373
 Hodges, W. A. (1941–), 139
 Homogeneous, 405
 Hopkins, C. (1902–39), 176
 Hopkins' theorem, 182
 Huntington, E. V. (1874–1952), 60
 Hyperbolic pair, plane, 209

 IBN (=invariant basis number), 143
 Ideal (in a lattice), 59
 Ideal class group, 327
 Ideal numbers, 305, 320
 Idempotent, 46, 172

 Identity morphism, 16
 Image, 126
 Incidence algebra, 54
 Inclusion–exclusion, 57
 Indecomposable ring, 172
 Independence property of tensor product, 159
 Index of a quadratic form, 210
 Indicator of a character, 244
 Induced module, representation, 253 f
 Induction, transfinite, 14
 Inductive ordered set, 11
 Inert extension, 91 f
 Inf (=infimum), 30
 Infinite, 2
 Information rate, 357
 Initial object, 20
 Initial vertex, 23
 Injective module, 147
 Inner product space, 197
 Input, 387
 Integers, 289 f
 Integral closure, 290
 Integral extension, 345
 Integral ideal, 321
 Intermediate value property, 219
 Intersection graph, 29
 Intertwining number, 235
 Interval, 34
 Invariant basis number, 143
 Invertible ideal, 321
 Inverting, 293, 311
 Involution, 205, 225
 Irreducible, 43, 224, 337
 Irredundant, 181, 342
 ISBN book number, 378
 Isolated component, 343
 Isometry, 198
 Isomorphism of categories, 19
 Isomorphism of field extensions, 82
 Isomorphism theorems, 126
 Isotropic, 205
 Isotypic, 133
 Iteration lemma, 414

 Jacobson, N. (1910–), 179
 Jacobson radical, 180
 Join, 30
 Join-irreducible, 43

 Kaplansky, I. (1917–), 334

- Kernel, 126, 232
 Kinds of complex representation, 243
 König's lemma (König, D.), 26
 König, J. (1849–1913), 5
 Königsberg bridge problem, 22
 Kraft–McMillan inequality, 400
 Krasner's lemma (Krasner, M.), 303
 Kronecker, L., 69, 95, 121
 Krull dimension, 345
 Krull intersection theorem, 354
 Krull's theorem, 43, 307 f
 (Krull, W., 1899–1971)
 Kummer, E. (1810–93), 305, 319 f, 327
- Lagrange resolvent, 111
 Language, 384
 Lasker, E. (1868–1941), 338
 Lattice, 30
 Laurent polynomial, series, 167
 (Laurent, P. A., 1803–54)
 Law of quadratic reciprocity, 123
 Least, 9
 Leech lattice (Leech, J.), 377
 Legendre symbol, 123
 Length of a chain, lattice, 40
 Lewin, J. (1940–), 406
 Limit ordinal, 14
 Lindemann, C. L. F. (1852–1939), 67
 Linear character, 233
 Linear code, 361
 Linearly dependent, 140
 Linearly disjoint, 185
 Local ring, 312, 408
 Localization, 312
 Locally finite, 54
 Locally nilpotent, 340
 Loop, 21
 Lower bound, 9
 Lower segment, 9
 Luca, A. De, 403 f
- Machine, 387
 Mackey, G. W. (1916–), 256
 McMillan, B., 399 f
 MacWilliams identity (MacWilliams, F. J.), 367
 Marriage theorem, 24
 Maschke's theorem, 196, 226, 228, 242
 (Maschke, H., 1853–1908)
 Matrix ring, 135 ff
 Matrix units, 136, 185
- Maximal, minimal, 9
 Maximal right ideal, 42
 Maximum condition, 38
 Maxterm, 48
 Measure, 398
 Meet, 30
 Meet-irreducible, 43, 341
 Minimal generating set, 140
 Minimal polynomial, 66
 Minimum condition, 39
 Minterm, 48
 Möbius function, inversion, 55 f
 (Möbius, A. F., 1790–1868)
 Modular law, lattice, 33
 Module, 124 ff, 222
 Molien's theorem, 241, 378
 (Molien, T., 1861–1941)
 Monoid, 379
 Moore, E. H. (1862–1932), 98
 Morita equivalence (Morita, K.), 139, 169
 Morphism, 16
 Multiplication, 183
 Multiplicative set, 293, 308
 Multiplicative system of representatives, 284
- Nagata, M. (1927–), 353
 Nagata's theorem, 315
 Nakayama's lemma, 182, 354
 (Nakayama, T., 1912–64)
 Natural irrationality, 109
 Natural isomorphism, transformation, 18
 Negative, 212
 Neumann, J. von (1903–57), 247
 Newton–Fourier rule, 282 f
 Nilideal, 181
 Nilpotent, 178
 Nilradical, 309
 Noether, E. (1882–1935), 176, 322 f, 339
 Noether normalization lemma, 350
 Noetherian induction, 39
 Noetherian module, ring, 41, 317
 Non-singular quadratic form, 199
 Norm, 104, 187
 Normal closure, 73
 Normal equation, 89
 Normal extension, 72
 Normalized valuation, 266
 Normed vector space, 275
 Null sequence, 215, 271

- Nullstellensatz (= solution set theorem), 351 f
- Object, 16
- Opposite category, 17
- Opposite ring, 125
- Optimal code, 359
- Order isomorphism, type, 10, 214
- Order of a power series, 405
- Ordered field, ring, 212
- Ordering, 8
- Ordinal number, 12
- Orthogonal basis, 202
- Orthogonal complement, 201
- Orthogonal group, transformation, 204 f., 241
- Orthogonal idempotents, 172
- Orthogonality relation, 230
- Ostrowski's first and second theorems, 276, 278
- (Ostrowski, A., 1893–1986)
- Output, 387
- p -adic integer, valuation, 268, 272
- Packing, 360
- Parallelogram law, 126
- Parity check, 357, 362
- Partial order, 8
- Path, 22, 388
- PDA (= pushdown acceptor), 393
- Pentagon lattice, 35 f
- Perfect code, 360
- Perfect field, 77
- Perlis, S. (1913–), 179
- Permutation module, 238
- Phrase-structure grammar, 384 f
- PID (= principal ideal domain), 330 f., 406
- Pierce, R. S. (1927–), 173
- Place, 292
- Plotkin bound (Plotkin, M.), 366, 378
- Point, 21
- Positive cone, 213
- Power of the continuum, 7
- Power series, 183, 404
- Power set, 7
- Prefix set, code, 396
- Preordering, 8
- Primary decomposition, 341
- Primary submodule, ideal, 332, 340
- Prime, 91, 267, 307
- Prime ideal, 192, 307
- Prime subfield, 63
- Primitive element, 98, 102 f
- Primitive idempotent, 172
- Primitive polynomial, 91, 281
- Primitive root of unity, 93
- Principal character, 233
- Principal ideal, 305
- Principal valuation (ring), 265, 268
- Principle of domination, 279
- Product formula, 64
- Projective coordinate system, 149
- Projective-free, 169
- Projective module, 147
- Proper orthogonal, 204
- Pullback, 127, 314
- Pumping lemma, 393
- Puncturing, 361
- QR-code, 375
- Quadratic extension, 88
- Quadratic form, 197
- Quadratic residue, 123
- Quadrature of the circle, 67
- Quasi-inverse, 180
- Quiver, 23
- Rabinowitsch trick, 352
- Radical extension, 111
- Radical of an ideal, 309
- Radical of an inner product space, 199
- Radical of a ring, 171 f
- Ramification index, 294, 328
- Ramsey number, theorem, 25 f
(Ramsey, F. P., 1903–30)
- Rank of a free module, 142
- Rank of a quadratic form, 202
- Rational power series, 408
- Ray-Chaudhuri, D. K., 373
- Reciprocal equation, 122
- Reduced acceptor, 390
- Reduced ring, 173
- Reducible algebraic set, 337
- Reducible representation, 224
- Reflexion, 205
- Regular language, 385
- Regular quadratic form, 199
- Regular representation, 168, 381
- Represent, 202
- Representation, 168, 221
- Residue class field, 267, 312, 408

- Residue degree, 294
 Resolvent, 111 f, 121
 Restivo, A., 403 f
 Retraction, 132
 Reutenaer, C., 406
 Rewriting rules, 384
 Rigid monoid, 381
 Root, 66
 Root tower, 113
 Rota, G.-C. (1932–), 54
 Rotation, 204
 Ruffini, P. (1765–1822), 113
 Ruler-and-compass construction, 67, 117
 Russell, B. A. W., 11

 S-function, 252
 Sarges, H., 318
 Saturated multiplicative set, 308
 Schreier–Lewin formula, 408
 Schreier refinement theorem, 34, 41
 Schreier set, 397, 406
 Schröder–Bernstein theorem, 4, 58
 (Schröder, E., 1841–1902)
 Schubfachprinzip (= box principle), 2
 Schur, I., 94, 176, 229
 Schur’s lemma, 174, 176, 229
 Schützenberger, M.-P. (1923–), 402,
 409, 411 f
 Section, 132
 Seidenberg, A., 348
 Self-embedding, grammar, 414
 Semi-Artinian module, 163
 Semimodular (upper, lower), 59
 Semisimple module, 130, 225
 Semisimple ring, 174
 Sentence, 384
 Separable extension, polynomial, 79
 Separable states, 390
 Sequential machine, 386
 Shannon, C. E. (1916–), 356 ff
 Sheffer, H. M. (1883–1964), 60
 Shortening, 361
 Sierpiński, W. (1882–1969), 13, 27
 Sieve principle, 57
 Simple field extension, 65, 102
 Simple module, 130
 Simple ring, 174
 Singleton bound, 361
 Singular quadratic form, 199
 Skeleton, 20
 Skew field, 167

 Small category, 17
 Socle, 133
 Solubility criterion, 114, 119
 Solution, 335
 Source, 16
 Special orthogonal group, 204
 Sperner’s lemma, 28
 (Sperner, E., 1903–80)
 Sphere packing bound, 360
 Split exact sequence, 132
 Split inner product space, 209
 Splitting field, 71, 74
 Standard array, 364
 State, 387 ff
 Steinitz, E. (1871–1928), 103, 332
 Stochastic matrix, algebra, 196
 Stone, M. H. (1903–), 53
 Strong approximation theorem, 327
 Structure constants, 167
 Successful path, 388
 Suffix, 396
 Sup (= supremum), 30
 Support of a module, 315
 Support of a power series, 405
 Symbolic prime power, 344
 Symmetric difference, 47
 Symmetrizer, 250
 Symmetry, 206
 Syndrome decoding, 363
 Szekeres, G., 27

 Target, 16
 Tartaglia, N. (1499–1557), 113
 Tensor product, 156
 Terminal letter, 384
 Thompson, J. G., 258, 262
 Torsion-free, 152, 331
 Torsion (sub)module, 331
 Totally ordered, 9
 Tower of fields, 64
 Trace, 104, 187
 Transcendental, 65 f
 Transfinite, 2
 Transfinite induction, 14
 Transition function, 387
 Transitivity formula, 105, 189
 Tree, 24
 Triangle inequality, 212, 270, 358
 Triangular matrix ring, 137, 167
 Trim automaton, 389
 Trisection of an angle, 68

- Trivial, 46, 222, 233, 265
 Turing machine, 379, 394
 (Turing, A. M., 1912–54)
 Type component, 133
 Type of a language, 385
 Type of a module, 133, 334
- UFD (= unique factorization domain),
 91, 307, 315 ff, 326, 406
 UGN (= unbounded generating
 number), 145
 Ultrametric inequality, 270
 Unambiguous, 397
 Uniform measure, 398
 Uniformizer, 267
 Unipotent, 226
 Unit, 382
 Unitary group, representation, 241
 Universal mapping property, 20, 61
 Universal quadratic form, 203
 Upper bound, 9
 Upper bound property, 218
- Valency, 22
 Valuation (ring), 265
 Valued ring, field, 270
 Vandermonde determinant, matrix,
 107, 113, 372
 Variety, 337
 Vertex, 21
 Viète, F. (1540–1603), 62
- Weakly finite ring, 143
- Wedderburn, J. H. M. (1882–1948), 176,
 190
 Wedderburn theorem (finite fields), 100 f
 Wedderburn theorem (semisimple rings),
 174 f
 Weight, 361
 Weight enumerator, 367
 Well-ordered, 9
 Well-ordering theorem, 12
 Whaples, G. (1914–81), 274
 Witt, E. (1911–), 100, 208
 Witt cancellation theorem, 208
 Witt decomposition, index, 210
 Witt extension theorem, 211
 Witt group, ring, 211
 Word, 384
- Yoneda's lemma (Yoneda, N.), 28
 Young, A. (1873–1940), 247
 Young diagram, tableau, 248
 Young symmetrizer, 250
- Zariski topology, 338, 352
 (Zariski, O., 1899–1986)
 Zech logarithm, 99
 Zermelo, E., 10
 Zero, 66
 Zero-delay code, 397
 Zerodivisor, 339
 Zeta function, matrix, 55, 58
 Zorn's lemma, 11
 (Zorn, M., 1906–)